Genet. Res., Camb. (1990), **55**, *pp.* 199–221 *With 5 text-figures Printed in Great Britain*

Mutation-selection balance and the evolutionary advantage of sex and recombination

BRIAN CHARLESWORTH
Department of Ecology and Evolution, The University of Chicago, 1103 E. 57th St, Chicago, IL 60637, USA

(*Received 21 November 1989 and in revised form 14 February 1990*)

Summary

Mutation-selection balance in a multi-locus system is investigated theoretically, using a modification of Bulmer's infinitesimal model of selection on a normally-distributed quantitative character, taking the number of mutations per individual (n) to represent the character value. The logarithm of the fitness of an individual with n mutations is assumed to be a quadratic, decreasing function of n. The equilibrium properties of infinitely large asexual populations, random-mating populations lacking genetic recombination, and random-mating populations with arbitrary recombination frequencies are investigated. With 'synergistic' epistasis on the scale of log fitness, such that log fitness declines more steeply as n increases, it is shown that equilibrium mean fitness is least for asexual populations. In sexual populations, mean fitness increases with the number of chromosomes and with the map length per chromosome. With 'diminishing returns' epistasis, such that log fitness declines less steeply as n increases, mean fitness behaves in the opposite way. Selection on asexual variants and genes affecting the rate of genetic recombination in random-mating populations was also studied. With synergistic epistasis, zero recombination always appears to be disfavoured, but free recombination is disfavoured when the mutation rate per genome is sufficiently small, leading to evolutionary stability of maps of intermediate length. With synergistic epistasis, an asexual mutant is unlikely to invade a sexual population if the mutation rate per diploid genome greatly exceeds unity. Recombination is selectively disadvantageous when there is diminishing returns epistasis. These results are compared with the results of previous theoretical studies of this problem, and with experimental data.

1. Introduction

The problem of the evolutionary significance of sexual reproduction and genetic recombination has recently attracted a great deal of attention. Under many biological circumstances, an asexual lineage is expected to have a reproductive advantage over an asexual random-mating, ancestral population, due to the 'cost of sex' (Lloyd, 1980; Maynard Smith, 1978). Many theories of possible countervailing advantages of sex have been devised, and tests attempted using both experimental and comparative methods (Charlesworth, 1989; Michod & Levin, 1988; Stearns, 1987). One of the most attractive theories depends on the generation of genetic variation by mutation to deleterious alleles at a large number of loci scattered throughout the genome (Kondrashov, 1988). If each locus affects fitness independently of the others (i.e. fitnesses are multiplicative across loci), then the mean fitness of an infinite population at equilibrium between selection and mutation is known to be independent of the presence or absence of sexual reproduction, being equal to $\exp -U$, where U is the per genome rate of mutation to deleterious mutations (Crow, 1970; Kimura & Maruyama, 1966). It appears likely, but has never been formally proved except in the case of two loci (Feldman *et al.* 1980), that under this mode of selection a modifier of the rate of genetic recombination in a sexual species will be neutral.

If, however, fitness effects deviate from multiplicativity, then equilibrium mean fitness depends on the mode of reproduction. From the study of special cases, it seems that 'synergistic' epistasis, such that log fitness falls off faster than linearly with the number of deleterious alleles carried by an individual, leads to the mean fitness of a sexual population exceeding $\exp -U$, whereas the mean fitness of an asexual population remains the same (Crow, 1970; Kimura & Maruyama, 1966; Kondrashov, 1982). Truncation selection is a special form of this type of selection, and has been much studied in this context (Crow & Kimura, 1979; Kondrashov, 1982, 1984, 1985, 1988).

The converse appears to be true when log fitness falls off more slowly than linearly with the number of deleterious alleles carried by an individual ('diminishing returns' epistasis). From numerical studies, it seems that a genetic modifier increasing the rate of recombination will often be favoured in a random-mating population under the synergistic model (Kondrashov, 1984). Analytic results on this have only been obtained for the case of a pair of loci (Feldman *et al.* 1980). The consequences of mutation-selection balance with synergistic epistasis have also been explored in relation to the evolution of transformation in bacteria (Redfield, 1988) and divided genomes in RNA viruses (Chao, 1988; Nee, 1989; Nee & Maynard Smith, 1990).

Since mutation to deleterious alleles is a universal phenomenon, and the net rates per genome of such mutations are known to be substantial in higher organisms (Simmons & Crow, 1977), it is important to have a clear understanding of the consequences of this process for the evolution of breeding systems (Kondrashov, 1988). In addition, evidence for the existence of synergistic effects of mutant genes on fitness components (Crow, 1970; Simmons & Crow, 1977) provides grounds for believing that an advantage to sexual reproduction and genetic recombination arises from the maintenance of deleterious alleles by mutation.

The purpose of the present paper is to study both population mean fitness, and selection on modifiers of the rate of recombination, under mutation-selection balance in a multi-locus system. Quantitative genetic theory for an infinite number of loci (Bulmer, 1980) is used, treating the number of deleterious mutations carried by individuals as a normally distributed, metrical trait. In order to preserve normality of the distribution after selection, the special case when log fitness is a quadratic function of the number of mutations is studied. The adequacy of the normal approximation is tested by comparing population properties calculated by the approximate method with those obtained for some limiting cases by exact recursion relations. It is found that synergistic selection can indeed confer a large advantage of sexual versus asexual reproduction at the level of population mean fitness, if the per genome mutation rate is sufficiently large. However, a large component of this arises from the effects of segregation, not recombination. Accordingly, selection on modifiers of recombination is generally weak, especially when there are several chromosomes per genome, although selection always favours non-zero rates of recombination under biologically reasonable conditions.

2. Properties of equilibrium populations under mutation-selection balance

(i) *Assumptions of the model*

An infinitely large population of diploid individuals is assumed. The cases of strictly asexual reproduction, and of random mating sexual diploid individuals, are considered. Recurrent mutation to deleterious mutations at a large number of loci is assumed to occur at a rate U per diploid genome per generation ($U = 2m\mu$, where m is the total number of mutable loci and μ is the mutation rate per locus from wild-type to mutant allele). Mutations are assumed to be partially dominant and present at low frequency in the population, so that their elimination exclusively involves heterozygotes. It is assumed that the number of new mutant genes present in a new diploid individual follows a Poisson distribution with mean U.

The fitness of an individual carrying n heterozygous mutations is assumed to be given by the function

$$w(n) = \exp -(\alpha n + \tfrac{1}{2}\beta n^2). \quad (1)$$

This is related to the quadratic fitness function studied by Kimura & Maruyama (1966) and Crow (1970), but has the advantage that normal distribution theory can be used in the development of the model (see below). Since the case of deleterious mutations is being considered, α is assumed to be positive. Positive β corresponds to synergistic interactions between mutations at different loci on the scale of log fitness; $\beta = 0$ corresponds to the case of multiplicative fitnesses; negative β corresponds to 'diminishing returns' epistasis (cf. Kimura & Maruyama, 1966). This fitness expression is, however, only realistic for negative β values when $n < -\alpha/\beta$, since fitness starts to increase with n above this threshold. All numerical of studies of diminishing returns epistasis were conducted with parameter values such that nearly all individuals in the population fell below the threshold. The exact numerical results reported below (Tables 1–3) were little affected by replacing eqn (1) with a function in which fitness is independent of n for $n > -\alpha/\beta$.

These selection parameters can be related as follows to empirical estimates obtained by Crow (1970) from data of Mukai (1969) on the decline in viability of homozygous second chromosomes of *Drosophila melanogaster*, due to the accumulation of deleterious mutations. Crow fitted a quadratic model relating egg-to-adult viability to the number of homozygous mutations, such that a and b are the linear and quadratic coefficients respectively. Both coefficients were found to be close to 0·01, for the detrimental mutations of minor effect constituting the bulk of newly arising mutational variation. With such weak effects, the logarithm of fitness behaves similarly to fitness, unless the number of mutations is high. If the coefficient of dominance is h, on the assumption that heterozygous effects on fitness follow the same curve as homozygous ones, we can thus write $\alpha = ha$ and $\beta = 2h^2b$ (Crow, 1970). Since there is strong evidence that most detrimental mutations are partially rather than completely recessive (Simmons & Crow, 1977), h

https://doi.org/10.1017/S0016672300025532 Published online by Cambridge University Press

can reasonably be assumed to be substantially greater than zero. Many of the numerical studies described below assume $h = 0{\cdot}2$. With $a = b = 0{\cdot}01$, this yields values of α and β of 2×10^{-3} and 8×10^{-4} respectively. These will be referred to as the standard values of the selection parameters in what follows.

The approximate analytic treatment of both the asexual and sexual cases assumes that the total number of mutations per individual at the start of a generation follows a normal distribution $\phi(n)$, with mean $\bar{n}$ and variance V. As a result of selection according to eqn (1), the distribution within this generation after selection is again normal, with mean $\bar{n}^*$ and variance V^*. Expressions for the changes in mean and variance are given in the Appendix, as is the mean fitness of the population, $\bar{w}$. The effect of mutation is to change the mean and variance by U.

The variance of n at the start of a generation can be partitioned into the genic (additive) variance V_A, which is the sum of the variances contributed by each locus, and a residual term. In a random-mating population, the residual term $C_L = V - V_A$ is attributable to linkage disequilibrium (Bulmer, 1980, p. 158). Given the assumption of low frequencies of mutant alleles at each locus, copy number would obey a Poisson distribution in the absence of linkage disequilibrium, and so V_A in a given generation is equal to the mean copy number $\bar{n}$. The effects of selection and reproduction on the transition between generations depends on the mode of reproduction, and so the cases of sexual and asexual reproduction must be considered separately.

(ii) *Asexual populations*

Asexual reproduction is assumed here to mean that offspring genotypes are identical to those of their parents. The exact recurrence relations for this case for a general fitness function are given by Kimura & Maruyama [1966, eqn (3.1)]. They are intractable as far as the equilibrium solution for the distribution of the numbers of mutations is concerned, except when fitnesses are multiplicative. In this case, a Poisson distribution of number of mutations per individual with mean and variance U/α is maintained at equilibrium (Haigh, 1978). It is also known that the mean fitness of an equilibrium population is $\exp - U$, regardless of the form of the fitness function (Kimura & Maruyama, 1966).

Using the normal approximation, an equilibrium solution under the fitness function of eqn (1) can be derived as follows. The change in mean copy number per generation is

$$\Delta\bar{n} = -\frac{V(\alpha+\beta\bar{n})}{1+\beta V} + U. \tag{2}$$

Equating this to zero, we obtain the equilibrium relation

$$\bar{n} = \frac{U - V(\alpha - \beta U)}{\beta V}. \tag{3}$$

Using equation (A 2) of the Appendix, and the fact that $\ln \bar{w} = -U$, we also have

$$2U = \ln(1+\beta V) - \frac{1}{1+\beta V}\{\alpha^2 V - 2\alpha\bar{n} - \beta\bar{n}^2\}. \tag{4}$$

These equations allow the equilibrium values of the mean and variance in number of mutations per individual to be calculated, substituting the expression for the mean from eqn (3) into eqn (4), and solving the resulting equation by Newton–Raphson iteration. The accuracy of the normal approximation can be checked by comparing the results of these calculations

Table 1. *Equilibrium properties of asexual populations*

	Exact results					Approximate results		
	(*a*) Effect of mutation rate [$a = 0{\cdot}01$, $b = 0{\cdot}01$, $h = 0{\cdot}2$ ($\alpha = 0{\cdot}002$, $\beta = 0{\cdot}0008$)]							
U	$\bar{w}$	$\bar{n}$	V	Skewness	Kurtosis	$\bar{w}$	$\bar{n}$	V
5·0	0·007	111	56·0	0·046	2·992	0·007	112	57·3
2·0	0·136	69·0	36·4	0·058	3·007	0·135	69·0	36·0
1·5	0·224	59·3	31·4	0·091	3·009	0·223	59·3	31·1
1·0	0·369	47·8	25·5	0·101	3·010	0·368	47·8	25·3
0·5	0·608	32·9	17·9	0·120	3·015	0·607	32·9	17·9
0·1	0·906	13·2	7·88	0·184	3·032	0·905	13·3	7·96
	(*b*) Effect of deviations from multiplicativity [$U = 1{\cdot}0$, $a = 0{\cdot}1$ ($\alpha = 0{\cdot}02$)]							
b	$\bar{w}$	$\bar{n}$	V	Skewness	Kurtosis	$\bar{w}$	$\bar{n}$	V
−0·001 25	0·368	59·3	71·7	0·178	3·046	0·368	59·3	70·6
−0·001	0·368	57·0	65·3	0·165	3·037	0·368	57·0	64·4
−0·000 5	0·368	53·4	56·5	0·150	3·025	0·368	53·3	55·8
0	0·368	50·5	50·5	0·141	3·020	0·368	50·0	50·0
0·01	0·368	31·2	22·8	0·122	2·994	0·368	31·2	22·6
0·05	0·368	18·2	11·4	0·162	3·027	0·368	18·2	11·3
0·1	0·368	13·8	8·32	0·190	3·036	0·368	13·8	8·19

https://doi.org/10.1017/S0016672300025532 Published online by Cambridge University Press

with the exact equilibrium solutions obtained by iteration of eqn (3.1) of Kimura & Maruyama (1966). Some examples are shown in Table 1, where the skewness and kurtosis of the distributions for the exact solution are also shown, as a check on the validity of the normal approximation. Only a narrow range of negative b values is shown, as no equilibrium exists when $b \ll 0$ [see section 2(iv) below]. It will be seen that, providing selection is fairly weak, there is remarkably good agreement between the exact and approximate solutions even when the mutation rate is so small that there is an appreciable departure from normality for the distribution of numbers of new mutations. The mean fitness of the population depends only on the mutation rate, as expected, despite the fact that the equilibrium means and variances of the number of mutations per individual are strongly affected by the selection parameters.

(iii) *Sexual populations with no recombination*

In this situation, the haploid genome is composed of a single chromosome with no recombination. Segregation takes place in diploid individuals, in the absence of any recombination. New diploid individuals are formed by random union of gametes, such that the frequency of individuals carrying i mutations from one parent and j from the other is $x_i x_j$, where x_i represents the frequencies of gametes carrying i mutations. The exact recurrence relations for the gamete frequencies are as follows (Kimura & Maruyama, 1966):

$$x_i^* = \frac{x_i}{\bar{w}} \sum_{j=0}^{m} x_j w_{ij}, \tag{5a}$$

$$x_i' = (\exp -\tfrac{1}{2}U) \sum_{j=0}^{i} \frac{(\frac{1}{2}U)^{(i-j)} x_j^*}{(i-j)!}, \tag{5b}$$

where w_{ij} is the fitness of individuals with i and j mutations, and x_i^* and x_i' are the gamete frequencies following selection and mutation respectively.

A normal approximation to the equilibrium predicted by these equations can be obtained by using the recurrence relation from eqn (5b) for $i = 0$ to write

$$\bar{w} = (\exp -\tfrac{1}{2}U) \sum_{j=0}^{m} x_j w_{j0}. \tag{6}$$

With the fitness model used above, we have $w_{j0} = \exp -j(\alpha + \frac{1}{2}\beta j)$. Hence, the sum in eqn (6) can be approximated by the integral of the product of w_{x0} with the probability density of the normal variate x, drawn from a distribution with mean $\frac{1}{2}\bar{n}$ and variance $\frac{1}{2}V$, the mean and variance for the gametes. Using eqn (A 2) for the mean fitness in equation (6), we obtain the following equation, analogous to eqn (4)

$$U = \ln \frac{1+\beta V}{1+\frac{1}{2}\beta V} + \{\alpha^2 V - 2\alpha\bar{n} - \beta\bar{n}^2\} \left\{ \frac{1}{2(1+\frac{1}{2}\beta V)} - \frac{1}{(1+\beta V)} \right\}. \tag{7}$$

This can be combined with the expression for $\bar{n}$ given by eqn (3), and solved numerically.

