RATES AND PROBABILITIES OF FIXATION FOR TWO LOCUS RANDOM MATING FINITE POPULATIONS WITHOUT SELECTION

SAMUEL KARLIN AND JAMES McGREGOR[1,2]

Department of Mathematics, Stanford University, Stanford, California 94305

Received June 22, 1967

A large body of literature has accumulated over the past decade concerned with the interaction between selection and linkage in two locus random mating populations. BODMER and PARSONS (1962) discuss the significance of this problem, in particular the effects of linkage on the fate of newly arisen gene complexes.

Conditions for the existence of stable polymorphic equilibria in the 2 locus infinite population size model were investigated by WRIGHT (1952), KIMURA (1956), LEWONTIN and KOJIMA (1960), BODMER and FELSENSTEIN (1967) and others under a variety of restrictions of the selection coefficients. The nonlinear equations relating genotypic frequencies of successive generations obtained in these problems are unlikely to yield in general to theoretical solution unless special assumptions are made on the selection coefficients e.g., symmetric viabilities, no epistasis (the additive model), multiplicative effects, etc.

Recourse to numerical computations and simulation techniques has been made where exact analysis has appeared impossible. However interpretations of the results of these numerical calculations must be made cautiously since the initial conditions and the length of trial run affect the numerical results decisively (see e.g., EWENS 1966, 1967). Numerical calculations for the two locus case occur also in KOJIMA (1965), LEWONTIN (1964a,b,c), PARSONS (1963a,b), KIMURA (1965), JAIN and ALLARD (1965) and SINGH and LEWONTIN (1966) and elsewhere. All of these works concentrate exclusively on infinite population deterministic models.

Monte Carlo studies of aspects of the two locus model for small populations have recently been reported by HILL and ROBERTSON (1966). These authors were interested primarily in ascertaining the influence of the linkage parameter on response to artificial selection where the loci were assumed to have additive effects on the character under selection and not to interact. As is not uncommon with simulation procedures, the interpretations are moot and somewhat vague. LATTER (1966) conducted similar Monte Carlo programs endowed with corresponding difficulties.

Conflicts in the results of computation between different Monte Carlo studies confuse the issue further. For example, compare EWENS (1966), JAIN and ALLARD (1965) and BODMER and PARSONS (1962). It is indicated (confirmed in HILL and ROBERTSON [1966]) that the transient behavior of the process is highly sensitive

[1] We express our great indebtedness to Mr. M. FELDMAN for many constructive criticisms and discussions.

[2] Research supported in part under contract NIH 10452 Stanford University.

to the initial conditions, the range of values of the parameters (recombination fraction, selection coefficients and population size) and the formulation of the model. The exact analysis of some two locus stochastic models would aid in establishing the detailed nature of the dependence of rates and probabilities of fixation on the structure and parameters of the model.

In this paper we examine a finite state stochastic frequency model (of constant population size) for a pair of linked loci with alleles A, a and B, b at the first and second locus respectively. The four possible gametes are, of course, AB, Ab, aB and ab. The model proposed to describe the fluctuations of the gamete types is given the canonical formulation namely that analogous to the one locus random sampling model of FISHER and WRIGHT. Specifically, suppose the current generation is composed of N diploid individuals with gametic numbers as listed in the table.

(1)

Gamete	Numbers
AB	i_1
Ab	i_2
aB	i_3
ab	i_4
Total population of gametes	$2N$

The i_ν ($\nu = 1, 2, 3, 4$) are non-negative integers obeying the constraint $i_1 + i_2 + i_3 + i_4 = 2N$. We determine the next generation by performing random sampling (i.e., multinomial trials) $2N$ times on the gametic output of the population produced by *random union of gametes.* (This will be referred to henceforth as random mating. This model of random union of gametes is different from that in which random union of zygotes is postulated. More details on this difference are given in the discussion.) Random mating yields the genotypes $\frac{AB}{AB}, \frac{AB}{Ab}, \frac{AB}{ab}, \ldots \frac{ab}{ab}$, in the proportions

