THE RATE OF PRODUCTION OF RECOMBINANTS BETWEEN LINKED GENES IN FINITE POPULATIONS

SAMUEL KARLIN, JAMES McGREGOR, and WALTER F. BODMER
STANFORD UNIVERSITY

1. Introduction

Problems concerning the rate of production of recombinants between linked loci in finite populations are important in applications to natural and artificial populations (for a review see [1]). We shall be mainly concerned with applications to the theory of breeding experiments with laboratory populations. Closely related (but not identical) problems arise in the study of the speed of evolution in wild populations.

We consider two linked loci in a diploid population with alleles A, a and B, b at the first and second locus, respectively. The normal or "wild type" gamete is ab but the mutant chromosomes Ab and aB are also assumed to be present in the population. The gametic output of a double heterozygote aB/Ab contains the four types, Ab, aB, AB, ab in the proportions $(1 - r)/2$, $(1 - r)/2$, $r/2$, $r/2$, where r is the recombination fraction, $0 \leqq r \leqq 1/2$.

Recombination is one of the main constituent processes of evolution. To analyze the relative importance of recombination one asks questions of the following kind. Suppose the normal gamete ab occurs with a high frequency in a population while the mutant gametes Ab, aB are maintained at low frequencies by mutation selection balance. Then, in the long time scale of evolution, the eventual formation of an AB recombinant is certain. How long is the time between successive appearances of an AB recombinant? If the population is large, a satisfactory answer for this problem can be found from simple deterministic models. Our main interest lies with the similar problems which arise in finite captive populations.

The concrete problem we investigate takes the following form. Consider a two locus diallelic population consisting initially of N diploid individuals including only the genotypes Ab/Ab, Ab/aB, aB/aB. Each generation is produced by random sampling from the gametic output of the preceding generation, maintaining a fixed population size of N. An individual with a recombinant gamete may eventually appear, but this outcome is not certain. If no recombination ever occurs (for example, when $r = 0$), then there is effectively only a single locus

Research supported in part by Contract NIH USPHS 10452 at Stanford University.

and if there are no mutation or migration pressures, it is clear that ultimate fixation is achieved for one of the two homozygotes Ab/Ab or aB/aB. Due to the limitation of finite population size, even when recombination is possible, fixation may occur prior to the formation of a recombinant type. Our aim is to determine the probability R that a recombinant appears before fixation, and also to study the distribution of the time involved until it appears.

The model proposed to describe the fluctuations of the gamete types aB and Ab is given the canonical formulation [3], [7]. Specifically, suppose the current generation is composed of N diploid individuals with gametic numbers shown in Table I.

TABLE I

GAMETIC NUMBERS OF N DIPLOID INDIVIDUALS

Total Population	$2N$
Ab	i
aB	$2N - i$

We determine the next generation by random sampling from the gametic output of the present generation. If N is not too small, the genotype frequencies of the next generation can be obtained with sufficient accuracy from the gene frequencies of the present generation by assuming a Hardy-Weinberg distribution. If the makeup of the present population is as indicated in table I, then the genotypes Ab/Ab, aB/Ab, aB/aB occur in the present generation with the Hardy-Weinberg frequencies

$$\left(\frac{i}{2N}\right)^2, \quad 2\frac{i}{2N}\left(1-\frac{i}{2N}\right), \quad \left(1-\frac{i}{2N}\right)^2. \tag{1.1}$$

A gamete, selected at random from the output of the present generation, is therefore of type Ab with probability

$$\begin{aligned} p_i &= \left[N\left(\frac{i}{2N}\right)^2 2 + N\,2\frac{i}{2N}\left(1-\frac{i}{2N}\right)(1-r)\right]\Big/2N \\ &= \frac{i}{2N} - r\frac{i}{2N}\left(1-\frac{i}{2N}\right), \end{aligned} \tag{1.2}$$

of type aB with probability

$$q_i = 1 - \frac{i}{2N} - r\frac{i}{2N}\left(1-\frac{i}{2N}\right), \tag{1.3}$$

and is of recombinant type (ab or AB) with probability $1 - p_i - q_i$. When $2N$ gametes are chosen, by repeated sampling, to form the N individuals of the next generation, the probability that j gametes of type Ab and $2N - j$ of type aB are obtained is

$$P_{ij} = \binom{2N}{j} p_i^j q_i^{2N-j}, \qquad j = 0, 1, \cdots, 2N. \tag{1.4}$$