Some examples are shown under 'Approximate results from eqns (6) and (7)' in Table 2, together with the results of the exact equilibrium solutions given by eqn (5). It will be seen that the normal approximation is very accurate, when selection is sufficiently weak. Comparisons with Table 1 indicate that the mean fitnesses of the equilibrium populations are considerably larger in this case than in the corresponding asexual cases (see discussion in section 4). In contrast to the asexual case, the population mean fitness is affected by the selection parameters as well as the mutation rate. It can be seen that, for fixed a, mean fitness increases as b increases. Conversely, mean fitness seems to decrease with a if b is fixed. A small positive value of b is sufficient to produce a large increase in equilibrium mean fitness, provided that the mutation rate is sufficiently high, e.g. with $U = 1{\cdot}5$ and $a = 0{\cdot}01$ ($\alpha = 0{\cdot}002$) there is a 58% increase in mean fitness between $b = 0$ and $b = 0{\cdot}01$ ($\beta = 0{\cdot}0008$).

(iv) *Sexual populations with arbitrary recombination frequencies*

In the case of sexual populations, the equilibrium variance can be found by a modification of the approach developed by Bulmer for a quantitative character controlled by a large number of genes (1980, pp. 158–159). Following his argument, the covariance in the number of mutant alleles present at a given pair of loci i and j among the haploid gametes is written as C_{ij}. This reflects the effect of linkage disequilibrium between this pair of loci, for which the frequency of recombination is denoted by r_{ij}. Since a diploid individual is formed by a pair of gametes, the total variance V at the start of a generation is equal to the genic variance V_A, plus $C_L = 4\sum_{i<j} C_{ij}$, the total covariance due to linkage disequilibrium. Within a generation, selection changes the total variance by an amount that reflects its action in inducing changes in genotype frequencies at the individual loci, and in altering the covariance between pairs of loci, both within gametes and between loci on opposite gametes of the same individual. At equilibrium the first term is equal to $-U$, since the corresponding increase in mean and variance due to mutation is U. The normal approximation to the change in variance due to selection refers only to the between-loci component (Bulmer, 1980, p. 150).

The equilibrium composition of the population can be calculated as follows. Let the variance at the start of a generation be V. Using normal theory, selection changes this variance by an amount $\Delta = -\beta V^2/(1+\beta V)$. A corresponding equation can be written for the change in mean due to selection and mutation, and is identical to eqn (2). According to the argument

https://doi.org/10.1017/S0016672300025532 Published online by Cambridge University Press

Table 2. *Equilibrium properties of sexual populations with no genetic recombination*

	Exact results					Approximate results [from eqns (6), (7)]			Approximate results [from eqns (10), (11)]		
	(*a*) Effect of mutation rate [$a = 0{\cdot}01$, $b = 0{\cdot}01$, $h = 0{\cdot}2$ ($\alpha = 0{\cdot}002$, $\beta = 0{\cdot}0008$)]										
U	$\bar{w}$	$\bar{n}$	V	Skewness	Kurtosis	$\bar{w}$	$\bar{n}$	V	$\bar{w}$	$\bar{n}$	V
5·0	0·032	92·1	75·6	0·086	2·642	0·032	92·3	69·6	0·040	89·4	71·9
2·0	0·251	57·0	43·9	0·098	3·007	0·251	57·0	43·5	0·270	55·4	44·7
1·5	0·353	49·0	37·8	0·106	3·008	0·353	49·0	37·5	0·372	47·7	38·5
1·0	0·497	39·5	30·6	0·118	3·010	0·497	39·5	30·5	0·513	38·5	31·2
0·5	0·701	27·3	21·3	0·142	3·015	0·701	27·3	21·3	0·710	26·7	21·8
0·1	0·927	11·2	9·01	0·261	2·959	0·927	11·2	9·21	0·928	11·1	9·24
	(*b*) Effect of deviations from multiplicativity [$U = 1{\cdot}0$, $a = 0{\cdot}1$ ($\alpha = 0{\cdot}02$)]										
b	$\bar{w}$	$\bar{n}$	V	Skewness	Kurtosis	$\bar{w}$	$\bar{n}$	V	$\bar{w}$	$\bar{n}$	V
−0·001 25	0·330	67·3	75·8	0·148	3·027	0·330	67·3	74·8	0·342	64·6	73·3
−0·001	0·342	61·9	66·9	0·144	3·024	0·342	61·9	66·1	0·350	60·4	65·6
−0·0005	0·357	55·0	56·7	0·141	3·021	0·357	55·0	56·1	0·362	54·3	56·0
0	0·368	50·5	50·5	0·141	3·020	0·368	50·0	50·0	0·368	50·0	50·0
0·01	0·425	27·8	24·3	0·158	3·020	0·425	27·8	24·1	0·432	27·5	24·3
0·05	0·458	15·6	12·9	0·204	3·018	0·457	15·6	12·8	0·466	15·4	12·9
0·1	0·466	11·7	9·58	0·254	3·975	0·466	11·7	9·46	0·475	11·5	9·58

of Bulmer (1980, p. 159), the value of C_{ij} after selection is given by

$$C_{ij}^{*} = (1-r_{ij})\,C_{ij} + \frac{\Delta}{4m(m-1)}. \tag{8}$$

Equating the new and old values of the C_{ij}, and summing over all locus pairs, the equilibrium value of C_L at the start of a generation can readily be calculated (Bulmer, 1980, p. 159). This is equal to the difference between V and $V_A = \bar{n}$. Hence,

$$V - \bar{n} = C_L, \tag{9}$$

where the equilibrium value of $\bar{n}$ is given by eqn (3).

In the present case, where low recombination rates are of interest, Bulmer's equilibrium expression for C_L is clearly not adequate, as it predicts an infinite value when the frequency of recombination tends to zero for all pairs of loci. The discrepancy arises from the fact that changes in allele frequencies induce changes in linkage disequilibria among the loci concerned (Thomson, 1977). Using her results, it is easy to see that the coefficient of linkage disequilibrium D between a neutral locus and a linked locus with a rare allele subject to a selection coefficient hs against the heterozygote is reduced by hsD each generation, due to the change in allele frequency at the selected locus. Hence, a pair of linked loci both segregating for rare deleterious alleles will experience a reduction of approximately $2hsD$, due to changes in allele frequencies at each locus. In the present case, hs can be estimated as the ratio of the reduction due to selection in the mean number of mutations per individual ($-\Delta_s \bar{n}$) to the mean number $\bar{n}$; if the population is at equilibrium, we thus have $hs = U/\bar{n}$. C_{ij} is thus reduced by the additional term $2hsC_{ij}$. The equilibrium value of C_{ij} is given by

$$C_{ij} = \frac{\Delta}{4m(m-1)\,(r_{ij}+2hs)}. \tag{10a}$$

We thus obtain the equilibrium value of C_L as

$$C_L = \frac{E\Delta}{2}, \tag{10b}$$

where E is the expectation over all pairs of loci of $1/(r_{ij}+2hs)$. It follows from this result and eqn (9) that the equilibrium total variance V is less than the genic variance ($\bar{n}$) when $\beta > 0$. If $\beta < 0$, a non-negative variance after selection requires $-\beta V < 1$. Under this condition, it is obvious that $V > \bar{n}$ at equilibrium when $\beta < 0$. As would be expected, when $\beta = 0$ we have $V = \bar{n}$.

Given knowledge of the distribution of recombination values for all pairs of loci in the genome, eqns (3), (9) and (10) can be combined to yield the following equation:

$$f(V) = V^3\beta^2\left(1+\frac{E}{2}\right) + V^2\beta(1+\alpha-\beta U) - V(2\beta U-\alpha) - U = 0. \tag{11}$$

The equilibrium mean copy number can be obtained by substituting the solution of this equation into eqn (3).

In general, numerical analysis is needed to obtain information concerning the properties of the equilibria from these results. However, some useful insights into the dependence of the equilibrium composition of the population on the frequency of recombination can be obtained as follows, from the properties of the derivatives of $\bar{n}$, V and $\bar{w}$ with respect to a measure of the frequency of recombination. If the r_{ij} are increasing functions of a parameter r that measures the total frequency of recombination in the genome (so that $\partial r_{ij}/\partial r \geqslant 0$), E in equation (10) is a decreasing function of r. Provided that $\partial r_{ij}/\partial r > 0$ for at least some pairs of loci, $\partial E/\partial r < 0$. The derivatives of the population parameters with respect to E can thus be used to assess the nature of their dependence on r.

Using implicit differentiation of equation (11), we obtain

$$-\frac{dV}{dE} = \frac{\beta^2 V^3}{2g(V)}, \tag{12}$$

where

$$g(V) = 3V^2\beta^2\left(1+\frac{E}{2}\right) + 2V\beta(1+\alpha-\beta U) - (2\beta U-\alpha) + \frac{V^3\beta^2}{2}\frac{\partial E}{\partial V}.$$

When $\beta > 0$, the minimum value of $\partial E/\partial V$ is attained when $r_{ij} = 0$ for all loci, and is equal to $-1/2\beta V^2$. We thus have

$$Vg(V) > f(V) + V^2\beta(1+\alpha-\beta U) + U - \frac{V^2\beta}{4} = V^2\beta(\tfrac{3}{4}+\alpha-\beta U) + U.$$

Therefore, the conditions $\frac{3}{4}+\alpha-\beta U > 0$ or $(\beta V)^2 \leqslant 1$ are sufficient for $dV/dr > 0$. Since

$$\partial\bar{n}/\partial V = -U/\beta V^2 < 0,$$

it follows that $d\bar{n}/dr < 0$ under these conditions. Hence, the mean number of mutations decreases with an increase in the recombination frequency, and the variance increases.

This suggests that, at least when β is sufficiently small, the equilibrium mean fitness increases in response to an increase in recombination if $\beta > 0$. This can be examined as follows. Using equation (A 2), we obtain

$$\frac{d\ln\bar{w}}{dr} = \frac{d\ln\bar{w}}{dV}\frac{\partial V}{\partial r} = \frac{\partial V}{\partial r}\frac{1}{1+\beta V_1} \times\left\{\frac{\beta}{2}\left(\frac{\alpha\bar{n}}{1+\beta V}+\frac{\bar{n}U}{V}-1-\frac{\alpha^2 V}{1+\beta V}\right)+\frac{\alpha^2}{2}+\frac{(\alpha+\beta\bar{n})\,U}{\beta V^2}\right\}. \tag{13}$$

With $\beta > 0, \bar{n} > V$. Hence, when $\alpha < 1$, as can reasonably be assumed, we have $\bar{n}\alpha > \alpha^2 V$. The only

https://doi.org/10.1017/S0016672300025532 Published online by Cambridge University Press

negative contribution to the term in braces therefore arises from $-\frac{1}{2}\beta$. It is obvious that $U \geqslant 1$ is sufficient for this term to be overcome by the positive term $\beta\bar{n}U/2V$. A less stringent condition, likely to be valid for $U < 1$ under most circumstances of biological interest, is provided by noting that eqn (3) implies that $U = \beta V\bar{n} + V(\alpha - \beta U) > \beta V\bar{n}$, provided that $\beta U < \alpha$. Hence, under this condition $U\beta\bar{n}/\beta V^2 > \beta\bar{n}$, and so the term in braces is positive if $\bar{n} > \frac{1}{2}$. We can therefore conclude that the population mean fitness increases with recombination frequency when $\beta > 0$, under the stated, rather light, conditions.

A similar argument can be used to show that V and $\bar{n}$ are increasing functions of recombination frequency when $\beta < 0$, whereas mean fitness is a decreasing function, under light conditions. Provided that $-\beta V \leqslant 1$, (which is required for a non-negative variance, from the equation for ΔV), eqn (2) implies that $\alpha + \beta\bar{n} \geqslant 0$, so that $-\beta \leqslant \alpha/\bar{n}$. This condition is also necessary for fitness to be a decreasing function of near the equilibrium [see section 2(i) above]. In addition, $\bar{n} > U/\alpha$ (the solution for $\beta = 0$), and $V > \bar{n}$. Hence, the magnitude of β that is compatible with the existence of an equilibrium and with a meaningful fitness function is severely constrained by the relation $-\beta < \alpha^2/U$. Therefore, unless the mutation rate is very small, $-\beta$ must be much smaller than α, and the equilibrium solution is usually close to that for $\alpha = 0$.

Equations (11) and (12) yield the result

$$V\frac{\partial f}{\partial V} > \beta V^2(1 + \alpha - \beta U) + U.$$

Using the inequalities on β derived above, it follows that a sufficient condition for positivity of $\partial f/\partial V$ is $-\beta < U/(V^2[1 + \alpha - \beta U]) < \alpha^2/U$, which is needed for the equilibrium to exist. Hence, V is an increasing function of the frequency of recombination. Since $\partial\bar{n}/\partial V > 0$, the same is true for $\bar{n}$. The behaviour of mean fitness can be evaluated using eqn (13). For small β, the condition for the term in braces in this equation to be negative is found, after some rearrangement, to be close to

$$\alpha U + \frac{\beta^2 U^3}{\alpha^2} > \frac{\beta^2 U^2}{2\alpha}\left(U + \frac{1}{\alpha}\right) - \frac{\beta U^2}{\alpha}.$$

If $-\beta U/\alpha \ll 1$ (for which $\alpha \ll 1$ is sufficient), the right-hand side of this is dominated by the term $-\beta U^2/\alpha$ and the left-hand side by αU, and so the condition reduces approximately to $\alpha^2 > -\beta U$, which is necessary for the existence of an equilibrium.

The adequacy of the normal approximation used to obtain these results was tested for the cases of $r_{ij} = 0$ for all loci (no recombination) and $r_{ij} = 0{\cdot}5$ (free recombination among all loci). Exact results for the first case can be obtained as described above; exact results for the second case by numerical iteration of the recurrence relations derived by Kondrashov (1984). The approximate results for zero recombination derived by this approach are given in Table 2 under 'Approximate results from eqns (10) and (11)'. It is evident that they are less accurate than those derived in section 2(iii) above, although the fit is still quite good. The exact and approximate values of the equilibrium mean variance, skewness and kurtosis of the copy number distributions for the case of free recombination are displayed in Table 3. It will be seen from Table 3 that the best fit to the exact results for free recombination is obtained when selection is relatively weak, and the mutation rate is higher than 0·1. When there is synergistic gene action, the mean fitness of the population with free recombination is greater than with no recombination, although the effect is not as large as the difference between sexual

Table 3. *Equilibrium properties of sexual populations with free recombination*

	Exact results					Approximate results		
	(*a*) Effect of mutation rate [$a = 0{\cdot}01$, $b = 0{\cdot}01$, $h = 0{\cdot}2$ ($\alpha = 0{\cdot}002$, $\beta = 0{\cdot}0008$)]							
U	$\bar{w}$	$\bar{n}$	V	Skewness	Kurtosis	$\bar{w}$	$\bar{n}$	V
2·0	0·329	50·8	49·2	0·134	3·014	0·333	50·6	49·0
1·5	0·434	43·6	42·4	0·145	3·018	0·436	43·4	42·2
1·0	0·571	35·0	34·2	0·163	3·026	0·571	35·0	34·2
0·5	0·752	24·1	23·7	0·198	3·038	0·750	24·2	23·8
0·1	0·940	9·84	9·77	0·315	3·098	0·938	10·1	10·0
	(*b*) Effect of deviations from multiplicativity [$U = 1{\cdot}0$, $a = 0{\cdot}1$ ($\alpha = 0{\cdot}02$)]							
b	$\bar{w}$	$\bar{n}$	V	Skewness	Kurtosis	$\bar{w}$	$\bar{n}$	V
−0·001 25	0·248	90·7	91·6	0·107	2·987	0·259	86·9	87·6
−0·001	0·308	68·9	69·3	0·121	3·022	0·312	67·9	68·3
−0·0005	0·347	56·7	56·8	0·133	3·019	0·350	56·1	56·2
0	0·368	50·5	50·5	0·141	3·019	0·368	50·0	50·0
0·01	0·462	25·8	25·4	0·192	3·035	0·465	25·6	25·2
0·05	0·504	14·3	13·8	0·249	3·057	0·507	14·3	13·7
0·1	0·512	10·8	10·3	0·281	3·071	0·515	10·8	10·2

https://doi.org/10.1017/S0016672300025532 Published online by Cambridge University Press

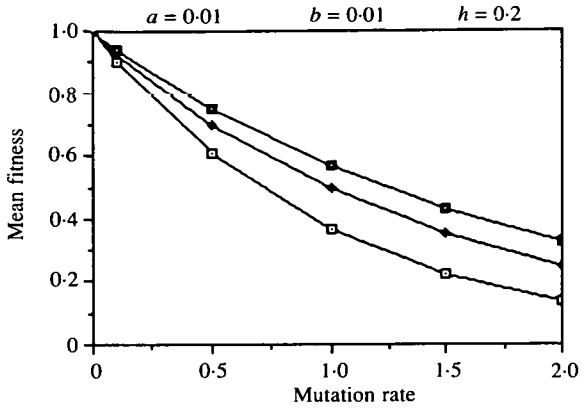

Fig. 1. Population mean fitness as a function of mutation rate, for asexual populations (⊡), sexual populations without recombination (◆), and sexual populations with free recombination (■). Standard values of the selection parameters are assumed.

reproduction with no recombination and asexual reproduction (Fig. 1).