$$\frac{i_1^2}{4N^2}, \frac{2i_1i_2}{4N^2}, \frac{2i_1i_3}{4N^2}, \ldots, \frac{i_4^2}{4N^2},$$

respectively. If r denotes the recombination fraction, a genotype Ab/aB has a gametic output described by the array

$$\frac{1}{2}r\,AB + \frac{1}{2}r\,ab + \frac{1}{2}(1-r)\,Ab + \frac{1}{2}(1-r)\,aB$$

and analogous segregation proportions apply in the other cases. Combining all the possibilities we find that a gamete chosen at random is AB with probability

(2)
$$p_1 = \left(\frac{i_1}{2N}\right)^2 + \frac{i_1i_2}{4N^2} + \frac{i_1i_3}{4N^2} + \frac{i_1i_4}{4N^2}(1-r) + \frac{i_2i_3}{4N^2}r$$
$$= \frac{i_1}{2N} + \frac{r}{4N^2}D$$

where

$$D = i_2 i_3 - i_1 i_4 \tag{3}$$

is the linkage deviation function corresponding to the present population makeup (D is more commonly called the linkage disequilibrium function).

More generally, each performance of random sampling produces one of the gametes with corresponding probabilities as follows

Gamete	Probability
AB	$p_1 = \frac{i_1}{2N} + \frac{r}{4N^2} D$
Ab	$p_2 = \frac{i_2}{2N} - \frac{r}{4N^2} D$
aB	$p_3 = \frac{i_3}{2N} - \frac{r}{4N^2} D$
ab	$p_4 = \frac{i_4}{2N} + \frac{r}{4N^2} D$

(4)

We are now ready to formulate the precise stochastic model underlying the fluctuations of the gamete frequencies over time. The state space of the process (symbolized by Δ) is described by vectors of 4 nonnegative integers.

$$\Delta = \{i = (i_1, i_2, i_3, i_4);\ i_r \text{ integers} \geq 0,\ i_1 + i_2 + i_3 + i_4 = 2N\},$$

where i_1 represents the number of AB gametes in the current generation and i_2, i_3, i_4 that of Aa, aB and ab respectively. The next generation is formed by $2N$ multinomial trials as follows. If the parent population is described by the vector $i = (i_1, i_2, i_3, i_4)$ then each trial results in AB, Ab, aB and ab with probabilities p_1, p_2, p_3 and p_4 respectively (p_i defined in (4)). Repeated samplings are made with replacement. By this procedure we generate a Markov chain

$$i^{(n)} = (i_1^{(n)}, i_2^{(n)}, i_3^{(n)}, i_4^{(n)}) \ \varepsilon\ \Delta,\ n = 0, 1, 2, \ldots$$

where $i^{(n)}$ is the vector which describes the population makeup in the nth generation. Let $i = (i_1, i_2, i_3, i_4) \ \varepsilon\ \Delta$ denote the current state variable and $j = (j_1, j_2, j_3, j_4)$ $\varepsilon\ \Delta$ the state variable of the next generation. The transition probability law which governs the fluctuations of the population structure from generation to generation is computed in accordance to the multinomial distribution to be

$$P_{i,j} = \frac{2N!}{j_1!\, j_2!\, j_3!\, j_4!} p_1^{j_1} p_2^{j_2} p_3^{j_3} p_4^{j_4} \tag{5}$$

There are $\binom{2N+3}{2N}$ states for the process corresponding to all possible 4-tuples of gamete combinations.