The probability that no recombinant gametes are obtained is therefore

$$\sum_{j=0}^{2N} P_{ij} = (p_i + q_i)^{2N} = \left[1 - 2r\frac{i}{2N}\left(1 - \frac{i}{2N}\right)\right]^{2N}, \tag{1.5}$$

and the probability that one or more recombinants appear is

$$1 - \left[1 - 2r\frac{i}{2N}\left(1 - \frac{i}{2N}\right)\right]^{2N}. \tag{1.6}$$

The Markov chain with transition probability matrix (1.4) ends either with the appearance of a recombinant type or with fixation. It is clear that one of these two events ultimately prevails with probability one. Our objective is to determine the probability R of recombination before fixation as a function of the initial gametic frequencies, the recombination fraction r, the population size N, selective values and the mating system. This problem can be solved formally as follows. Let $u = (u_0, u_1, \cdots, u_{2N})$ be the unique vector with components satisfying the linear system of equation

$$u_i = \sum_{j=0}^{2N} P_{ij}u_j, \qquad i = 0, 1, 2, \cdots, 2N, \tag{1.7}$$

coupled with the end conditions $u_0 = u_{2N} = 1$. A little reflection convinces one that u_i is the probability that the population fixes and therefore,

$$R = 1 - u_i, \tag{1.8}$$

where i is the initial number of the Ab gamete. To invert the matrix of (1.7) is an impossible task analytically, although numerically feasible but prohibitive for N large. To uncover qualitative insights we will pass from the Markov chain induced by the transition probability matrix (1.4) to an approximating diffusion process and solve the problem in this context. The general validity of diffusion approximations in studying finite population stochastic genetic models is discussed in Ewens [2] (see also Kimura [5], and Karlin and McGregor [4]).

The diffusion process in the present circumstance has the novel feature of containing a killing term. More precisely, we will establish later that $u(x)$, the probability of fixation, where x denotes the initial frequency of the Ab gamete satisfies the differential equation

$$\frac{1}{2}V(x)\frac{\partial^2 u}{\partial x^2} + M(x)\frac{\partial u}{\partial x} - K(x)u = 0, \qquad 0 < x < 1, \tag{1.9}$$

where V and M represent the usual infinitesimal variance and mean displacement. The term $K(x)$ corresponds to the rate of killing at x due to the contingency of recombination. In solving (1.9) it is essential to impose the boundary conditions $u(0) = u(1) = 1$ corresponding to the event of certain fixation. Numerical computation indicates that for N large (even $N = 10$) the solution

obtained from (1.9) fits remarkably well the exact solution determined from the equations (1.7).

Let $R(x) = 1 - u(x)$, with $0 < x < 1$, denote the probability of a recombinant appearing prior to fixation when the initial frequency of the Ab gamete is $x = i/2N$. From analysis of the relevant diffusion equations we obtain the explicit formula

$$R(x) = 1 - \frac{\sinh 4x\sigma^{1/2} + \sinh 4(1-x)\sigma^{1/2}}{\sinh 4\sigma^{1/2}}, \tag{1.10}$$

where σ is defined by the relation

$$r = \frac{\sigma}{N^2}. \tag{1.11}$$

The derivation of (1.10) is given in section 2.

We see immediately that when σ is large, that is, $r \gg 1/N^2$, then R is approximately 1 and recombination is virtually certain. On the other hand, if σ is small then $R(x) \sim 0$ and fixation almost certainly occurs before recombination.