More generally, the value of E can be determined by using the result of Morton (1955) that the probability of a map distance z between two loci sampled at random from a uniform distribution along a chromosome of total map length l is

$$\phi(z) = \frac{2(l-z)}{l^2}. \tag{14a}$$

Using Haldane's (1919) formula for the frequency of recombination between loci in the absence of interference, when there is a single chromosome of length l in the genome we obtain the following expression for E,

$$E_l = \frac{2}{l^2}\int_0^l \frac{l-z}{\{\frac{1}{2}(1-e^{-2z})+2s\}}\,dz. \tag{14b}$$

If there are j chromosomes of equal length l, we obtain

$$E \approx 2\left(1-\frac{1}{j}\right)+\frac{E_l}{j}. \tag{15}$$

This formula can be used to obtain numerical solutions to the equilibria, employing numerical integration to evaluate E. This enables studies to be made of the dependence of population mean fitness on the number of chromosomes, map length, and the selection and mutation parameters. Fig. 2 illustrates this dependence. Mean fitness is strongly affected by chromosome number, but is only strongly affected by map length when either j or l is small.

3. Spread of a modifier of recombination

(i) *General considerations*

The question of the direction and intensity of selection on a modifier of the frequency of recombination

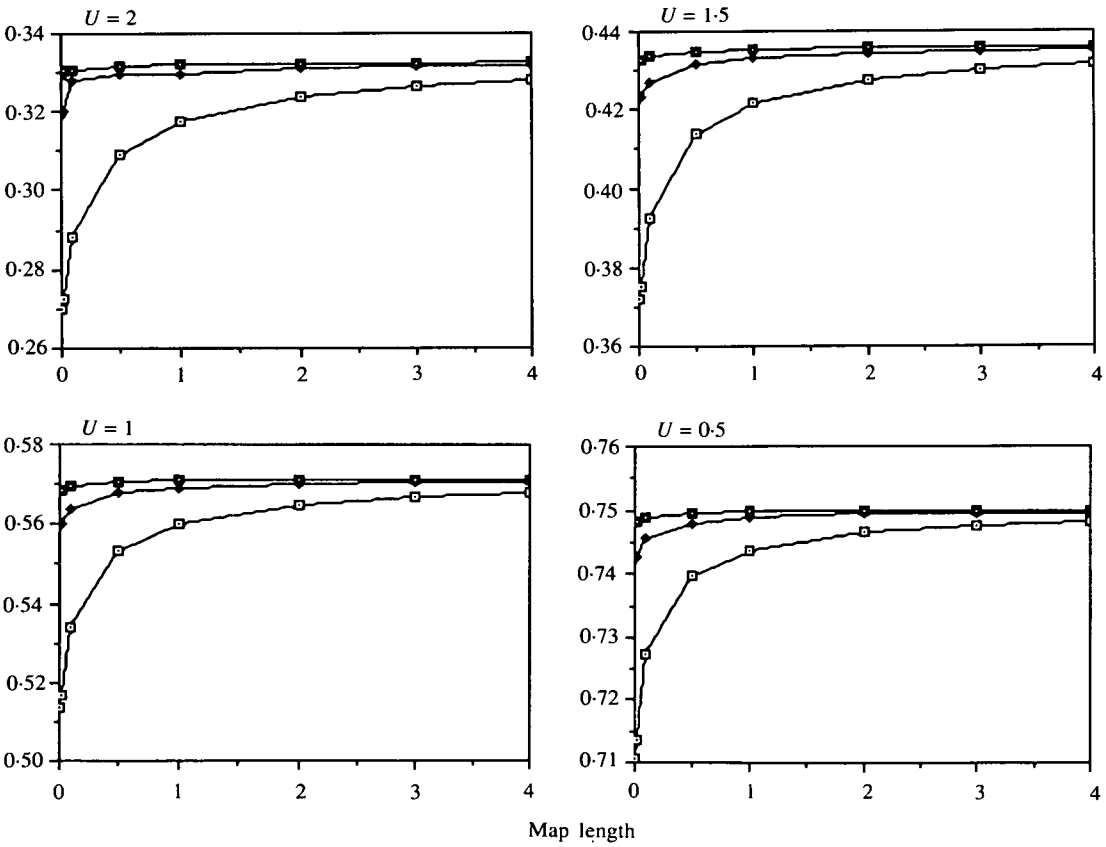

Fig. 2. Population mean fitness of sexual populations as a function of map length and number of chromosomes, for several different mutation rates. Standard values of the selection parameters are assumed. ⊡, 1 chromosome; ◆, 5 chromosomes; ■, 20 chromosomes.

https://doi.org/10.1017/S0016672300025532 Published online by Cambridge University Press

among the selected loci in a sexual population will be considered in this section. The initial population is assumed to be fixed for allele M_1 at the modifier locus; the new allele M_2 is assumed to be so rare that homozygotes can be ignored. The asymptotic values of the mean and variance in copy number for M_1/M_2 individuals will be calculated (denoted by $\bar{n}_2$ and V_2 respectively); this enables the asymptotic mean fitness of M_1/M_2 individuals ($\bar{w}_2$) to be calculated using eqn (A 2). The corresponding quantities for the initial M_1/M_1 population can be calculated by the methods described above, and will be designated by subscript 1.

There are two possibilities for the ordering of M and a pair of selected loci, i and j: (*a*) M is located outside i and j (with i being arbitrarily taken to be closer to M than j) (*b*) M is located between i and j. In case (*a*), let the frequency of recombination between M and i be ρ_i in M_1/M_2 individuals. The frequency of recombination between i and j is r_{ij1} in M_1/M_1 individuals, and r_{ij2} in M_1/M_2. The frequency of recombination between M and j in M_1/M_2 individuals is $\rho_j = \rho_i + r_{ij2} - 2\rho_i r_{ij2}$, assuming no interference. In case (*b*), r_{ij}, ρ_i and ρ_j retain their meaning, but the frequency of recombination between i and j in M_1/M_2 individuals is now $r_{ij2} = \rho_i + \rho_j - 2\rho_i\rho_j$. If the loci are on a different chromosome from M, then $\rho_i = \rho_j = \frac{1}{2}$.

Equations are derived in the Appendix that determine $\delta\bar{n}$ and δV, the asymptotic values of the deviations of $\bar{n}_2$ and V_2 from the equilibrium values for the initial M_1/M_1 population. These equations make use of the assumption that the modifier has a small effect on the frequency of recombination. They can be used to determine the difference in mean fitness between M_1/M_2 and M_1/M_1 individuals, and hence the direction and strength of selection on recombination. Unfortunately, in general these equations depend in a complex way on the distribution of frequencies of recombination between the pairs of selected loci and between the selected loci and the recombination modifier, so that it is not easy to obtain insight into the dynamics of the modifier other than by numerical solutions. Two limiting cases can be treated relatively easily, and these will be considered before discussion of the more general case.

(ii) *The low recombination limit*

The case when the frequency of recombination between all pairs of loci is close to zero for the initial M_1/M_1 population will be analysed first. This corresponds to an organism with a single chromosome with a short map, and is an appropriate framework for considering the initial evolution of recombination or the evolutionary stability of non-zero rates of recombination. A low frequency of recombination and a small effect of the modifier implies that the harmonic mean of the ρ_i that appears in eqn (A 12) will also be small. Hence, eqn (A 12) simplifies to

$$\delta\bar{n} \approx -\frac{U\delta V}{\beta V_1^2}. \tag{16}$$

[This is identical to the expression for $\delta\bar{n}$ given by eqn (3).]

The recombination terms in the denominators of eqns (A 6) and (A 8) can similarly be neglected in this case, leading to the following expression for δV when substituted into eqn (A 9):

$$\delta V - \delta\bar{n} \approx \frac{\left\{\delta\Delta - \frac{(\delta\bar{r} + 2\delta hs)\,\Delta_1}{2hs_1}\right\}}{8hs_1}, \tag{17}$$

where $\delta\Delta$ is the difference in Δ between M_1/M_2 and M_1/M_1 individuals, $\delta\bar{r}$ is the corresponding difference in the mean recombination frequency between all pairs of loci, and δhs is the difference in the selection coefficient against heterozygous mutant alleles at each locus. We have

$$\delta\Delta \approx -\frac{\beta V_1(2+\beta V_1)\,\delta V}{(1+\beta V_1)^2}, \tag{18a}$$

and

$$\delta hs \approx \frac{\delta\bar{n}}{\bar{n}_1}\left\{\frac{\beta V_1}{1+\beta V_1} - \frac{U}{\bar{n}_1}\right\} + \frac{U\delta V}{\bar{n}_1 V_1(1+\beta V_1)}. \tag{18b}$$

Combining these three equations, we obtain the following expression:

$$\frac{\partial V}{\partial\bar{r}} \approx \frac{\beta V_1^2 \bar{n}_1^2}{16U^2(1+\beta V_1)\left\{1 + \frac{U}{\beta V_1^2} - \frac{1}{8(1+\beta V_1)} + \frac{\beta V_1 \bar{n}_1(2+\beta V_1)}{8U(1+\beta V_1)^2}\right\}}. \tag{19}$$

A measure of the strength of selection on the rate of recombination can be obtained by using eqn (19) to calculate the partial derivative of $\ln\bar{w}$ with respect to $\bar{r}$, multiplying the expression for $d\ln\bar{w}/dV$ from eqn (13) by $\partial V/\partial\bar{r}$. We have:

$$d = \left(\frac{d\ln\bar{w}}{dV}\right)_{V_1}\left(\frac{\partial V}{\partial\bar{r}}\right)_{\bar{r}=0}. \tag{20}$$

This derivative is the selection gradient on mean recombination frequency at the low recombination limit. If recombination frequencies are controlled polygenically (with low frequency modifier alleles at each locus), the rate of change of the population mean of $\bar{r}$ is equal to the product of d and the additive genetic variance of $\bar{r}$ (cf. Lande, 1976). Combining eqns (19) and (16), we see that, when $\beta > 0$, modifiers which increase the frequency of recombination from a near-zero value will be associated with an increased variance in the number of mutant genes, and a reduced mean. The analysis of equation (13) indicates that in this case d is positive under biologically plausible conditions, and so selection will favour an

increase in the frequency of recombination away from zero. When $\beta < 0$, it is obvious from eqn (19), and the condition $-\beta < \alpha^2/U$ derived earlier, that $\partial V/\partial \bar{r} > 0$ when β is sufficiently small for V to be approximated by U/α. Hence, under this condition an increase in $\bar{r}$ is again associated with an increase in variance, but eqn (16) implies an increase in mean as well. As shown above, under these conditions, there is a reduction in mean fitness with an increase in $\bar{r}$, so that selection favours modifiers that reduce recombination.

Some numerical examples confirming these conclusions are shown in Fig. 3. The value of E in eqns (12) and (13) was calculated using the approximation $E \approx \frac{1}{2}s(1-\frac{1}{2}s\bar{r})$, which is valid for small $\bar{r}$. The gradients are calculated with respect to map length l, rather than $\bar{r}$, as l is a more fundamental variable. For a single chromosome with a low frequency of recombination, it follows from eqn (14) that $l \approx 3\bar{r}$. The gradients with respect to $\bar{r}$ are thus about 3 times the values shown in Fig. 3. It can be seen that a modifier with a small effect in increasing recombination away from zero appears always to have a selective advantage when $\beta > 0$; the reverse is true for $\beta < 0$.

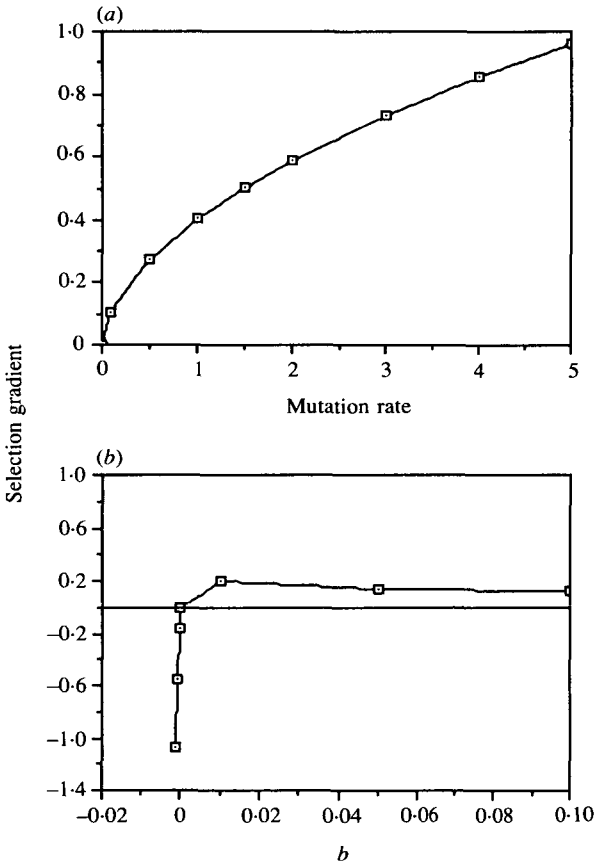

Fig. 3. Selection gradients on map length (l) for a modifier increasing map length away from zero, in an organism with a single chromosome. (*a*) Shows the dependence of the selection gradient on mutation rate for the standard values of the selection parameters. (*b*) shows the dependence of the selection gradient on the coefficient (b) of the quadratic term relating log fitness to number of homozygous mutations, when the linear coefficient (a) is 0·1 and the mutation rate (U) is 1.