The model is now fully structured. The ultimate goal would be to describe as completely as possible the probabilistic distribution of the population vector $i^{(n)}$ or various probability laws of functionals of these vectors. The Markov chain

although simple in formulation is extremely complex in behavior and cannot be resolved in classical mathematical terms. We shall be content to extract special information of some genetic interest. Previously certain special functionals of the stochastic process corresponding to the mechanism of random union of zygotes have been examined by WRIGHT and KIMURA. WRIGHT (1933) used the method of path coefficients while KIMURA (1963) has developed some elegant refinements on the use of the method of identity by descent due to MALÉCOT (1948) (see also MALÉCOT 1951, 1959a, 1959b). The stochastic model underlying the work of WRIGHT and KIMURA differs from that outlined above. The models are compared in the discussion.

Every finite interbreeding genetic population not subject to gene mutation will eventually become homozygous. From classical properties of Markov chains and by the nature of the specific process at hand, the population ultimately fixes (due to random drift) in one of the pure states $I_1 = (2N, 0, 0, 0)$, $I_2 = (0, 2N, 0, 0)$, $I_3 = (0, 0, 2N, 0)$ or $I_4 = (0, 0, 0, 2N)$ consisting exclusively of AB, Ab, aB or ab gametes respectively.

Notice we have postulated no selection differences in the formation of the next generation. This is, of course, a limitation on the model but even so the analysis of this simple model is complicated. This discussion may be helpful toward the end of carrying out a theoretical treatment of a corresponding stochastic model involving selection differences among the genotypes.

Our objectives in this paper are fourfold.

(I) We will determine the probabilities of fixation in states I_1, I_2, I_3 and I_4 as a function of the initial population makeup $i = (i_1, i_2, i_3, i_4)$, the recombination fraction r and the population size N.

(II) We will ascertain the precise rate of approach to fixation. Indeed, the theory of finite Markov chains tells us that the probability f_n that the population at the nth generation includes at least two types of gametes is asymptotically of the order

$$f_n \sim c(i)\ \lambda^n \qquad n \to \infty$$

for some λ, $0 < \lambda < 1$ where $c(i)$ is a number depending on the initial population makeup but independent of the generation time n. The number λ is called the *rate of approach to fixation.* The following interpretation can be ascribed to λ. Consider a large number of independent populations each of size $2N$ governed independently by the same transition probability law (5) and assume many generations have already passed by. Then in each succeeding generation a proportion $1-\lambda$ of these populations become fixed. The quantity $1/(1-\lambda)$ is also related to the concept of effective population size and can be interpreted in these terms (see KIMURA and CROW 1963).

(III) Fixation takes place in two stages. First one of the two loci becomes homozygous. Thereafter the fluctuations of the process are completely described by the WRIGHT-FISHER binomial sampling model involving two alleles at one locus. We will determine the probabilites (as a function of the initial

population makeup $i = (i_1, i_2, i_3, i_4)$ and the parameter r) that the population fixes first in one of the four alleles A, a, B or b.

(IV) Finally and of *most importance* we will determine the rate of loss of some allele from the population. More specifically, let π_n denote the probability that in the nth generation all four gametes are present. From standard theory concerning Markov chains, we know that π_n has the asymptotic form $\pi_n \sim K(i)\ \mu^n$ as $n \to \infty$ where $0 < \mu < 1$ and $K(i)$ depends on the initial population numbers $i = (i_1, i_2, i_3, i_4)$ but is independent of n. Thus the chance per generation that the population loses its original polymorphic state is of the order $1-\mu$.

The importance of the answers to the problems posed in III and IV pertaining to the effects of small population size on evolutionary theory is discussed in WRIGHT (1931), MORAN (1962), and elsewhere. For example, in the case of a small population (say laboratory stock) subject to an inbreeding mechanism, it is important to know how long the original 4 gametes will remain together in the population for different sets of initial numbers of gametic types. Inbreeding models in this connection are usually easier to analyse than those involving random mating. The problems in the former are in some sense linear, while those in the latter are at best quadratic.

The explicit solutions of problems (I) through (IV) are summarized and interpreted in Section 2. The rate of decrease of the variance of the linkage disequilibrium function is also determined. The rigorous proofs of the results of Section 2 involve careful use of the transformation properties of the probability transition matrix (5) when applied to certain polynomials in the variables i_1, i_2, i_3 and i_4. The detailed argument will be presented elsewhere. Extensions of some of the results of Section 2 to take account of a general progeny distribution per mating are indicated in Section 3. Section 4 concludes with a general discussion of the significance and limitations of the results of this paper and some discussion of previous work.