If we are interested in the creation of the AB recombinant gamete, while ab gametes are discarded when formed, then the formula (1.10) has to be modified to

$$R(x) = 1 - \frac{\sinh x\,(8\sigma)^{1/2} + \sinh\,(1-x)(8\sigma)^{1/2}}{\sinh\,(8\sigma)^{1/2}}. \tag{1.12}$$

In particular, we have the especially simple relation

$$R\left(\frac{1}{2}\right) = 1 - \frac{1}{\cosh(2\sigma)^{1/2}} = 1 - \frac{1}{\cosh N(2r)^{1/2}} \sim 1 - 2e^{-N(2r)^{1/2}} \tag{1.13}$$

if $N(2r)^{1/2}$ is not too small. For any initial x, where $0 < x < 1$, it is easy to calculate the size of the population required to guarantee with a given probability that a recombinant type will be created. For example, for $x = 1/2$ the relation (1.13) is easily solved for N to yield

$$N = \frac{1}{(2r)^{1/2}} \cosh^{-1} \frac{1}{1 - R(\frac{1}{2})}, \tag{1.14}$$

and interpreted as the population size N required to obtain a recombinant with a specified confidence $R(1/2)$ when the initial frequency is $x = 1/2$. Thus, to achieve $R(1/2) = 0.99$ we need a population size of the order $N = [1/(2r)]\log 200 = 5.3/(2r)^{1/2}$ and for $R(1/2) = 0.75$, we require $N = 2.2/(2r)^{1/2}$.

The preceding model can also be analyzed taking account of selection effects. The details are set forth in section 3. We shall now set up a selection model and indicate some of the main results.

Assume the frequencies of the three genotypes Ab/Ab, aB/Ab, aB/aB, are x^2, $2x(1-x)$, $(1-x)^2$, where $x = i/2N$, and that the respective selective coeffi-

cients are $1 + a$, 1, $1 + b$. Then the gametic output contains the four types of gametes in relative proportions as follows:

$Ab: 2x^2(1 + a) + 2x(1 - x)(1 - r)$;
$aB: 2(1 - x)^2(1 + b) + 2x(1 - x)(1 - r)$;
$ab: 2x(1 - x)r$;
$AB: 2x(1 - x)r$.

Hence, (1.2) and (1.3) must be replaced by

$$p_i = \frac{x^2(1 + a) + x(1 - x)(1 - r)}{1 + ax^2 + b(1 - x)^2},$$
$$q_i = \frac{(1 - x)^2(1 + b) + x(1 - x)(1 - r)}{1 + ax^2 + b(1 - x)^2}, \tag{1.15}$$

and we have

$$p_i + q_i = 1 - \frac{2rx(1 - x)}{1 + ax^2 + b(1 - x)^2}. \tag{1.16}$$

The probability $R = R(x, a, b, r,$ mating system) of producing the recombinant type AB generally depends on the parameters in a complicated manner. However, for the special case $a = -b = c > 0$, in which the Ab chromosome confers an additive selective advantage relative to aB, the function $R(x)$ satisfies a differential equation which is amenable to a complete solution. We obtain the formula

$$R(x) = 1 - \frac{e^{2\gamma(1-x)} \sinh x\,(8\sigma + 4\gamma^2)^{1/2} + e^{-2\gamma x} \sinh(1 - x)(8\sigma + 4\gamma^2)^{1/2}}{\sinh\,(8\sigma + 4\gamma^2)^{1/2}}, \tag{1.17}$$

where $\sigma = rN^2$ and $\gamma = Nc$. In this case fixation tends to occur at $x = 1$. Equivalently, inspection of the function $R(x)$ reveals that its graph is skewed toward $x = 0$. An approximation to (1.17) valid for $Nr^{1/2}$ and Nc not too small for the special initial gamete frequency $x = 1/2$ is