(iii) *The high recombination limit*

The case of a modifier reducing recombination from the maximum value of one-half for all loci can also be studied analytically. Assuming that the modifier is unlinked to all the selected loci, equation (A 12) yields the following result

$$\delta\bar{n} \approx \frac{-U\delta V}{(1+2\beta V_1)\,V_1}. \tag{21}$$

Equations (10) and (A 6) give

$$\delta C_{ij} \approx \frac{4}{3}\left\{\frac{\delta\Delta}{4m(m-1)} - \frac{(\delta r_{ij}+2\delta hs)\,\Delta}{2m(m-1)}\right\}. \tag{22}$$

Summing over all pairs of loci, using eqn (21) to calculate δhs, and noting that eqn (10) in this case implies that $V_1(1+2\beta V_1) \approx \bar{n}_1(1+\beta V_1)$, we obtain

$$\frac{\partial V}{\partial \bar{r}} \approx \frac{2\beta V_1^2}{3(1+\beta V_1)\left\{1+\frac{1}{1+\beta V_1}\left(hs_1+\frac{\beta\bar{n}_1}{3}\right) - \frac{4\beta s_1(1+hs_1)\,V_1}{3(1+\beta V_1)(1+2\beta V_1)}\right\}}. \tag{23}$$

When $0 < \beta \ll 1$, the negative component of the term in braces is smaller in magnitude than the positive term, and so $\partial V/\partial \bar{r} > 0$, whereas equation (21) implies $\partial\bar{n}/\partial\bar{r} < 0$. When $\beta < 0$, the term in braces is positive when $-\beta \ll 1$, and so $\partial V/\partial\bar{r} < 0$ and $\partial\bar{n}/\partial\bar{r} > 0$.

The selection gradient with respect to $\bar{r}$ can be calculated by the same approach as before, and we obtain

$$d \approx \frac{1}{1+\beta V_1}\left\{\frac{\beta}{2}\left(\frac{\alpha\bar{n}_1}{1+\beta V_1}+\frac{\bar{n}_1 U}{V_1}-[1+\alpha^2 V_1]\right) + \frac{\alpha^2}{2}+\frac{(\alpha+\beta\bar{n}_1)\,U}{(1+2\beta V_1)\,V_1}\right\}. \tag{24}$$

The conditions for the term in braces to be positive when $\beta > 0$ are somewhat more stringent than the corresponding condition on eqn (13), since the right-hand term is considerably smaller than the corresponding term in eqn (13). When the mutation rate is sufficiently small, the quantity in braces can be negative even when $\beta > 0$, indicating that selection does not always push the frequency of recombination to its maximum (see section 4).

The stability of the state of free recombination to invasion by a modifier allele that completely suppresses recombination, and which is itself completely linked to the selected loci, can be studied similarly. This corresponds to the case of a chromosome rearrangement such as an inversion, in a species with a single chromosome with a very high rate of recombination. As before, second-order terms in $\delta\bar{n}$ will be ignored; this is justified by the fact that $\delta\bar{n}$ usually turns out to be small relative to the initial

https://doi.org/10.1017/S0016672300025532 Published online by Cambridge University Press

value, even in this extreme case. Using the same approach as before, we obtain

$$\delta V = \frac{-\beta V_1^2\left\{\frac{1}{8hs_1}-\frac{1}{2}\right\}}{(1+\beta V_1)^2\left\{1+\frac{U}{\beta V_1^2}+\frac{\beta V_1(2+\beta V_1)}{8hs_1(1+\beta V_1)^2}-\frac{1}{8(1+\beta V_1)}\right\}}. \quad (25)$$

The selection coefficient for the suppressor, t, can be determined by using the derivative of log mean fitness given by equation (13) together with equation (25) to calculate $\delta \ln \bar{w}$. With $\beta > 0$, it is easily seen by the same type of argument as used previously that $t < 0$ under light conditions, implying that recombination suppression is disfavoured regardless of the mutation rate. The reverse is true if $\beta < 0$. From standard theory, the frequency of the suppressor changes from x to $x(\bar{w}_2/\bar{w}_1)$ each generation, and so $t \approx \bar{w}_2/\bar{w}_1 - 1$ provides an adequate measure of the strength of selection in this case.

Table 4 gives some numerical examples of the strength of selection on both modifiers of small effect and on recombination suppressors, in initial populations with free recombination. In this case the selection gradient is expressed with respect to the mean frequency of recombination, $\bar{r}$, not map length, as the number of chromosomes and their map lengths are not specified in this case. The values of $\delta\bar{n}$ and δV for the case of modifier of small effect are calculated assuming a $\delta\bar{r}$ of 0·01.

(iv) *The general case*

The case of an organism with j chromosomes of equal length l (where l is the value for the initial population) will be considered here, using an approach similar to that used to obtain eqn (15). Provided that the initial map length is non-zero, it is reasonable to assume that the modifier allele M_2 changes the map length of each chromosome by an amount $\delta l = \epsilon l$, and that this increase is spread uniformly across the chromosome, so that the frequency of recombination $r(z)$ between two loci initially separated by map distance z is changed by $\delta r(z) = \epsilon z \exp -2z$, assuming the Haldane mapping function. Equations (A 6), (A 8) and (A 12) can then be used as follows to obtain $\delta\bar{n}$ and δV.

A fraction $(1-1/j)$ of the total of $\frac{1}{2}m(m-1)$ loci pairs involve loci on different chromosomes, so that the modifier has no effect on them, and they contribute nothing to $\delta\bar{n}$ and δV in eqn (A 9). Of the remaining $\frac{1}{2}m(m-1)/j$ loci pairs, a fraction $(1-1/j)$ are unlinked to the modifier locus, and so we can apply equation (A 6) with $\rho_i = \frac{1}{2}$. Neglecting second-order terms, and using eqn (15), this gives the following contribution to $\delta V - \delta\bar{n}$ in eqn (A 9)

$$\left(1-\frac{1}{j}\right)\frac{1}{jl^2}\int_0^l \frac{\left\{\delta\Delta - \frac{[\delta r(z)+2\delta hs_1]\,\Delta_1}{[r(z)+2hs_1]}\right\}}{1+r(z)+2hs_1}(l-z)\,dz \quad (26a)$$

δhs in this equation can be evaluated by the methods used in section 3(iii) above.

A fraction $1/j^2$ of loci pairs are located on the same chromosome as the modifier. The nature of the contributions from this case to $\delta V - \delta\bar{n}$ depend on the location of the modifier. Assume that a fraction p of loci are located to the left of the modifier, and a fraction $q = 1-p$ to the right. In a proportion $(1-2pq)$ of cases, the modifier will be located outside a randomly chosen pair of loci, and in a proportion $2pq$ of cases it will be between the loci. When the modifier is to the right of the selected loci (frequency p^2), we

Table 4. *Selection on recombination at the high recombination limit*

	Modifier of small effect			Suppressor of recombination		
	(*a*) Effect of mutation rate [$a = 0{\cdot}01$, $b = 0{\cdot}01$, $h = 0{\cdot}2$ ($\alpha = 0{\cdot}002$, $\beta = 0{\cdot}0008$)]					
U	$d(\times 10^{-2})$	$\delta\bar{n}(\times 10^{-4})$	$\delta V(\times 10^{-2})$	t	$\delta\bar{n}$	δV
5·0	1·52	16·3	−2·88	−0·222	3·40	−3·33
2·0	0·234	4·43	−1·17	−0·099	2·41	−2·31
1·5	0·126	2·93	−0·881	−0·076	2·13	−2·02
1·0	0·050	1·63	−0·589	−0·052	1·78	−1·67
0·5	0·007	0·585	−0·289	−0·027	1·28	−1·12
0·1	−0·001	0·051	−0·052	−0·005	0·538	−0·429
0·05	−0·0008	0·018	−0·003	−0·002	0·349	−0·255
0·01	−0·0001	0·005	−0·002	−0·0004	0·104	−0·052
b	(*b*) Effect of deviations from multiplicativity [$U = 1{\cdot}0$, $a = 0{\cdot}1$ ($\alpha = 0{\cdot}02$)]					
−0·00125	−0·013	−0·595	0·512	0·228	−20·2	−15·5
−0·00100	−0·010	−0·365	0·247	0·064	−4·36	−1·63
−0·00050	−0·004	−0·148	0·083	0·016	−0·926	−0·117
0	0	0	0	0	0	0
0·01	0·061	1·22	0·318	−0·356	0·892	−0·456
0·05	0·245	2·89	0·439	−0·037	0·498	−0·037
0·10	0·436	3·92	0·465	−0·033	0·329	−0·033

https://doi.org/10.1017/S0016672300025532 Published online by Cambridge University Press

can apply eqn (A 6). The probability density that the closest of the pair is map distance z from M is $1/lp$ and the probability density that the second is located a distance y from the first is $1/lpz$. The net contribution from this case is thus

$$\frac{p}{4j^2l}\int_0^{pl}\int_0^{pl-z}\frac{\left\{\delta\Delta-\dfrac{[\delta r(y)+2\delta hs]\Delta_1}{r(y)+2hs_1}\right\}}{(pl-z)\{r(z)+[1-r(z)][r(y)+2hs_1]\}}dy\,dz. \tag{26b}$$

A similar contribution accrues from the case when the modifier is to the left of the selected loci, except that q replaces p.

Using eqn (A 8), we obtain the following contribution from the case when the modifier is located between the selected loci. Here, we write

$$r(z,y) = r(z)+r(y)-2r(z)r(y),$$

and

$$\delta r(z,y) = \epsilon z[1-2r(y)](\exp -2z) + \epsilon y[1-2r(z)](\exp -2y)$$

$$\frac{1}{2j^2l^2}\int_0^{pl}\int_0^{ql} \times \frac{\left\{\delta\Delta-\dfrac{[\delta r(z,y)+2\delta hs_1]\Delta_1}{r(z,y)+2hs_1}\right\}}{\left\{r(z,y)+r(z)\,r(y)+\dfrac{2s_1[1-r(z)][1-r(y)]}{[1-r(z,y)]}\right\}}dy\,dz. \tag{26c}$$

The sum of these four contributions yields the value of $\delta V-\delta\bar{n}$. As before, eqn (A 12) provides an independent equation for $\delta\bar{n}$. The harmonic mean of the ρ_i in eqn (A 12) is given by the expression

$$\left.\begin{aligned}\frac{1}{\rho_H} &= 2\left(1-\frac{1}{j}\right)+\frac{2}{pjl}\int_{z_M}^{pl}\frac{dz}{1-e^{-2z}}+\frac{2}{qjl}\int_{z_M}^{ql}\frac{dz}{1-e^{-2z}}\\ &\approx 2+\frac{1}{jl}\ln\left\{\frac{(1-e^{-2pl})(1-e^{-2ql})}{(1-e^{-2z_M})^2}\right\},\end{aligned}\right\} \tag{27}$$

where z_M is the distance between the closest selected locus and the modifier. On the assumption of even spacing between loci, we have $z_M = lj/(m-j)$.

These relations provide closed expressions in $\delta\bar{n}$ and δV that can be used to obtain solutions by Newton–Raphson iteration and numerical integration. A copy of the MACFORTRAN program for performing these calculations will be provided on request, if a 3·5″ disc is supplied. The value of t, the selection coefficient on a modifier of specified effect, can be obtained from the corresponding value of $\delta \ln \bar{w}$. The selection gradient on the proportion by which the map length of each chromosome is increased by the modifier is approximated by $d = t/\epsilon$, provided that ϵ is small. Table 5 shows the results of some calculations of this selection gradient for the standard selection parameter values in the case of a modifier with $\epsilon = 0{\cdot}001$, located in the middle of a chromosome ($p = 0{\cdot}5$). The rows of the table correspond to different initial map lengths (l) and the columns to different chromosome numbers (j). The selection gradient on map length itself is d/l, which is much larger than the gradient for ϵ when l is small. It will be seen that for the lowest mutation rate shown ($U = 0{\cdot}1$), there can actually be selection against increased recombination if l and j are sufficiently large, indicating the existence of an evolutionarily stable (ESS) map length that is characteristic of a given number of chromosomes [cf. section 3(iii)]. For higher mutation rates, there is always selection for increased recombination, although the

Table 5. *General model of selection on recombination modifiers: effect of mutation rate on strength of selection*

	$a = 0{\cdot}01$, $b = 0{\cdot}01$, $h = 0{\cdot}2$ ($\alpha = 0{\cdot}002$, $\beta = 0{\cdot}0008$)							
l	1	2	5	20	1	2	5	20
		$U = 0{\cdot}1$				$U = 0{\cdot}5$		
0·05	$1{\cdot}5\times10^{-1}$	$1{\cdot}4\times10^{-4}$	$1{\cdot}3\times10^{-5}$	$2{\cdot}2\times10^{-7}$	$6{\cdot}2\times10^{-1}$	$9{\cdot}8\times10^{-4}$	$1{\cdot}5\times10^{-4}$	$8{\cdot}1\times10^{-6}$
0·10	$6{\cdot}1\times10^{-3}$	$8{\cdot}4\times10^{-5}$	$8{\cdot}4\times10^{-6}$	$-2{\cdot}4\times10^{-8}$	$3{\cdot}6\times10^{-1}$	$9{\cdot}1\times10^{-4}$	$1{\cdot}3\times10^{-4}$	$6{\cdot}6\times10^{-6}$
0·50	$3{\cdot}2\times10^{-3}$	$8{\cdot}5\times10^{-6}$	$-1{\cdot}1\times10^{-7}$	$-2{\cdot}5\times10^{-7}$	$4{\cdot}4\times10^{-2}$	$2{\cdot}8\times10^{-4}$	$3{\cdot}8\times10^{-5}$	$2{\cdot}8\times10^{-6}$
1·0	$6{\cdot}0\times10^{-4}$	$1{\cdot}0\times10^{-6}$	$-6{\cdot}4\times10^{-7}$	$-2{\cdot}3\times10^{-7}$	$1{\cdot}4\times10^{-2}$	$1{\cdot}1\times10^{-4}$	$1{\cdot}8\times10^{-5}$	$1{\cdot}8\times10^{-6}$
2·0	$5{\cdot}7\times10^{-5}$	$-7{\cdot}8\times10^{-7}$	$-6{\cdot}0\times10^{-7}$	$-1{\cdot}8\times10^{-7}$	$4{\cdot}1\times10^{-3}$	$3{\cdot}9\times10^{-5}$	$7{\cdot}8\times10^{-6}$	$1{\cdot}1\times10^{-6}$
3·0	$-1{\cdot}5\times10^{-5}$	$-7{\cdot}7\times10^{-7}$	$-4{\cdot}7\times10^{-7}$	$-1{\cdot}4\times10^{-7}$	$2{\cdot}1\times10^{-3}$	$1{\cdot}8\times10^{-5}$	$4{\cdot}5\times10^{-6}$	$7{\cdot}9\times10^{-7}$
4·0	$-3{\cdot}3\times10^{-5}$	$-5{\cdot}9\times10^{-7}$	$-3{\cdot}6\times10^{-7}$	$-1{\cdot}1\times10^{-7}$	$1{\cdot}3\times10^{-3}$	$9{\cdot}5\times10^{-6}$	$3{\cdot}0\times10^{-6}$	$6{\cdot}0\times10^{-7}$
5·0	$-3{\cdot}9\times10^{-5}$	$-4{\cdot}4\times10^{-7}$	$-2{\cdot}9\times10^{-7}$	$-9{\cdot}8\times10^{-8}$	$9{\cdot}8\times10^{-4}$	$9{\cdot}2\times10^{-6}$	$2{\cdot}1\times10^{-6}$	$4{\cdot}9\times10^{-7}$
		$U = 1{\cdot}0$				$U = 2{\cdot}0$		
0·05	1·1	$2{\cdot}0\times10^{-3}$	$3{\cdot}5\times10^{-4}$	$2{\cdot}5\times10^{-5}$	1·7	$3{\cdot}8\times10^{-3}$	$7{\cdot}5\times10^{-4}$	$6{\cdot}9\times10^{-5}$
0·10	$6{\cdot}8\times10^{-1}$	$2{\cdot}1\times10^{-3}$	$3{\cdot}5\times10^{-4}$	$2{\cdot}3\times10^{-5}$	1·2	$4{\cdot}6\times10^{-3}$	$8{\cdot}6\times10^{-4}$	$7{\cdot}2\times10^{-5}$
0·50	$1{\cdot}2\times10^{-1}$	$9{\cdot}1\times10^{-4}$	$1{\cdot}5\times10^{-4}$	$1{\cdot}4\times10^{-5}$	$2{\cdot}8\times10^{-1}$	$2{\cdot}7\times10^{-3}$	$5{\cdot}0\times10^{-4}$	$5{\cdot}3\times10^{-5}$
1·0	$4{\cdot}2\times10^{-2}$	$4{\cdot}3\times10^{-4}$	$7{\cdot}9\times10^{-5}$	$9{\cdot}9\times10^{-6}$	$1{\cdot}1\times10^{-1}$	$1{\cdot}4\times10^{-3}$	$2{\cdot}9\times10^{-4}$	$4{\cdot}0\times10^{-5}$
2·0	$1{\cdot}4\times10^{-2}$	$1{\cdot}6\times10^{-4}$	$3{\cdot}9\times10^{-5}$	$6{\cdot}5\times10^{-6}$	$4{\cdot}3\times10^{-2}$	$6{\cdot}0\times10^{-4}$	$1{\cdot}5\times10^{-4}$	$2{\cdot}7\times10^{-5}$
3·0	$7{\cdot}7\times10^{-3}$	$8{\cdot}1\times10^{-5}$	$2{\cdot}4\times10^{-5}$	$4{\cdot}8\times10^{06}$	$2{\cdot}5\times10^{-2}$	$3{\cdot}0\times10^{-4}$	$9{\cdot}6\times10^{-5}$	$2{\cdot}0\times10^{-5}$
4·0	$5{\cdot}2\times10^{-3}$	$4{\cdot}4\times10^{-5}$	$1{\cdot}6\times10^{-5}$	$3{\cdot}7\times10^{-6}$	$1{\cdot}8\times10^{-2}$	$1{\cdot}7\times10^{-4}$	$6{\cdot}7\times10^{-5}$	$1{\cdot}6\times10^{-5}$
5·0	$3{\cdot}9\times10^{-3}$	$2{\cdot}5\times10^{-5}$	$1{\cdot}2\times10^{-5}$	$3{\cdot}0\times10^{-6}$	$1{\cdot}4\times10^{-2}$	$9{\cdot}2\times10^{-5}$	$4{\cdot}9\times10^{-5}$	$1{\cdot}3\times10^{-5}$