A few comments on previous related studies may facilitate interpretation of the ensuing mathematical results. With no selection differences present among genotypes and infinite populations undergoing random mating it is a classical result that the gamete frequencies converge geometrically fast with rate $1-r$ (note the dependence on r) to a set of frequencies which are at zero linkage deviation (see LI 1955 or KEMPTHORNE 1957 for appropriate historical references concerning this result).

The rates of approach to homozygosity were determined for certain monoecious and dioecious one locus random mating finite population stochastic models by WATTERSON (1959a), (1959b), MORAN and WATTERSON (1959), and MORAN (1962). An approximate value for the rate of approach to homozygosity was obtained by KIMURA (1955) in his analysis of the standard diffusion approximation to the finite population process. A general method for determining the rate of approach to statistical equilibrium, or homozygosity and the probabilities of fixation of one or the other gene under a variety of mating systems allowing for

the pressures of mutation, migration and finite population size has been elaborated in KARLIN and MCGREGOR (1964a, 1965a), see also KARLIN (1966a, Chap. 13). We mentioned earlier the simulation studies of the interaction between effective population size and linkage intensity under artificial selection given by LATTER (1966) and HILL and ROBERTSON (1966), see also EWENS (1966). A novel form of diffusion approximations involving a killing term, was used by KARLIN, MCGREGOR and BODMER (1965) to determine the probability of recombination before fixation as a function of the initial gametic frequencies, the recombination fraction, the population size, selective values and the mating system.

The present paper seems to be the first to resolve precisely the questions (I) to (IV), with special emphasis on (III) and (IV), for the classical FISHER-WRIGHT random sampling of a two locus two allele finite population size stochastic model. The qualitative consequences are intriguing and in the light of previous work perhaps surprising.

RESULTS AND GENERAL DISCUSSION

In this section formulae and conclusions are set forth which fulfill the objectives of Section 1. The formulae obtained are exact; the methods by which they were obtained involve, as mentioned before, some detailed analysis of the Markov chains involved. We feel this is not the appropriate forum to record these arguments. For the details the reader is referred to BODMER, FELDMAN and KARLIN (1968). There is some conflict between the interpretations of these results with some of the inferences derived from the simulation studies reported by HILL and ROBERTSON (1966).

I. *Probabilities of Fixation in a Pure Gamete State.* Every finite interbreeding genetic population not subject to gene mutation will eventually become homozygous. Fixation can occur in one of four pure states consisting entirely of one of the four gametes AB, Ab, aB or ab. The probabilities of these events are listed in Table 1 and obviously depend on the initial population numbers, the recombination fraction r, and the total population size.

We are grateful to DR. M. KIMURA who pointed out to us that the results of

TABLE 1

Initial population vector $i = (i_1, i_2, i_3, i_4)$ $D = i_2 i_3 - i_1 i_4$

Pure state	Probability of fixation as a function of i, r and N
AB gamete $(2N, 0, 0, 0)$	$\frac{i_1}{2N} + \frac{r}{2N[1-r+r2N]} D$
Ab gamete $(0, 2N, 0, 0)$	$\frac{i_2}{2N} - \frac{r}{2N[1-r+r2N]} D$
aB gamete $(0, 0, 2N, 0)$	$\frac{i_3}{2N} - \frac{r}{2N[1-r+r2N]} D$
ab gamete $(0, 0, 0, 2N)$	$\frac{i_4}{2N} + \frac{r}{2N[1-r+r2N]} D$

Table 1 can be obtained by adapting the method of identity by descent as in KIMURA (1963). This method, however, cannot be applied to obtain the results in the later sections (especially parts III and IV) of this paper.

The following are easily inferred from the table.