$$R(\tfrac{1}{2}) \sim 1 - 2 \exp\{-N[(2r + c^2)^{1/2} - c]\}. \tag{1.18}$$

It follows that to achieve $R(1/2) = 0.99$ we need a minimum population size

$$N = \frac{5.3}{(2r + c^2)^{1/2} - c} > \frac{5.3}{(2r)^{1/2}}. \tag{1.19}$$

Another important special situation is the case of heterotic selection, that is, where the selection coefficients are related as $a = b = -c$, $-\infty < c < 1$. The heterozygote is at an advantage if $c > 0$ and at a disadvantage if $c < 0$ compared to either homozygote. In this case we cannot obtain an explicit formula for $R(x)$. Recourse to the asymptotic theory of differential equations leads to the asymptotic expression (for N large but r and c fixed)

$$1 - R(x) \sim e^{-2Ncx(1-x)}\,[\phi'(x)]^{-1/2}\, w_0[\phi(x)], \tag{1.20}$$

where

$$\phi(x) = \frac{1}{(2\rho)^{1/2}} \int_0 \left[2r + c^2(1-2\xi)^2 - \frac{c}{N}\right]^{1/2} d\xi,$$

$$(2\rho)^{1/2} = \int_0^1 \left[2r + c^2(1-2\xi)^2 - \frac{c}{N}\right]^{1/2} d\xi, \tag{1.21}$$

and $w_0(t)$ is the solution of the equation

$$w_0''(t) - 8N^2\rho w_0(t) = 0, \qquad w_0(0) = w_0(1) = [\phi'(0)]^{1/2}. \tag{1.22}$$

It can be verified that $R(x)$ increases if either c or r increases. Furthermore, $R(x)$ is symmetric about $x = 1/2$.

For the special initial frequency $x = 1/2$, we obtain

$$1 - R\left(\frac{1}{2}\right) = \left(\frac{2r + c^2 - \frac{c}{N}}{2r - \frac{c}{N}}\right)^{1/4} 2 \exp\left\{-N\left[(2\rho)^{1/2} + \frac{c}{2}\right]\right\}. \tag{1.23}$$

We call the selection *relatively weak* if $c^2 \leqq 2r$. In this case, if $c > 0$, further examination of (1.23) yields the good approximation

$$N \sim \frac{1}{(2\rho_\infty)^{1/2} + \frac{1}{2}c} \log \frac{2}{1 - R\left(\frac{1}{2}\right)}, \tag{1.24}$$

where

$$\rho_\infty = (2r)^{1/2} \int_0^1 \left[1 + \frac{c^2}{2r}(1-2\xi)^2\right]^{1/2} d\xi, \tag{1.25}$$

and if $c^2/2r \ll 1$ is close to zero (1.24) can be replaced by

$$N \sim \frac{1}{(2r)^{1/2} + \frac{c}{2}} \log \frac{2}{1 - R\left(\frac{1}{2}\right)}. \tag{1.26}$$

Further discussion and indications of the derivation of the above facts are given in the body of the paper. We also review briefly procedures for calculating the moments of the distribution function of the time until a recombinant is formed. This distribution is representable by an integral formula involving exponential functions. We omit the formal elaboration of these results.

The methods of this paper reveal a novel form of diffusion approximations, natural and appropriate for the study of certain genetic problems. These same techniques can be applied to analyze the effects of artificial selection schemes versus random elimination operating on finite population genetic models. We could also handle the model of this paper with random mating replaced by certain inbreeding systems, especially selfing or a mixture of selfing and random mating. These considerations and other applications are deferred to a future publication.

2. Random mating without selection

Consider in each generation, a population of N diploids consisting of only the genotypes Ab/Ab, Ab/aB, aB/aB.

The finite state Markov chain whose state variable denotes the number of Ab gametes governed by the transition probability matrix (1.4) is the standard model used to study the fluctuations of the gamete frequencies accountable to random sampling and recombination effects. The process ends if either fixation occurs or a new recombinant type appears. We proceed to the analysis of the given Markov chain.