https://doi.org/10.1017/S0016672300025532 Published online by Cambridge University Press

Table 6. *General model of selection on recombination modifiers: effect of position of modifier on strength of selection*

				(*a*) $l = 1{\cdot}0$ [$a = 0{\cdot}01$, $b = 0{\cdot}01$, $h = 0{\cdot}2$ ($\alpha = 0{\cdot}002$, $\beta = 0{\cdot}0008$)]				
P	1	2	5	20	1	2	5	20
		$U = 0{\cdot}1$				$U = 0{\cdot}5$		
0·01	$4{\cdot}8 \times 10^{-4}$	$6{\cdot}5 \times 10^{-7}$	$-6{\cdot}5 \times 10^{-7}$	$-2{\cdot}3 \times 10^{-7}$	$9{\cdot}6 \times 10^{-3}$	$8{\cdot}5 \times 10^{-5}$	$1{\cdot}6 \times 10^{-5}$	$1{\cdot}7 \times 10^{-6}$
0·10	$4{\cdot}8 \times 10^{-4}$	$6{\cdot}5 \times 10^{-7}$	$-6{\cdot}5 \times 10^{-7}$	$-2{\cdot}3 \times 10^{-7}$	$1{\cdot}2 \times 10^{-2}$	$1{\cdot}1 \times 10^{-4}$	$1{\cdot}7 \times 10^{-5}$	$1{\cdot}8 \times 10^{-6}$
0·50	$6{\cdot}0 \times 10^{-4}$	$1{\cdot}0 \times 10^{-6}$	$-6{\cdot}4 \times 10^{-7}$	$-2{\cdot}3 \times 10^{-7}$	$1{\cdot}4 \times 10^{-2}$	$1{\cdot}1 \times 10^{-4}$	$1{\cdot}8 \times 10^{-5}$	$1{\cdot}8 \times 10^{-6}$
		$U = 1{\cdot}0$				$U = 2{\cdot}0$		
0·01	$3{\cdot}0 \times 10^{-2}$	$3{\cdot}7 \times 10^{-8}$	$6{\cdot}7 \times 10^{-4}$	$9{\cdot}3 \times 10^{-6}$	$8{\cdot}6 \times 10^{-2}$	$1{\cdot}2 \times 10^{-3}$	$2{\cdot}6 \times 10^{-4}$	$3{\cdot}8 \times 10^{-5}$
0·10	$3{\cdot}6 \times 10^{-2}$	$3{\cdot}9 \times 10^{-4}$	$7{\cdot}4 \times 10^{-4}$	$9{\cdot}7 \times 10^{-6}$	$9{\cdot}9 \times 10^{-2}$	$1{\cdot}3 \times 10^{-3}$	$2{\cdot}8 \times 10^{-4}$	$3{\cdot}9 \times 10^{-5}$
0·50	$4{\cdot}2 \times 10^{-2}$	$4{\cdot}3 \times 10^{-4}$	$7{\cdot}9 \times 10^{-4}$	$9{\cdot}9 \times 10^{-6}$	$1{\cdot}1 \times 10^{-1}$	$1{\cdot}4 \times 10^{-3}$	$2{\cdot}9 \times 10^{-4}$	$4{\cdot}0 \times 10^{-5}$
				(*b*) $l = 0{\cdot}1$ [$a = 0{\cdot}01$, $b = 0{\cdot}01$, $h = 0{\cdot}2$ ($\alpha = 0{\cdot}002$, $\beta = 0{\cdot}008$)]				
		$U = 0{\cdot}1$				$U = 0{\cdot}5$		
0·01	$4{\cdot}0 \times 10^{-3}$	$4{\cdot}9 \times 10^{-5}$	$3{\cdot}9 \times 10^{-6}$	$-1{\cdot}1 \times 10^{-7}$	$2{\cdot}7 \times 10^{-1}$	$6{\cdot}3 \times 10^{-7}$	$9{\cdot}0 \times 10^{-5}$	$5{\cdot}0 \times 10^{-6}$
0·10	$4{\cdot}9 \times 10^{-3}$	$6{\cdot}4 \times 10^{-5}$	$5{\cdot}5 \times 10^{-6}$	$-5{\cdot}0 \times 10^{-8}$	$3{\cdot}1 \times 10^{-1}$	$7{\cdot}4 \times 10^{-7}$	$1{\cdot}1 \times 10^{-4}$	$5{\cdot}9 \times 10^{-6}$
0·50	$6{\cdot}1 \times 10^{-3}$	$8{\cdot}4 \times 10^{-5}$	$8{\cdot}4 \times 10^{-6}$	$-2{\cdot}4 \times 10^{-8}$	$3{\cdot}6 \times 10^{-1}$	$9{\cdot}1 \times 10^{-7}$	$1{\cdot}3 \times 10^{-4}$	$6{\cdot}6 \times 10^{-6}$
		$U = 1{\cdot}0$				$U = 2{\cdot}0$		
0·01	$5{\cdot}5 \times 10^{-1}$	$1{\cdot}5 \times 10^{-3}$	$2{\cdot}6 \times 10^{-4}$	$1{\cdot}9 \times 10^{-5}$	1·0	$3{\cdot}5 \times 10^{-3}$	$6{\cdot}9 \times 10^{-4}$	$6{\cdot}1 \times 10^{-5}$
0·10	$5{\cdot}9 \times 10^{-1}$	$1{\cdot}8 \times 10^{-3}$	$3{\cdot}0 \times 10^{-4}$	$2{\cdot}1 \times 10^{-5}$	1·1	$3{\cdot}9 \times 10^{-3}$	$7{\cdot}7 \times 10^{-4}$	$6{\cdot}7 \times 10^{-5}$
0·50	$6{\cdot}8 \times 10^{-1}$	$2{\cdot}1 \times 10^{-3}$	$3{\cdot}5 \times 10^{-4}$	$2{\cdot}3 \times 10^{-5}$	1·2	$4{\cdot}6 \times 10^{-3}$	$8{\cdot}6 \times 10^{-4}$	$7{\cdot}2 \times 10^{-5}$

strength of this falls off rapidly with increasing initial map length and chromosome number, even with a mutation rate as high as 2. Table 6 displays the effect of the chromosomal position of the modifier on the selection gradient. As might be expected, a modifier towards the end of the chromosome is selected for more strongly than one towards the middle, but the effect is usually small. Fig. 4 displays the effect of varying the quadratic term in the fitness equation. Negative b values lead, as expected, to selection against increased map length. When $b > 0$, the strength of selection for increased map length increases with b but is weak when the number of chromosomes is large.

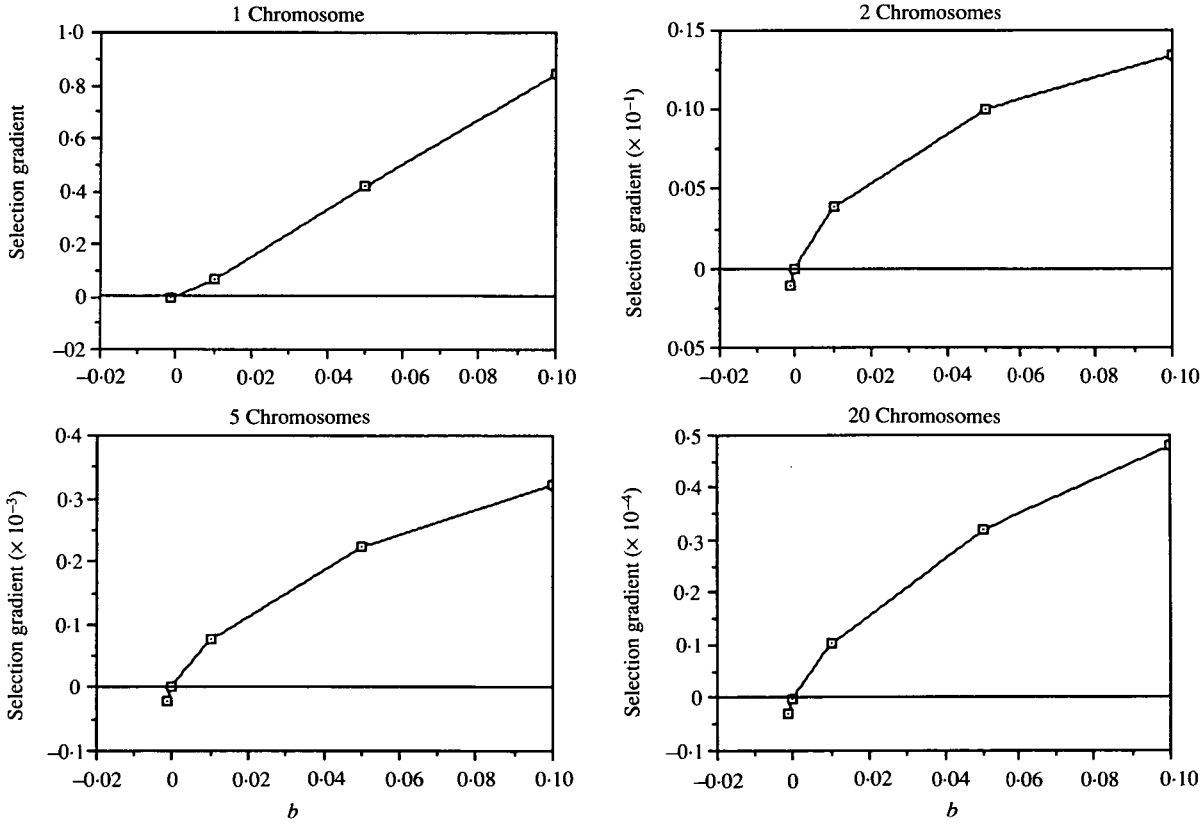

Fig. 4. Selection gradients on the proportional effect of a modifier on map length (ϵ), as a function of the coefficient (b) of the quadratic term relating lot fitness to number of homozygous mutations, for different numbers of chromosomes. An initial map length of 1 and a mutation rate of 1 are assumed. The linear selection term (a) is 0·1.

https://doi.org/10.1017/S0016672300025532 Published online by Cambridge University Press

4. Discussion

(i) *Population mean fitness with epistasis*

The results of this study are in agreement with the earlier studies of Kimura and Maruyama (1966), Crow (1970) and Kondrashov (1982, 1984) as far the behaviour of population mean fitness under multilocus mutation-selection balance is concerned. The method used here has the advantage that equilibrium solutions to a wide variety of breeding systems can be obtained straightforwardly, without making too many specialising assumptions. With asexual reproduction and a per genome mutation rate of U, the population mean fitness at equilibrium is $\exp - U$, independent of the mode of selection. With sexual reproduction and random mating in the absence of recombination (i.e. with a single chromosome that does not cross over), in all cases studied the equilibrium mean fitness is greater than $\exp - U$ with synergistic epistasis, and less with diminishing returns epistasis (cf. Tables 1 and 2). The difference can be substantial if the mutation rate is sufficiently high (Fig. 1). For example, with the standard values of the selection parameters, the ratio of the respective mean fitnesses is 1·58 when $U = 1{\cdot}5$. This difference between asexuality and sexuality with no recombination has previously been noted by Kimura & Maruyama (1966). It arises from the fact that selection within a generation induces covariances between loci within the same gamete (C_L) and covariances between loci carried on the maternal and paternal gametes of the same individual (C_{HW}). Both covariances are transmitted intact to the next generation in the case of asexual reproduction, but the C_{HW} term is destroyed by random mating in the case of a sexual population. Hence, there is a larger effect of covariance terms in the case of asexuality than sex with no recombination. This can be seen in Tables 1 and 2; e.g. with $U = 1{\cdot}5$ and the standard values of the selection parameters, the ratio of the total variance in number of mutations to the genic variance is 0·53 for asexuality and 0·77 with sexuality and no recombination; the total variance is larger in the latter case. There is a parallel change in the efficiency of selection, leading to a lower equilibrium mean number of mutations with sexuality and no recombination (59·3 *vs.* 49·0). Although an analytical proof of these results has not been obtained, all numerical examples that have been studied agree with these conclusions.

The results on mean fitness imply that, with synergistic epistasis, there appears to be an advantage to sexual reproduction over asexual reproduction arising from segregation as opposed to recombination. Although the evolutionary advantages of sex and recombination are often equated, this is an oversimplification, and several models have been proposed in which sex is favoured as a result of the consequences of Mendelian segregation. Lloyd (1980) and Hamilton (1980) independently suggested that temporal fluctuations in the direction of selection on alleles at a single locus, generated by host–parasite interactions, could provide an advantage to sex. Barton & Post (1986) showed that sib-competition in a spatially variable environment could also produce such an effect. Kirkpatrick & Jenkins (1989) have proposed a model in which directional selection leads to a segregational advantage to sex.