(a) The probability of fixation of the population in a particular gamete (say Ab) depends on the initial population numbers of the four gametes only through the initial frequency of that particular gamete (Ab) and the linkage deviation function of the initial population vector.

(b) If we start the process with i_1 AB, i_2 Ab, i_3 aB and i_4 ab, the probability of fixation in a particular gamete depends on the recombination fraction if and only if $D = i_2 i_3 - i_1 i_4 \neq 0$, i.e., if and only if the initial frequencies are not in a state of zero linkage deviation.

The second observation is somewhat striking since stochastic fluctuations will most likely disturb the value of D in one generation to a nonzero value. Nevertheless in a certain average sense linkage equilibrium persists over succeeding generations. To explain this notion we let $\bar{\imath}' = (i_1', i_2', i_3', i_4')$ describe the population vector after one generation and let $\bar{\imath} = (i_1, i_2, i_3, i_4)$ denote the initial population vector. It is interesting to compare the linkage deviation function $D' = i_2' i_3' - i_1' i_4'$ for the second generation with that of $D = i_2 i_3 - i_1 i_4$ of the initial generation. Of course, D' is a random variable following a probability distribution that can in principle (although not in a practical fashion) be calculated from knowledge of the probabilistic laws (5) governing the process. The expected value of D' can be routinely determined and we obtain the simple formula

$$\mathcal{E}(D') = (1-1/2N)\ (1-r)\ D. \tag{6}$$

(We use the symbol $\mathcal{E}(X)$ to denote expectation of X.) The above was also noted by HILL and ROBERTSON. After n generations, we have

$$\mathcal{E}(D^n) = [(1-1/2N)\ (1-r)]^n\ D^{(0)} \tag{7}$$

which shows that the average value of $D^{(n)}$ approaches zero at a geometric rate decreasing by a factor of $(1-1/2N)\ (1-r)$ per generation. Notice that if $D^{(0)} = 0$ then $\mathcal{E}(D^n) = 0$ for all n. Later (see (19)) we shall describe the transient behavior of the variance of $D^{(n)}$ as $n \to \infty$. The rate of decrease differs markedly from that of $\mathcal{E}(D^n)$. In fact, $\mathrm{Var}(D^{(n)})$ tends to zero at a slower geometric rate suggesting that although the average value of $D^{(n)}$ may be close to zero the actual value of the linkage deviation function could be relatively appreciably different from zero.

(c) HILL and ROBERTSON (1966) and also LATTER (1966) in their computer runs always start with $D = 0$ for the initial population. Inspection of the formulae of Table 1 reveals that the probabilities of fixation depend on the initial value of D in an essential fashion (cf. also the first statement of paragraph (b) above). Several relevant genetic considerations which dictate against the assumption of initial value of $D = 0$ are indicated in the discussion at the conclusion of the paper.

(d) HILL and ROBERTSON (1966) claim that the probabilities of fixation depend on the recombination fraction r only as a function of Nr. The formulas of Table 1

contradict this assertion. HILL and ROBERTSON carried out a simulation program based on the stochastic model as formulated here, actually slightly more general in that small selection pressures operate favoring a specific gamete. If the selection differences are very small then continuity considerations dictate that the probabilities of fixation derivable from the Hill-Robertson studies should almost agree with the formulas of Table 1. Therefore, the dependence of r in any case does not occur in the combination Nr.