The gamete frequency in the nth generation, provided no recombinant has yet appeared, is a random variable $X(n)$ with one of the values $i/2N$, $i = 0, 1, \cdots, 2N$. If $X(n) = i/2N = x$, then it may happen that in the next generation no recombinant appears. In this case for the difference $\Delta X = X(n+1) - X(n)$ we have the expected values

$$
\begin{aligned}
E(\Delta X) &= [p_i - x(p_i + q_i)](p_i + q_i)^{2N-1}, \\
E[(\Delta X)^2] &= \left\{[p_i - x(p_i + q_i)]^2 + \frac{p_i q_i}{2N}\right\}(p_i + q_i)^{2N-2}.
\end{aligned} \tag{2.1}
$$

Alternatively, a recombinant may appear in the $(n+1)$st generation, and the probability of this is given by (1.6). In this case we say the gamete frequency process ends by recombination in the $(n+1)$st generation, and $X(k)$ is not defined for $k \geqq n+1$.

We introduce a new recombination parameter σ defined by

$$
r = \sigma/N^2. \tag{2.2}
$$

If σ is regarded as constant the above expected values are of the form

$$
E(\Delta X) = O(1/N^2), \qquad E[(\Delta X)^2] = \frac{x(1-x)}{2N} + O(1/N^2), \tag{2.3}
$$

and the probability that the process ends by recombination in the $(n+1)$st generation is

$$
\frac{4\sigma x(1-x)}{N} + O(1/N^2). \tag{2.4}
$$

We choose the time scale so that unit time is taken for the passage of N generations. Then the time elapsing between successive generations is $\Delta t = 1/N$ which $\to 0$ when $N \to \infty$. Hence,

$$
\begin{aligned}
M(x) &= \lim_{\substack{N\to\infty \\ \Delta t\to 0}} E\left(\frac{\Delta X}{\Delta t}\right) = 0, \\
V(x) &= \lim_{\substack{N\to\infty \\ \Delta t\to 0}} E\left[\frac{(\Delta X)^2}{\Delta t}\right] = \frac{x(1-x)}{2},
\end{aligned} \tag{2.5}
$$

and the limiting rate of appearance of recombinants is

$$
K(x) = 4\sigma x(1-x). \tag{2.6}
$$

The diffusion process which has the same limiting mean change $M(x)$, variance $V(x)$, and recombination rate $K(x)$ is that governed by the (backward) diffusion equation

$$\frac{\partial u}{\partial t} = \frac{1}{2} V(x) \frac{\partial^2 u}{\partial x^2} + M(x) \frac{\partial u}{\partial x} - K(x)u \tag{2.7}$$

$$= \frac{x(1-x)}{4} \frac{\partial^2 u}{\partial x^2} - 4\sigma x(1-x)u.$$

Diffusion equations like (2.7), which contain the term $-K(x)u$, called a killing term, are known in diffusion theory but apparently have not previously been used in genetics. For the diffusion process governed by (2.7), the states $x = 0$ and $x = 1$ are absorbing barriers corresponding to fixation of the population. If at some time t the process is in state x, $0 < x < 1$, then in the subsequent small time interval $(t, t + dt)$ the process ends or is killed, corresponding to the appearance of a recombinant, with probability $K(x)\,dt$. With probability one the process either ends in this way by recombination at some finite time, or else one of the two states of fixation is reached in finite time.

Let $Z(x)$ be the probability that the process eventually ends by recombination when it starts in state x, $0 < x < 1$. By standard methods of diffusion theory we deduce from (2.7) that $Z(x)$ satisfies

$$\frac{1}{2} V(x) \frac{d^2 Z}{dx^2} + M(x) \frac{dZ}{dx} - K(x)Z = -K(x), \qquad Z(0) = 0, Z(1) = 0. \tag{2.8}$$

In the special case at hand this simplifies to

$$\frac{1}{4} Z'' - 4\sigma Z = -4\sigma, \qquad Z(0) = Z(1) = 0, \tag{2.9}$$

and therefore,

$$Z(x) = 1 - \frac{\sinh 4x\sigma^{1/2} + \sinh 4(1-x)\sigma^{1/2}}{\sinh 4\sigma^{1/2}}. \tag{2.10}$$

This is the expression (1.10) of the introduction.