There is a further effect on mean fitness in the same direction in sexual populations with recombination, since, as discussed by Bulmer (1980, p. 158), recombination tends to break down the C_L term (although it also converts some of the C_{HW} term into C_L). Comparison of Tables 2 and 3 shows that free recombination ($r = \frac{1}{2}$ for all loci) is associated with an increase in the mean fitness of a sexual population with synergistic epistasis, and a decrease with diminishing returns epistasis. The effect is not as large as the difference between asexuality and sexuality without recombination (Fig. 1). For example, with the standard values of the selection parameters and $U = 1{\cdot}5$, the ratio of mean fitnesses for recombination and for no recombination is 1·23. Numerical studies of the equilibrium mean fitness of a population as a function of a number of chromosomes j and the map length of each chromosome l indicate that the mean fitness is very sensitive to chromosome number, and much less sensitive to map length unless the chromosome number or map length is small (Fig. 2). With 5 chromosomes or more, there is very little effect of map length unless $l \ll 0{\cdot}5$, and the mean fitness of the population is close to that for free recombination when $j \geqslant 5$ and $l \geqslant 0{\cdot}5$. Similarly, most of the effect of chromosome number occurs as j increases from 1 to 5. Nonetheless, with synergistic epistasis, mean fitness appears always to increase with j and l, as would be expected from the fact that increases in both variables lead to an increased efficiency of recombination in breaking down C_L. Again, a general analytical proof of these results has not been obtained, although the argument leading to eqn (13) suggests that they hold true under biologically reasonable conditions on the selection parameters α and β. With synergism, larger values of the linear term a seem to lead to lower equilibrium mean fitnesses, whereas larger values of the quadratic term b lead to higher mean fitnesses (Tables 2 and 3). This contrasts with the multiplicative case, where equilibrium mean fitness is independent of the selection coefficient (Kimura & Maruyama, 1966; Crow, 1970).

(ii) *The maintenance of sexual reproduction*

These effects of breeding system on mean fitness when there is epistasis have important implications for the evolutionary significance of sexual reproduction (Crow, 1988; Kondrashov 1982, 1988), and contradict the conclusion of other authors that there is no effect of breeding system on mean fitness. This conclusion was based on the assumption of multiplicative fitnesses

https://doi.org/10.1017/S0016672300025532 Published online by Cambridge University Press

(Hopf *et al.* 1987; Maynard Smith, 1978). If synergistic epistasis is widespread, then an asexual population derived from an ancestral population will equilibrate at a lower mean fitness than the ancestral population. The extent of this reduction depends on the degree of synergism and on the mutation rate per genome. As noted by Crow (1970), a very modest quadratic term is sufficient to change the equilibrium mean fitness drastically. As can be seen in Table 3, even if the quadratic coefficient b in the fitness expression for homozygotes is one-tenth of the linear coefficient a, the genetic load for a freely recombining population as measured by $-\ln \bar{w}$ is approximately $\frac{1}{2}U$ instead of U (the value for multiplicative fitnesses or asexuality), and is reduced only slightly as b increases (cf. Crow, 1970; Kimura & Maruyama, 1966). Thus, if $U \gg 2\ln 2 = 1{\cdot}4$, a mutation-free asexual variant entering a population with separate sexes will equilibrate at a fitness level less than one half that of the ancestral population. Initially, of course, such a variant will tend to increase in frequency due to the two-fold cost of sex (Lloyd, 1980; Maynard Smith, 1978). But from the numerical studies that have been carried out, it seems that equilibrium is approached from a starting state of zero mutations over a period of time of the order of a hundred generations or so, with mutation rates of this magnitude. Unless the size of the initial population is very small, the asexual lineage will be unlikely to have replaced the sexual population by the time that it has approached equilibrium. It will then diminish in frequency once its mean fitness has dropped below one-half that of the sexual population, and will eventually be eliminated.

If the initial population were hermaphroditic rather than dioecious, the cost of sex is only 1·5 in a random-mating population (Lloyd, 1980), and so $U \gg 2\ln 1{\cdot}5 = 0{\cdot}8$ would result in the maintenance of sex. From Tables 1 and 3, these conditions somewhat overestimate the advantage to the sexual population; critical U values of about 1·75 and 0·95 respectively are indicated with the standard selection parameters. With larger values of the selection parameters, mean fitnesses tend to be increased with synergistic selection, although the effect is not large. Studies of viability mutations in *Drosophila* by Mukai and his colleagues suggest that U is at least 0·8 for this species (Crow & Simmons, 1977); Kondrashov (1988) has argued that considerably higher values are plausible, since fitness components other than egg-to-adult viability are of great importance.

These conditions for the maintenance of sexuality are, in fact, conservative, since a new asexual variant is unlikely to arise in a mutation-free individual. If an asexual lineage arises in an individual carrying i mutations, the minimum number of mutations in its decendants is i, ignoring back-mutation. Thus, the initial mean fitness of the asexual lineage is $w(i)$, and the mean fitness of the asexual population steadily declines from this value as mutations accumulate. Exact calculations of the process of accumulation, using the equations of Kimura & Maruyama (1966), indicate that an asexual lineage in which $i > 0$ reaches an equilibrium with a mean number of mutations substantially larger than those displayed in Table 1, even if i is considerably less than the mean number of mutations per individual for the ancestral sexual population. The equilibrium mean fitnesses can be found as follows, by generalizing the mean fitness result of Kimura & Maruyama (1966). Their argument on the equilibrium frequency of the class with the lowest number of mutations implies that the equilibrium mean fitness of an asexual population started with i mutations is

$$\bar{w} = \exp -\{U + i(\alpha + \tfrac{1}{2}\beta i)\}. \qquad (28)$$

Given that the variance of a freely recombining sexual population is slightly less than the mean (Table 3), it is unlikely that i would be much less than $\bar{n} - 2\sqrt{\bar{n}}$, where $\bar{n}$ is the equilibrium number of mutations per individual for the sexual population. Thus, even under the most favourable circumstances, an initially rare asexual variant enjoying a two-fold fertility advantage will ultimately be eliminated from the population if the mean fitness given by the above expression with $i = \bar{n} - 2\sqrt{\bar{n}}$ is less than one-half the mean fitness of the sexual population.

Mean fitnesses obtained in this way are considerably lower than the Table 1 values. Table 7 gives the ratios of the maximum values of the equilibrium mean fitnesses of asexual lineages to the mean fitnesses of the sexual populations from which they are derived in this way. It will be seen that $U = 1$ is close to the threshold value needed to overcome a two-fold fertility advantage of asexuality. Fig. 5 displays the population trajectories of asexual variants introduced at low frequencies into sexual populations with different mutation rates, confirming that they are indeed eliminated when the suggested criterion is met. Given that asexual variants may often suffer often lower fertilities than their sexual competitors, due to imperfect functioning of cytological devices for bypassing meiosis (Lloyd, 1980), it seems that the

Table 7. *Equilibrium properties of asexual populations started with the minimum probable numbers of mutations per individual* (i)

U	i	$\bar{w}$	R*	$\bar{n}$	V
5·0	64	0·001	0·018	129·5	48·4
2·0	37	0·073	0·221	79·3	31·7
1·5	30	0·147	0·338	67·4	27·8
1·0	23	0·284	0·498	53·9	22·7
0·5	14	0·545	0·725	36·6	16·2
0·1	4	0·892	0·949	14·4	7·3

* R is the ratio of the equilibrium mean fitness for the asexual population to that for a freely recombining sexual population with the same parameters.

https://doi.org/10.1017/S0016672300025532 Published online by Cambridge University Press

process described here could play a major role in the maintenance of sex in higher organisms. If the population size is small, however, the initial frequency of the asexual variant may be sufficiently high that it spreads to fixation before its mean fitness declines below the level needed to overcome any fertility advantage (see Fig. 5). Large populations will thus be less vulnerable than small ones to invasion by asexual variants. This may be another reason for the well-known association of asexuality with sparse populations (Bell, 1982).

A somewhat different approach to the invasion of a sexual population by an asexual variant was taken by Kondrashov (1985), who assumed that asexual individuals initially had the same distribution of mutations per individual as the sexual population from which they were derived. In this case, the initial mean fitness of the asexual population would be the same as that for a population with the post-selection distribution of mutations characteristic of the sexual population from which it was derived. It would then decline asymptotically towards $\exp - U$. The initial fitness of such a population on the present model can readily be calculated from equation (A 2), using the same mean number of mutations as for the equilibrium sexual population and its post-selection variance. With the standard selection parameters, there is only a very small reduction in fitness to the asexuals, unless the mutation rate is implausibly high, in agreement with Kondrashov's results for his linear selection model (Kondrashov, 1985, Fig. 1), which has a comparable level of epistasis. This suggests that an immediate loss in fitness to asexual individuals is unlikely to be detectable by observation. This calculation is relevant

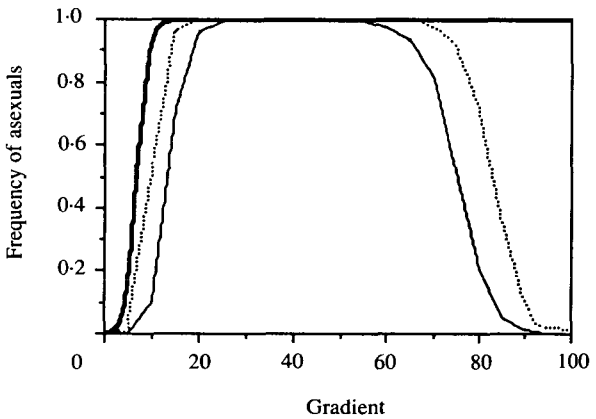

Fig. 5. Progress of an asexual variant introduced into a freely recombining sexual population with mutation rate $U = 1{\cdot}5$, and the standard selection parameters. Asexual females have a two-fold fertility advantage. The heavy line is for an initial frequency of asexual females of 10^{-3}, the dotted line for an initial frequency of 10^{-4}, and the full line for an initial frequency of 10^{-5}. The asexual variant is assumed to occur as a unique variant in a single female; the number of sexual females in the initial population is thus the reciprocal of the initial frequency of asexuals. Fixation of the asexual variant is assumed to occur if the frequency of sexual females falls below the initial frequency of asexuals.

to comparisons of the differences between arrays of sexual and asexual offspring produced experimentally in studies such as those of Kelley *et al.* (1988). They in fact found an approximately 40% higher fitness for sexually produced offspring of *Anthoxanthum odoratum* in the field, relative to offspring produced vegetatively. This result would not be expected on the present theory, and either reflects the operation of threshold selection with a high threshold and high mutation rate (cf. Kondrashov, 1985, Fig. 1), the effects of other selective forces maintaining sex, or simply the direct consequences of vegetative versus seed propagation. On the other hand experiments in *Drosophila* where suppression of genetic recombination was found to be associated with a fitness advantage (Kondrashov, 1988; Maynard Smith, 1978, chap. 5) are consistent with the present model, if other forces tending to produce a mutational load are operating in addition to synergistic selection (Maynard Smith, 1978).

(iii) *Evidence on the nature of fitness interactions*

A critical question concerns the extent to which synergistic fitness interactions occur in nature. First, it is useful to note that diminishing returns epistasis is compatible with equilibrium under mutation-selection balance only under a narrow range of parameter values, even if the modified fitness function with constant fitness for $n > -\alpha/\beta$ is used in order to avoid the artefact of fitnesses that increase with increasing number of mutations [section 2(i)]. The approximate threshold is $-\beta = \alpha^2/U$ [section 2(iv)]. If $-\beta$ exceeds this value, mutations will progressively accumulate, and mean fitness will decline indefinitely. Diminishing returns epistasis is unlikely to be observed among extant organisms, at least when the genome size is large. It is useful to note that the above relation corresponds to the 'error threshold' postulated by Eigen & Schuster (1979), in their hypercycle theory of the origin of self-replicating systems. Their model postulates a master genomic sequence of a self-replicating molecule; all mutational deviants from this, at any site in the sequence, are effectively assumed to be equally unfit (Maynard Smith, 1983; Nowak & Schuster, 1989). This is an extreme form of diminishing returns epistasis. Since there is no obvious biological justification for this mode of selection over one in which fitness is a decreasing function of the number of sites by which a mutated sequence differs from the master sequence, the necessity of invoking hypercycles to explain the evolution of relatively large genomes (Eigen & Schuster, 1979; Nee & Maynard Smith, 1990) is not apparent.

The tight constraint on diminishing returns epistasis suggests that multiplicative fitnesses or synergistic epistasis are likely to prevail. Direct evidence on the form of the relation between fitness and number of mutations is hard to obtain. While some authors have advocated truncation selection as a widespread mode

of selection that generates synergistic interactions (Kondrashov, 1988), there are essentially no direct data supporting its operation in nature. Data on the effects of minor detrimental mutations on egg-to-adult viability accumulated on the second chromosome of *D. melanogaster* suggest some weak synergism of the kind modelled here (Crow, 1970). Experiments in which the two major autosomes have simultaneously been made homozyous for wild-type chromosomes extracted from natural populations generally indicate weak or no synergistic effects (Clark, 1987; Crow & Simmons, 1977). Data on viability effects of the second and third chromosomes reported by Seager and Ayala (1982) suggest weak and doubtfully significant diminishing returns epistasis, however. Viabilities of chromosomal homozygotes relative to heterozygotes for the balancer chromosomes were estimated from the frequencies of double balancer, single balancer and wild-type homozygotes, segregating in cultures produced by intercrosses between flies simultaneously heterozygous for second and third chromosome balancers and wild-type chromosomes. Controls where the non-balancer individuals produced were heterozyous rather than homozygous for the chromosomes in question were also carried out, and used to calculate the homozygous viabilities relative to those of chromosomal heterozygotes. If the double balancer individuals perform disproportionately poorly in competition with wild-type heterozygotes, then the cross-product ratio *aBCd/AbcD* used as an index of the direction of epistasis between the viability effects of homozygosity for the two chromosomes will be biased in the direction of diminishing returns (lower case letters refer to crosses involving wild-type heterozygotes for each chromosome; upper case to crosses involving wild-type homozygotes; *a* refers to the frequency of wild-type flies, *b* and *c* refer to the frequency of single balancer genotypes, and *d* to the frequency of double-balancer genotypes). It seems likely that a similar criticism applies to the very large diminishing returns effect reported by Seager *et al.* (1982) for total fitness (where the effect of making the second and third chromosomes homozygous simultaneously was the same as that of making either homozygous singly). Such an extreme degree of diminishing returns epistasis is, indeed, theoretically incompatible with stability under mutation-selection balance (see above), and so one must infer that these data either have an artefactual explanation of the kind suggested, or are inapplicable to the question of mutation-selection balance. Thus, it would seem that greater weight should be attached to those *Drosophila* experiments which show synergistic effects. More data are clearly needed to resolve this question.

In summary, it is clear from the theoretical results that extremely small non-linear terms in the relation between log fitness and number of mutations can have large effects on equilibrium mean fitnesses. Such terms would be almost undetectable except in very large experiments. In groups other than *Drosophila*, careful studies are needed of the relationship between fitness components and the inbreeding coefficients of individuals produced in programmes of controlled inbreeding.

(iv) *Selection on recombination with epistasis*

The results presented in section [3[(iii) and (iv)] show that the effect of recombination on mean fitness does not always predict the direction of selection on a modifier of the rate of genetic recombination, since increased recombination can sometimes be selected against with synergistic epistasis despite the fact that mean fitness always seems to increase with increased recombination frequencies. [There are, of course, other precedents for a failure of mean fitness to predict the course of evolution of genes that modify the breeding system (Altenberg & Feldman, 1987; Karlin & McGregor, 1974).] For example, with the standard values of the selection parameters, it is apparent from Fig. 3 that a modifier increasing the rate of recombination away from zero is favoured with $U = 0{\cdot}1$. Table 4 shows that a modifier reducing the rate of recombination in a population with free recombination is favoured when $U = 0{\cdot}1$, whereas a gene totally suppressing recombination is eliminated. This suggests that an intermediate recombination frequency is evolutionarily stable under these conditions. For $U \geqslant 0{\cdot}5$, however, it seems that free recombination is selected for. More extensive calculations indicate that the critical value of U is approximately 0·29. The results of Table 5 confirm these conclusions, and enable the approximate location of the ESS map length to be determined as a function of the number of chromosomes, for modifiers located in the middle of a chromosome and which affect the map lengths of all chromosomes in the genome. For $U = 0{\cdot}1$, the ESS map length decreases with the number of chromosomes, as would be expected from the fact that the chromosome number has a large effect on the average frequency of recombination between pairs of genes. The location of the modifiers has little effect on the ESS map length (Table 6).