A related probability is of even greater interest. When a recombinant appears it may be either of type ab or AB. We assume that individuals with the ab gamete can be recognized and discarded from the population, and that the population is maintained at size N until either fixation occurs or an AB recombinant appears. By means of a simple renewal argument we find that the probability $R(x)$ of eventually obtaining an AB recombinant satisfies

$$\frac{1}{2} V(x) \frac{d^2 R}{dx^2} + M(x) \frac{dR}{dx} - K(x)R(x) = -K(x)\left[\frac{1}{2} + \frac{1}{2} R(x)\right], \tag{2.11}$$

$$R(0) = 0, \quad R(1) = 0.$$

For the special case at hand this leads to

$$R(x) = 1 - \frac{\sinh x(8\sigma)^{1/2} + \sinh (1-x)(8\sigma)^{1/2}}{\sinh (8\sigma)^{1/2}}. \tag{2.12}$$

Some interpretations and uses of formula (2.12) are discussed in the introductory section of this paper.

If $R(x, t)$ is the probability that an AB recombinant appears before time t when the process starts in state x at time $t = 0$, we find that the Laplace transform

$$\Phi(x, \lambda) = \int_0^\infty e^{-\lambda t} \frac{\partial R(x, t)}{\partial t}\, dt, \qquad \lambda > 0, \tag{2.13}$$

satisfies

$$\frac{1}{2} V(x) \frac{d^2\Phi}{dx^2} + M(x) \frac{d\Phi}{dx} - K(x)\,\Phi - \lambda\Phi = -K(x)\left[\frac{1}{2} + \frac{1}{2}\Phi\right], \tag{2.14}$$

$$\Phi(0, \lambda) = 0, \quad \Phi(1, \lambda) = 0.$$

Consider the moments

$$\Phi_n(x) = (-1)^n \frac{\partial^n \Phi}{\partial \lambda^n}\bigg|_{\lambda=0} = \int_0^\infty t^n \frac{\partial R(x, t)}{\partial t}\, dt; \tag{2.15}$$

it follows from the above that

$$\frac{1}{2} V(x) \frac{d^2\Phi_n}{dx^2} + M(x) \frac{d\Phi_n}{dx} - \frac{1}{2} K(x)\,\Phi_n = n\Phi_{n-1}, \tag{2.16}$$

$$\Phi_n(0) = 0, \quad \Phi_n(1) = 0.$$

Since $\Phi_0(x) = R(x)$ is known, the successive moments can be found by elementary methods.

3. Effects of selection

In this section we analyze the Markov chain model with transition probability matrix (1.4) whose parameters p_i and q_i (see (1.15)) involve the selection coefficients.

To find a diffusion approximation for the Markov chain at hand, we let

$$a = \frac{\alpha}{N}, \qquad b = \frac{\beta}{N}, \qquad r = \frac{\sigma}{N^2} \tag{3.1}$$

and assume that α, β, σ remain fixed and $N \to \infty$. Proceeding as in section 2, we obtain

$$V(x) = \frac{x(1 - x)}{2}, \tag{3.2}$$

$$M(x) = x(1 - x)[\alpha x - \beta(1 - x)], \qquad K(x) = 4\sigma x(1 - x).$$

Thus, the diffusion equation is

$$\frac{\partial u}{\partial t} = \frac{x(1 - x)}{4} \frac{\partial^2 u}{\partial x^2} + x(1 - x)[\alpha x - \beta(1 - x)] \frac{\partial u}{\partial x} - 4\sigma x(1 - x)u. \tag{3.3}$$

The probability $R(x)$ of obtaining an AB recombinant before fixation is the solution of

$$\frac{1}{4}R'' + [\alpha x - \beta(1-x)]R' - 2\sigma R = -2\sigma, \qquad R(0) = R(1) = 0. \tag{3.4}$$

Equation (3.4) is not an elementary differential equation (except if $\alpha + \beta = 0$, when it has constant coefficients). Nevertheless, since its coefficients are linear functions of x, it is manageable by classical methods. Its solution can be represented by Laplace integrals, and the asymptotic behavior of $R(x)$ when α, β and σ are large, can be determined.

Since $R(x)$ depends on the four parameters α, β, σ, x in a complicated way, it seems advisable to consider special cases, and we describe the results for two examples. In example 1, which is particularly simple, the complete solution is obtained. In the second example we find an asymptotic approximation of the solution, from which one can estimate the population size N required to achieve a specified certainty of getting a recombinant before fixation. The technique of the second example can be applied to other cases.

Example 1. Let $\alpha = -\beta = \gamma$. In this case the Ab chromosome has an additive advantage over aB. The differential equation simplifies to

$$\frac{1}{4}R'' + \gamma R' - 2\sigma R = -2\sigma, \tag{3.5}$$

and we find

$$R(x) = 1 - \frac{e^{2\gamma(1-x)} \sinh x(8\sigma + 4\gamma^2)^{1/2} + e^{-2\gamma x} \sinh (1-x)(8\sigma + 4\gamma^2)^{1/2}}{\sinh (8\sigma + 4\gamma^2)^{1/2}} \tag{3.6}$$

(see the discussion of (1.17)).

Example 2. Let $\alpha = \beta = -\gamma$. This is the case of heterotic selection. The heterozygote is at an advantage if $\gamma > 0$, a disadvantage if $\gamma < 0$, compared to either homozygote. When σ is fixed it is clear that $R(x)$ should be an increasing function of γ.

If we let $F(x) = 1 - R(x)$ be the probability of ultimate fixation then $F(x)$ is the solution of

$$\frac{1}{4}F'' + \gamma(1-2x)F' - 2\sigma F = 0, \qquad F(0) = F(1) = 1. \tag{3.7}$$

Our aim is to find the approximate value of $F(x)$ when $\sigma = N^2 r$, $\gamma = Nc$ with $-\infty < c < 1$, where c and r are fixed and N is large. We make the preliminary substitution

$$F(x) = e^{-2Ncx(1-x)}\, y(x), \tag{3.8}$$

which will have the effect of eliminating the first derivative term from the differential equation. Then $y(x)$ is the solution of

$$y'' - [8N^2 r + 4N^2 c^2 (1-2x)^2 - 4Nc]y = 0, \qquad y(0) = y(1) = 1. \tag{3.9}$$

The final substitution is now chosen so that the coefficient in the resulting differential equation is nearly constant. We assume, if $c > 0$, that $2r - c/N$ is positive, that is, $N > c/2r$. Let

$$(2\rho)^{1/2} = \int_0^1 \left[2r + c^2(1-2x)^2 - \frac{c}{N}\right]^{1/2} dx,$$

$$t = \phi(x) = \frac{1}{(2\rho)^{1/2}} \int_0^x \left[2r + c^2(1-2\xi)^2 - \frac{c}{N}\right]^{1/2} d\xi, \tag{3.10}$$

$$y(x) = [\phi'(x)]^{-1/2}\, w(t).$$

Then $w(t)$ is the solution of

$$\frac{d^2w}{dt^2} - [8N^2\rho + H(t)]w = 0, \tag{3.11}$$

$$w(0) = w(1) = [\phi'(0)]^{1/2} = \left(\frac{2r + c^2 - \frac{c}{N}}{2\rho}\right)^{1/4},$$

where

$$H(t) = \left[\frac{1}{2}\frac{\phi'''(x)}{\phi'(x)} - \frac{3}{4}\left(\frac{\phi''(x)}{\phi'(x)}\right)^2\right] \Big/ [\phi'(x)]^2. \tag{3.12}$$