In his numerical studies of selection on recombination modifiers, Kondrashov (1984) also found that the direction of selection on recombination frequency depends on the mutation rate. His interpretation of this finding involves the notion that a recombination modifier has its primary effect on the variance of number of mutant genes per genome; this is borne out in the left-hand section of Table 4, where it can be seen that the asymptotic value of the change in mean associated with a change in recombination frequency caused by a modifier of small effect is much smaller than the change in variance [cf. eqn (21)]. In the case of threshold selection, if the mean number of mutations is above the threshold for truncation, an

increase in variance will increase the fraction of the population that fall to the left of the truncation point, and hence survive. Conversely, if the mean is below the threshold, then an increase in the variance will increase the fraction of the population that fall to the right of the threshold, and hence are eliminated by selection (Kondrashov's fig. 1). A high mutation rate implies a high mean copy number and thus a higher chance of being to the right of the truncation point. In the present case, the curve relating fitness to the number of mutations n has an inflection point, such that the absolute value of the slope is an increasing function of n to the left of the inflection point, and decreases with n to the right. If the mean is far to the left of the inflection point, then an increase in the variance of n results in the production of a higher frequency of extreme individuals with n values in the right-hand part of the distribution, where fitness declines more sharply. These individuals thus have a disproportionate effect on the mean fitness, compared with the increase in frequency of extreme individuals in the left-hand portion of the distribution, where fitness decreases slowly with copy number. The result is a net reduction in mean fitness. The converse holds if the mean is to the right of the inflection point.

More formally, partial differentiation of equation (A 2) with respect to the variance V yields the following approximate expression, valid when $\beta V \ll 1$ and $\alpha^2 < \beta$:

$$\frac{\partial \ln \bar{w}}{\partial V} \approx \frac{\beta}{2}\{(2\alpha\bar{n}+\beta\bar{n}^2)-1\}+\frac{\alpha^2}{2}.$$

Clearly, if $\bar{n}$ is very small, this expression is dominated by $(\alpha^2-\beta)/2$, which is less than zero on the above assumptions. For sufficiently large values of $\bar{n}$, the derivative is positive. Thus, an increase in variance alone can lead to reduction in mean fitness with small $\bar{n}$, corresponding to a small mutation rate. This calculation overestimates the range in which there is selection against increased recombination, since it ignores the increase in $\bar{n}$ associated with increased recombination under synergistic epistasis. For example, with the standard selection parameters and the equilibrium values of $\bar{n}$ for free recombination given in Table 3, the values of the derivative given by the above formula for $U = 0{\cdot}1$, $0{\cdot}5$ and 1 are $-3{\cdot}1 \times 10^{-4}$, $-1{\cdot}7 \times 10^{-4}$, and 5×10^{-5} respectively, whereas Table 4 indicates selection in favour of free recombination with $U = 0{\cdot}5$.

Kondrashov's (1984) numerical results for truncation selection in a sexual haploid population led him to suggest that there is selection for free recombination when the 'genome degradation rate' $v = U/\sqrt{V}$ exceeds 0·35. With the standard selection parameters of the present model, the critical U value of ·29 corresponds to $v = 0{\cdot}07$, a much smaller value than Kondrashov's truncation selection value. One possible explanation of this difference is that selection in a haploid organism is less favourable for the evolution of increased recombination than in a diploid. This was investigated by modifying the model to apply to a sexual haploid. The results of an investigation of the equilibrium population mean fitnesses and the ESS map lengths for different mutation rates for a haploid are shown in Table 8, for the standard values of the selection parameters (note that haploidy means that the dominance coefficient h is effectively 1 when calculating α and β, and that the U values are one-half those for the corresponding diploid case). The mean fitnesses tend to be somewhat higher than those in the diploid case, due to the stronger selection on mutations expressed in haploids. The behaviour of the direction of selection on recombination as a function of map length, number of chromosomes, and mutation rate

Table 8. *Selection on recombination modifiers in a haploid sexual population*

				$a = 0{\cdot}01$, $b = 0{\cdot}01$				
l	1	2	5	20	1	2	5	20
		$U = 0{\cdot}05$				$U = 0{\cdot}25$		
0·01	0·949	0·952	0·954	0·955	0·802	0·819	0·829	0·824
	4·6	$2{\cdot}8 \times 10^{-4}$	$4{\cdot}8 \times 10^{-5}$	$1{\cdot}4 \times 10^{-6}$	12	$1{\cdot}4 \times 10^{-3}$	$3{\cdot}3 \times 10^{-4}$	$4{\cdot}2 \times 10^{-5}$
0·50	0·953	0·954	0·955	0·955	0·819	0·827	0·832	0·835
	$5{\cdot}4 \times 10^{-1}$	$9{\cdot}9 \times 10^{-5}$	$-7{\cdot}4 \times 10^{-6}$	$-7{\cdot}3 \times 10^{-6}$	3·7	$2{\cdot}0 \times 10^{-3}$	$4{\cdot}2 \times 10^{-4}$	$3{\cdot}8 \times 10^{-5}$
1·0	0·954	0·954	0·955	0·955	0·824	0·830	0·833	0·835
	$1{\cdot}5 \times 10^{-1}$	$7{\cdot}7 \times 10^{-6}$	$-2{\cdot}0 \times 10^{-5}$	$-7{\cdot}5 \times 10^{-6}$	1·9	$1{\cdot}20 \times 10^{-3}$	$2{\cdot}6 \times 10^{-4}$	$2{\cdot}6 \times 10^{-5}$
		$U = 0{\cdot}5$				$U = 1{\cdot}0$		
0·05	0·656	0·685	0·703	0·712	0·443	0·484	0·510	0·523
	19	$2{\cdot}5 \times 10^{-3}$	$6{\cdot}6 \times 10^{-4}$	$1{\cdot}0 \times 10^{-4}$	27	$4{\cdot}3 \times 10^{-3}$	$1{\cdot}3 \times 10^{-3}$	$2{\cdot}3 \times 10^{-4}$
0·50	0·683	0·699	0·708	0·713	0·478	0·502	0·517	0·525
	7·0	$4{\cdot}9 \times 10^{-3}$	$1{\cdot}3 \times 10^{-3}$	$1{\cdot}7 \times 10^{-4}$	12	$1{\cdot}1 \times 10^{-2}$	$3{\cdot}3 \times 10^{-3}$	$5{\cdot}2 \times 10^{-4}$
1·0	0·693	0·704	0·711	0·714	0·493	0·510	0·520	0·526
	4·0	$3{\cdot}6 \times 10^{-3}$	$9{\cdot}4 \times 10^{-4}$	$1{\cdot}4 \times 10^{-4}$	7·8	$9{\cdot}1 \times 10^{-3}$	$2{\cdot}7 \times 10^{-3}$	$4{\cdot}8 \times 10^{-4}$

The upper entries in each row are the equilibrium mean fitnesses; the lower entries are the selection gradients on recombination for a modifier located in the middle of a chromosome.

https://doi.org/10.1017/S0016672300025532 Published online by Cambridge University Press

differs little from that with diploidy. The values of the genome degradation rate are close to those for the corresponding diploid cases e.g. with haploidy and $U = 0{\cdot}25$, $v = 0{\cdot}14$ compared with $v = 0{\cdot}10$ for the diploid case with $U = 0{\cdot}5$. It thus does not seem likely that haploidy is the cause of the discrepancy.

Another possibility is that truncation selection is less favourable to selection for increased recombination than the less extreme mode of synergistic selection assumed here. At first sight, this seems contrary to expectation, based on the fact that more extreme epistasis seems to select more strongly for recombination (e.g. Table 4). However, a study of the relationship between the critical value of U for selecting for free recombination, and the degree of synergism as measured by the ratio b/a, indicates that the critical value of U can decrease with the degree of synergism. With $a = 0{\cdot}02$ and $b = 0{\cdot}001$, for example, U need exceed only $0{\cdot}2$ ($v = 0{\cdot}03$) for free recombination to be evolutionarily stable, compared with a value of $0{\cdot}29$ with $b = 0{\cdot}01$ ($v = 0{\cdot}07$). With $a = 0{\cdot}01$ and $b = 0{\cdot}1$ on the other hand, the critical value of U is $0{\cdot}32$ ($v = 0{\cdot}12$). This is consistent with Kondrashov's finding that, with his 'intermediate' selection model (in which fitness is a quadratic function of number of mutations), the critical value of v is approximately $0{\cdot}1$ (Kondrashov, 1984, p. 206). The probable explanation for this somewhat counterintuitive effect of the level of synergism on the critical mutation rate is that a larger quadratic term lowers the equilibrium mean number of mutations (see Table 3). Hence, the distribution of the number of mutations is shifted to the left, towards the region where a reduction in variance lowers mean fitness. At all events, with the present model selection for free recombination occurs at much smaller mutation rates and genome degradation rates than with truncation selection.

Despite these complications, the analysis of the conditions for spread of modifiers of recombination at the low recombination limit [section 3(ii)] indiates that, with synergistic epistasis, non-zero recombination rates are favoured whatever the mutation rate, under biologically reasonable conditions. This is consistent with the numerical findings presented in Table 5. Kondrashov (1984) found examples in which an allele causing zero recombination could approach fixation. In contrast to the present model (where an allele associated with near-zero recombination is itself closely linked to all selected loci), he assumed free recombination between the modifier locus and all the selected loci, even when one modifier allele at the modifier locus caused zero recombination. As pointed out by Nee (1988) this is somewhat unrealistic, and tends to underestimate the strength in favour of increased recombination. As can be seen from Fig. 3 and Table 4, with the present model there is quite strong selection in favour of modifiers increasing the recombination rate away from zero, and against suppressors of recombination in a population with free recombination, over a wide range of mutation rates. The strength of selection increases markedly with U, as expected. The difference in behaviour from the case of modifiers of small effect in populations with a high frequency of recombination arises from the fact that the changes in mean number of mutations associated with the invading alleles are much greater than with freer recombination, and overhelm the effect of variance on mean fitness [see eqns (16), (21), and Table 4].

The analysis of the strength of selection on the proportional effect of modifiers on map length (ϵ), as measured by the selection gradient (Lande, 1976), indicates that selection to increase the map length of each chromosome is strongest in a genome with a single chromosome, and diminishes rapidly as the number of chromosomes increases (see Tables 5–8, and Fig. 4), even when the initial map length is as small as $0{\cdot}05$. As noted earlier, the strength of selection on map length itself is $1/l$ times the value for ϵ, so that it is substantially higher than that shown when l is small. Above a map length of 1 or so, the selection gradient falls off rather slowly with map length (Table 5). These findings are in general agreement with those of Kondrashov (1984), on a modifier that is unlinked to the selected loci. Selection is stronger in a haploid population than in the case of a diploid population with the same mutation rate per haploid genome, reflecting the stronger selection against deleterious alleles with haploidy. This effect may be negated by the fact that haploid species tend to have smaller genome sizes than diploids (Cavalier-Smith, 1985), and so values of the order of $0{\cdot}05$ or less may be more appropriate for haploids other than RNA viruses, which appear to have very high mutation rates per nucleotide (Nee & Maynard Smith, 1990; Pressing and Reanney, 1984).

With diploidy and $U = 1$, selection gradients on ϵ of the order of 10^{-4} to -10^{-5} are found for a map length of 1 when the number of chromosomes is greater than 5 (Table 5). With haploidy, the corresponding selection gradients are of the order of 10^{-3} to 10^{-4}. The selection gradients for $l > 1$ and $j = 20$ are in approximate agreement with the selection gradients on mean recombination frequency with free recombination, given in Table 4, when the appropriate change of variable is made. These results indicate that the strength of selection maintaining large map lengths is weak in genomes with more than one chromosome, unless the per genome mutation rate is very high, in agreement with the results of Kondrashov (1984). This makes it difficult to judge the plausibility of synergistic selection selection on deleterious mutations as a mechanism for maintaining the map lengths of 100 centimorgans or more that are characteristic of higher organisms. Selection gradients of the magnitude observed are ineffective in populations of less than 10^4 or so individuals (Lande, 1976), although this number refers to the species population size rathe than local

population size in species which have even a modest amount of migration between populations (Crow & Kimura, 1970, chap. 9). Clearly, the selection coefficients at individual loci affecting recombination rates will be very low with these kinds of selection gradients, so that allele frequencies at modifier loci will be strongly affected by mutation and drift, and will often be far from their equilibria under selection. This may account for the genetic variability often observed for recombination frequencies (Brooks, 1988).

The present model also enables the strength of selection on chromosome number to be studied. A centric fusion between two chromosomes of equal length behaves formally as a dominant gene that reduces the frequency of crossing over between pairs of loci that were formerly located on two different chromosomes (Charlesworth, 1985). The selection coefficient on a rare centric fusion provides a measure of the intensity of selection for maintaining high chromosome numbers. Calculations for the standard values of the selection parameters indicate that centric fusions are weakly selected against with synergistic selection. For example, with $U = 1$ and a map length of 1, the selection coefficients against a centric fusion are $-1{\cdot}8 \times 10^{-4}$, $-1{\cdot}6 \times 10^{-5}$ and $-5{\cdot}5 \times 10^{-7}$ with 2, 5 and 20 chromosomes respectively. Thus, although the number of chromosomes has a major effect on the equilibrium mean fitness of the population (Fig. 2), this does not seem to translate into a strong selection force for maintaining chromosome number. It is thus not surprising that centric fusions are a major mode of karyotypic evolution in a variety of groups (White, 1973).

Appendix

(i) *Effect of selection on the distribution of a normal variate within a generation*

Assume that the distribution $\phi(n)$ before selection is normal, with mean $\bar{n}$ and variance V. The selection function of equation (1) is assumed to be applied to this distribution. Transforming to the standardized normal deviate $z = n - \bar{n})/\sqrt{V}$, we have $w(z) \propto \exp -(\tilde{\alpha}z + \frac{1}{2}\tilde{\beta}z^2)$, where $\tilde{\alpha} = (\alpha + \beta\bar{n})\sqrt{V}$, and $\tilde{\beta} = \beta V$. The distribution of z after selection is

$$\left.\begin{aligned}\phi^*(z) &= \frac{w(z)\,\phi(z)}{\int_{-\infty}^{\infty} w(z)\,\phi(z)\,\mathrm{d}z} \\ &= \frac{\exp -\{\tilde{\alpha}z + \frac{1}{2}(1+\tilde{\beta})\,z^2\}}{\int_{-\infty}^{\infty} \exp -\{\tilde{\alpha}z + \frac{1}{2}(1+\tilde{\beta})\,z^2\}\,\mathrm{d}z}.\end{aligned}\right\} \quad \text{(A 1)}$$

Transforming to $u = z(1+\tilde{\beta})^{\frac{1}{2}} + \tilde{\alpha}(1+\tilde{\beta})^{-\frac{1}{2}}$, we obtain a standardized normal deviate on substitution into this equation. This implies that z after selection is normally distributed with mean $-\tilde{\alpha}/(1+\tilde{\beta})$ and variance $1/(1+\tilde{\beta})$. Hence, $\bar{n}$ is changed by $\Delta_s\bar{n} = -(\alpha + \beta\bar{n})V/(1+\beta V)$, and V by $\Delta = -\beta V^2/(1+\beta V)$. The mean fitness of the population is given by $1/\sqrt{2\pi}$ times the denominator of equation (A 1), and reduces to

$$\bar{w} = (1+\beta V)^{-\frac{1}{2}} \exp \frac{1}{2(1+\beta V)}\{\alpha^2 V - 2\alpha\bar{n} - \beta\bar{n}^2\}. \quad \text{(A 2)}$$

(ii) *Covariances between loci in gametes carrying modifier allele M_2*

Consider first the case where the modifier locus is outside all of the selected loci. For loci i and j, the state of an M_1/M_2 individual can be represented as $M_1 X_{i1} X_{j1}/M_2 X_{i2} X_{j2}$, where X_{i1} indicates the state of the allele at locus i contributed by the M_1 gamete, X_{j1} indicates the state of the allele at locus j contributed by the M_1 gamete, etc. ($X = 0$ for wild type, $X = 1$ for mutant). A gamete produced by this individual and which carries M_2 can have the following genotypic states, with probabilities calculated on the basis of no interference:

$X_{i2} X_{j2} (P = [1-\rho_i][1-r_{ij2}])$,

$X_{i2} X_{j1} (P = [1-\rho_i] r_{ij2})$,

$X_{i1} X_{j2} (P = \rho_i r_{ij2})$,

$X_{i1} X_{j1} (P = \rho_i [1-r_{ij2}])$.

The new mean value of X_i for these gametes, ignoring the effects of selection, is given by

$$\bar{X}'_{i2} = (1-\rho_i)\,\bar{X}_{i2} + \rho_i\,\bar{X}_{i1} = \bar{X}_i + (1-\rho_i)\,\delta\bar{X}_i, \quad \text{(A 3)}$$

where $\delta\bar{X}_i$ is the deviation of the mean of X_i for M_2 gametes from the mean for M_1 gametes. It will be assumed that the effect of the recombination modifier is sufficiently small that second-order terms in the $\delta\bar{X}_i$ can be ignored. This means that deviations from Hardy-Weinberg frequencies in the population of M_1/M_2 individuals, induced by differences between the M_1 and M_2 gametes can be neglected, enabling the change in covariance between X_i and X_j in M_2 gametes to be calculated on lines similar to those used in the text for a homogeneous population.

The new value of the covariance between X_i and X_j in M_2 gametes following recombination, but ignoring selection, is given by

$$\left.\begin{aligned}C'_{ij2} &= E\{X'_i X'_j\} - \bar{X}'_{i2}\,\bar{X}'_{j2} = (1-r_{ij2}) \\ &\quad \times [(1-\rho_i)\,E\{X_{i2} X_{j2}\} + \rho_i\,E\{X_{i1} X_{j1}\}] \\ &\quad + r_{ij2}[(1-\rho_i)\,E\{X_{i2} X_{j1}\} + \rho_i\,E\{X_{i1} X_{j2}\}] \\ &\quad - \bar{X}'_{i2}\,\bar{X}'_{j2} \\ &= (1-r_{ij2})[(1-\rho_i)(C_{ij2} + \bar{X}_{i2}\,\bar{X}_{j2}) \\ &\quad + \rho_i(C_{ij1} + \bar{X}_{i1}\,\bar{X}_{j1})] + r_{ij2}[(1-\rho_i)(\bar{X}_{i2}\,\bar{X}_{j1}) \\ &\quad + \rho_i(\bar{X}_{i1}\,\bar{X}_{j2})] - \bar{X}'_{i2}\,\bar{X}'_{j2}.\end{aligned}\right\} \quad \text{(A 4)}$$

https://doi.org/10.1017/S0016672300025532 Published online by Cambridge University Press

Substituting from eqn (A 3) and neglecting terms in $\delta\bar{X}_i\,\delta\bar{X}_j$, this reduces after simplification to

$$C'_{ij2} \approx (1-r_{ij2})[(1-\rho_i)\,C_{ij2}+\rho_i\,C_{ij1}]. \qquad \text{(A 5}a\text{)}$$

In addition to the effect of recombination, the effect of gene frequency changes on the covariance need to be included [see section 2(iv) of the text], together with the direct effect of selection in changing linkage disequilibrium. Let hs_2 be the selection coefficient against heterozygous mutant alleles in the population of M_1/M_2 individuals ($hs_2 = -\Delta_s\bar{n}_2/\bar{n}_2$, where $\Delta_s\bar{n}_2$ is calculated using $\bar{n}_2$ and V_2 in the equation for the change in mean due to selection). There is a change of $-2hs_2\,C_{ij2}$ due to change in the disequilibrium among the M_2 gametes, and $-2hs_2\,C_{ij1}$ due to change in the disequilibrium among the M_1 gametes. These changes are calculated from the portion of the general expression for change in linkage disequilibrium in a two-locus system, considering only gametes that are non-recombinant with respect to loci i and j. The contribution from such gametes is weighted by the reciprocal of their frequency $(1-r_{ij})$, so that the net contribution has a weight of one [cf. Kojima & Lewontin 1970, eqn (15)]. Hence, M_2 will be associated with the indirect change in covariance among M_2 gametes with probability $1-\rho_i$, and with the change among M_1 gametes with probability ρ_i.

The direct effect of selection on M_1/M_2 individuals can be calculated in the same way as for eqn (8) of the text, noting that the relevant change in variance due to linkage disequilibrium is $\Delta_2 = -\beta V_2^2/(1+\beta V_2)$.

The final expression for the new value of C_{ij2} is thus

$$C'_{ij2} \approx (1-r_{ij2})[(1-\rho_i)\,C_{ij2}+\rho_i\,C_{ij1}]-2(1-\rho_i) \times hs_2\,C_{ij2}-2\rho_i hs_2\,C_{ij1}+\frac{\Delta_2}{4m(m-1)}. \qquad \text{(A 5}b\text{)}$$

This can be further simplified by using the first terms in the Taylor's expansion of the quantities in question, writing δz for the deviation of a quantity z in M_2 gametes or M_1/M_2 individuals from its corresponding value for the M_1/M_1 population. Substituting into equation (A 5b), assuming that C_{ij1} is at the equilibrium value given by equation (10), and neglecting second-order terms, the asymptotic value of δC_{ij} is given by

$$\delta C_{ij} = \frac{1}{\rho_i+(1-\rho_i)(r_{ij1}+2hs_1)} \times\left\{\frac{\delta\Delta}{4m(m-1)}-(\delta r_{ij}+2\delta hs)\,C_{ij1}\right\}. \qquad \text{(A 6)}$$

This approach is easily extended to the case when the modifier locus is located between a pair of selected loci, i and j. The frequencies of recombination between M and the other loci are ρ_i and ρ_j, and the frequency of recombination between loci i and j is $r_{ij2} = \rho_i+\rho_j-2\rho_i\rho_j$. In this case, the following M_2 gametes are produced:

$X_{i2}\,X_{j2}(P = [1-\rho_i][1-\rho_j])$,

$X_{i2}\,X_{j1}(P = [1-\rho_i]\rho_j)$,

$X_{i1}\,X_{j2}(P = \rho_i[1-\rho_j])$,

$X_{i1}\,X_{j1}(P = \rho_i\rho_j)$.

Equation (A 3) gives the mean of X_i among M_2 gametes following recombination. The analogue of eqn (A 5a) is:

$$C'_{ij2} \approx (1-\rho_i)(1-\rho_j)\,C_{ij2}+\rho_i\rho_j\,C_{ij1}. \qquad \text{(A 7)}$$

The indirect effect on covariance of changes in gene frequencies can be found as before, noting that the probability that M_2 is associated with an $X_{i2}\,X_{j2}$ gamete is $(1-\rho_i)(1-\rho_j)$, and the probability that M_2 is associated with an $X_{i1}\,X_{j1}$ gamete is $\rho_i\rho_j$. The net indirect change in C_{ij2} is thus

$$\frac{-2hs(\rho_i\rho_j\,C_{ij1}+[1-\rho_i][1-\rho_j]\,C_{ij2})}{1-r_{ij2}}.$$

The direct effect of selection on covariance can be calculated exactly as before.

These considerations yield the analogue of eqn (A 6) for the asymptotic values of δC_{ij}:

$$\delta C_{ij} \approx \frac{1}{r_{ij1}+\rho_i\rho_j+\dfrac{2hs_1(1-\rho_i)(1-\rho_j)}{(1-r_{ij1})}} \times\left\{\frac{\delta\Delta}{4m(m-1)}-(\delta r_{ij}+2\delta hs)\,C_{ij1}\right\}. \qquad \text{(A 8)}$$

Given that M_1/M_2 individuals are formed from the fusion of M_1 and M_2 gametes, we have

$$\delta V-\delta\bar{n} \approx 2\sum_{i<j}\delta C_{ij}. \qquad \text{(A 9)}$$

(iii) *Mean number of mutations in M_1/M_2 individuals*

The asymptotic value of $\delta\bar{n}$ can be found as follows. Modifying eqn (A 3) to include the effects of selection and mutation, and assuming that each locus contributes equally to the change in mean due to selection, we have

$$\bar{X}'_{i2} = \bar{X}_{i1}+(1-\rho_i)\,\delta\bar{X}_i+\frac{\Delta_s\bar{n}_2+U}{2m}. \qquad \text{(A 10)}$$

Writing $\Delta_s\bar{n}_2 = \Delta_s\bar{n}_1+\delta(\Delta_s\bar{n})$, and assuming that the M_1/M_1 population is at equilibrium, we obtain the following expression

$$\left.\begin{aligned}\rho_i\,\delta\bar{X}_i &= \frac{\Delta_s\bar{n}_1+U+\delta(\Delta_s\bar{n})}{2m}\\ &= \frac{\delta(\Delta_s\bar{n})}{2m}.\end{aligned}\right\} \qquad \text{(A 11)}$$

Noting that $\delta\bar{n} = \sum\delta\bar{X}_i$, and writing ρ_H for the harmonic mean of the ρ_i, we obtain

$$\delta\bar{n} \approx -k_1\,\delta V, \qquad \text{(A 12)}$$

https://doi.org/10.1017/S0016672300025532 Published online by Cambridge University Press

where

$$k_1 = \frac{(\alpha+\beta\bar{n}_1)\delta V}{(1+\beta V_1)\{2\rho_H(1+\beta V_1)+\beta V_1\}}$$

$$= \frac{U\delta V}{V_1\{2\rho_H(1+\beta V_1)+\beta V_1\}}.$$

Equations (A 6), (A 8), (A 9) and (A 12) completely determine $\delta\bar{n}$ and δV. In turn, these equations enable the mean fitness of M_2/M_1 individuals to be calculated from eqn (A 2).

This work was supported by NSF grant BSR-8817976. I thank Deborah Charlesworth and Nicholas Barton for useful discussions. Michael Bulmer, James Crow, Alexei Kondrashov, Russell Lande, Sean Nee and an anonymous referee provided valuable suggestions for improving the manuscript.

References

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

Barton, N. H. & Post, R. J. (1986). Sibling competition and the advantage of mixed families. *Journal of Theoretical Biology* **120**, 381–387.

Bell, G. (1982). *The Masterpiece of Nature*. London, Croom-Helm.

Brooks, L. D. (1988). The evolution of recombination rates. In *The Evolution of Sex* (ed. B. R. Levin and R. E. Michod), pp. 87–106. Sunderland, Mass.: Sinauer.

Bulmer, M. G. (1980). *The Mathematical Theory of Quantitative Genetics*. Oxford: Oxford University Press.

Cavalier-Smith, T. (1985). *The Evolution of Genome Size*. Chichester; John Wiley.

Chao, L. (1988). Evolution of sex in RNA viruses. *Journal of Theoretical Biology* **133**, 99–112.

Charlesworth, B. (1985). Recombination, genome size and chromosome number. In *The Evolution of Genome Size* (ed. T. Cavalier-Smith), pp. 489–513. Chichester: John Wiley.

Charlesworth, B. (1989). The evolution of sex and recombination. *Trends in Ecology and Evolution* **4**, 264–267.

Clark, A. G. (1987). A test of multilocus interaction in *Drosophila melanogaster*. *American Naturalist* **130**, 283–299.

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

Crow, J. F. (1988). The importance of recombination. In *The Evolution of Sex* (ed. B. R. Levin and R. E. Michod), pp. 56–73. Sunderland, Mass.: Sinauer.

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.

Elgen, M. & Schuster, P. (1979). *The Hypercycle*. Berlin: Springer-Verlag.

Feldman, M. W., Christiansen, F. B. & Brooks, L. D. (1980). Evolution of recombination in a constant environment. *Proceedings of the National Academy of Sciences USA* **77**, 4824–4827.

Haigh, J. (1978). The accumulation of deleterious genes in a population. *Theoretical Population Biology* **14**, 251–267.

Haldane, J. B. S. (1919). The combination of linkage values and the calculation of distance between loci of linked factors. *Journal of Genetics* **8**, 299–309.

Hamilton, W. D. (1980). Sex versus non-sex versus parasite. *Oikos* **35**, 282–290.

Hopf, R., Michod, R. E. & Sanderson, M. (1987). On the effect of reproductive system on mutation load and the number of deleterious mutations. *Theoretical Population Biology* **33**, 243–265.

Karlin, S. & McGregor, J. (1974). Towards a theory of the evolution of modifier genes. *Theoretical Population Biology* **5**, 59–103.

Kelley, S. E., Antonovics, J. & Schmitt, J. (1988). A test of the short-term advantage of sexual reproduction. *Nature* **331**, 714–716.

Kimura, M. & Maruyama, T. (1966). The mutational load with epistatic gene interactions in fitness. *Genetics* **54**, 1303–1312.

Kirkpatrick, M. & Jenkins, C. D. (1989). Genetic segregation and the maintenance of sexual reproduction. *Nature* **339**, 300–301.

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.

Kondrashov, A. S. (1985). Deleterious mutations as an evolutionary factor. II. Facultative apomixis and selfing. *Genetics* **111**, 635–653.

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

Lande, R. S. (1976). Natural selection and random genetic drift in phenotypic evolution. *Evolution* **30**, 314–334.

Lloyd, D. G. (1980). Benefits and handicaps of sexual reproduction. *Evolutionary Biology* **13**, 69–111.

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

Maynard Smith, J. (1983). Models of evolution. *Proceedings of the Royal Society of London* **B219**, 315–325.

Michod, R. E. & Levin, B. R. (1988). *The Evolution of Sex*. Sunderland, Mass.: Sinauer.

Morton, N. E. (1955). Sequential tests for the detection of linkage. *American Journal of Human Genetics* **7**, 277–318.

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

Nee, S. (1988). Deleterious mutation and genetic recombination. *Nature* **331**, 308.

Nee, S. (1989). On the evolution of sex in RNA viruses. *Journal of Theoretical Biology* **138**, 407–412.

Nee, S. & Maynard Smith, J. (1990). The evolutionary biology of molecular parasites. *Parasitology* (in the press).

Nowak, M. & Schuster, P. (1989). Error thresholds of replication in a finite population- mutation frequencies and Muller's ratchet. *Journal of Theoretical Biology* **137**, 375–398.

Pressing, J. & Reanney, D. C. (1984). Divided genomes and intrinsic noise. *Journal of Molecular Evolution* **20**, 135–146.

Redfield, R. J. (1988). Evolution of bacterial transformation: is sex with dead cells ever better than no sex at all? *Genetics* **119**, 213–221.

Seager, R. D. & Ayala, F. J. (1982). Chromosome interactions in *Drosophila melanogaster*. I. Viability studies. *Genetics* **102**, 467–483.

Seager, R. D., Ayala, F. J. & Marks, R. W. (1982). Chromosome interactions in *Drosophila melanogaster*. II. Total fitness. *Genetics* **102**, 485–502.

Simmons, M. J. & Crow, J. F. (1977). Mutations affecting fitness in *Drosophila* populations. *Annual Review of Genetics* **11**, 49–78.
Stearns, S. C. (1987). *The Evolution of Sex and its Consequences*. Basel: Birkhaüser.
Thomson, G. (1977). The effect of a selected locus on linked neutral loci. *Genetics* **85**, 753–788.
White, M. J. D. (1973). *Animal Cytology and Evolution*, 3rd edn. Cambridge, UK: Cambridge University Press.