When $N \to \infty$, the quantity $(2\rho)^{1/2}$ approaches a positive limit, and $\phi(x)$ and its derivatives converge uniformly, and $\phi'(x)$ remains bounded away from zero. Hence, $H(t)$ is uniformly bounded on $0 \leqq t \leqq 1$. We can show that the solution $w(t)$ of (3.11) is nearly the same, for large N, as the solution of

$$\frac{d^2w_0}{dt^2} - 8N^2\rho w_0 = 0, \qquad w_0(0) = w_0(1) = [\phi'(0)]^{1/2}. \tag{3.13}$$

More precisely,

$$w(t) = w_0(t)[1 + O(1/N)]. \tag{3.14}$$

Using this we have as the approximate solution of (3.7)

$$F(x) = 1 - R(x) \tag{3.15}$$

$$\sim e^{-2Ncx(1-x)}[\phi'(x)]^{-1/2}\, w_0[\phi(x)],$$

and, since $\phi(1/2) = 1/2$,

$$1 - R(1/2) \sim \left(\frac{2r + c^2 - \frac{c}{N}}{2r - \frac{c}{N}}\right)^{1/4} 2e^{-N[(2\rho)^{1/2} + c/2]}. \tag{3.16}$$

4. Discussion

A striking example of repeated accelerated responses to selection for increased bristle number in *Drosophila melanogaster* was reported by Thoday and Boam [6]. One explanation they suggested for the responses was based on the production of a particularly favorable AB recombinant gamete from a repulsion heterozygote Ab/aB. Thus, suppose that the gamete Ab is initially rare, but has a positive effect on bristle number relative to aB. Then heterozygotes will initially

be rare, and so the probability of producing AB recombinants is low, but this probability will increase as the frequency of the gamete Ab increases by selection. The mean time to the onset of the accelerated response will be the mean time to the production of an AB recombinant. The models considered in this paper are applicable to such an explanation for an accelerated response to selection. The expression for $R(x)$ given in equation (1.17) shows the dependence of the probability of producing a recombinant before fixation, on the population size, recombination fraction, and selective values. It is clear that population sizes must be quite small, the recombination fraction must be very small, and the initial frequency of one gamete very small for the risk of fixation to be significant. However, laboratory populations are often small enough for these considerations to be important and linkage values may be small enough in natural populations to retard significantly the rate of production of appropriate recombinant gametes. Further work is clearly needed to correlate the theoretical results more precisely with experimental findings.

5. Summary

Some results are given for the probability that in a finite population containing initially only gametes Ab and aB (alleles A, a and B, b at two linked loci), the recombinant gametes AB or ab are produced before the population fixes on either of the initial gametes Ab or aB.

REFERENCES

[1] W. F. BODMER and P. A. PARSONS, "Linkage and recombination in evolution," *Adv. Genet.*, Vol. 11 (1962), pp. 1–100.
[2] W. J. EWENS, "Numerical results and diffusion approximations in a genetic process," *Biometrika*, Vol. 50 (1963), pp. 241–249.
[3] R. A. FISHER, *The Genetical Theory of Natural Selection*, New York, Dover, 1958.
[4] S. KARLIN and J. MCGREGOR, "On some stochastic models in genetics," *Stochastic Models in Medicine and Biology*, Madison, University of Wisconsin Press, 1964, pp. 245–271.
[5] M. KIMURA, "Diffusion models in population genetics," *J. Appl. Probability*, Vol. 1 (1964), pp. 177–232.
[6] J. M. THODAY and T. B. BOAM, "Regular responses to section I. Description of responses," *Genet. Res.*, Vol. 2 (1961), pp. 161–176.
[7] S. WRIGHT, "Evolution in Mendelian populations," *Genetics*, Vol. 16 (1931), pp. 97–159.