(e) Suppose N is large and r is small such that $2Nr = \gamma$ is moderate. The formulae of Table 1 then can be effectively approximated as follows:

Probability of fixation in the gamete in question is

$$\begin{aligned} \text{AB gamete} &\sim x_1 + \frac{\gamma}{1+\gamma}\tilde{D} \\ \text{Ab gamete} &\sim x_2 - \frac{\gamma}{1+\gamma}\tilde{D} \\ \text{aB gamete} &\sim x_3 - \frac{\gamma}{1+\gamma}\tilde{D} \\ \text{ab gamete} &\sim x_4 + \frac{\gamma}{1+\gamma}\tilde{D} \end{aligned} \tag{8}$$

where

$$\tilde{D} = x_2 x_3 - x_1 x_4$$

and x_1, x_2, x_3, x_4 represent the initial frequencies of the AB, Ab, aB and ab gamete respecitvely. Notice that if Nr is very large then $\gamma/1+\gamma)$ is near 1. In this circumstance we obtain the good approximations for the probability of fixation in the gamete indicated,

$$\begin{aligned} \text{AB} &\sim p_A\, p_B \\ \text{Ab} &\sim p_A\, p_b \\ \text{aB} &\sim p_a\, p_B \\ \text{ab} &\sim p_a\, p_b \end{aligned} \tag{9}$$

where $p_A = x_1 + x_2$ denotes the initial frequency of the A allele with similar meaning ascribed to p_a, p_B and p_b. The result expressed in (9), in essence, agrees with the corresponding classical result from deterministic theory.

II. *Rate of Approach to Fixation.* As pointed out in the Introduction, the probability f_n that the nth generation includes at least two types of gametes is of the order

$$f_n \sim c(\bar{\imath})\ \lambda^n$$

where $0 < \lambda < 1$ and $c(\bar{\imath})$ is a constant depending on the initial population vector $\bar{\imath}$ but not on the time index n. The value λ is referred to as the rate of approach to fixation. We can interpret $1/(1-\lambda)$ approximately as the expected number of elapsed generations required before the population comprises a single pure type.

For the model at hand the rate of approach to fixation (= homozygosity) is determined to be

$$\lambda = 1 - 1/2N. \tag{10}$$

It is perhaps surprising that the rate of approach to fixation is independent of the recombination fraction and the number of loci involved and coincides with the

rate of approach to fixation for the corresponding stochastic model of two alleles at one locus. This is also the case where the model is formulated involving a general fertility distribution (i.e., different from the Poisson progeny distribution implied by multinomial sampling, see section 3). This is not true of the rate at which the first allele is lost.

It is also worthwhile to underscore the fact that the expected value of $D^{(n)}$ tends to zero at a faster rate (provided $r > 0$), by a factor $1-r$, than the rate of approach to fixation.

III. *Probabilities of First Fixation of a Specified Allele.* In Table 2 below we record the probabilities of the events that the allele A, a, B or b respectively becomes fixed first. The following interpretations emerge from examination of the formulas in Table 2.

(a) The probabilities that any given allele fixes first are independent of the recombination fraction only in the special circumstance that initially there is zero linkage deviation, i.e., $D^{(0)} = 0$.

(b) The probability that the (A,a) locus fixes prior to the (B,b) locus is obviously

$$Q_A + Q_a = \frac{(i_1+i_3)\ (i_2+i_4)}{(i_1+i_2)\ (i_3+i_4) + (i_1+i_3)\ (i_2+i_4)} \tag{11}$$

Notice that this expression is always independent of r.

(c) The probabilities Q_A, Q_B, Q_a, Q_b generally depend on the initial population numbers (i_1, i_2, i_3, i_4) as a ratio of a cubic to a quadratic polynomial in these variables.

(d) Let x_1, x_2, x_3, x_4 denote as before the initial frequencies of the AB, Ab, aB, ab gametes respectively and let $\tilde{D} = x_2\, x_3 - x_1\, x_4$ represent the linkage deviation

TABLE 2

$(i_1\, i_2\, i_3\, i_4)$ is the initial population vector

Allele	Q = Probability of the Allele in Question Fixing First
A	$Q_A = \dfrac{i_1 i_2 + i_2 i_3 + \dfrac{(1-r)}{(2N-2)(1-r) - 2N} D + \dfrac{(i_1+i_3) r D}{(2N-2)(1-r) - 2N}}{(i_1+i_2)(i_3+i_4) + (i_1+i_3)(i_2+i_4)}$
a	$Q_a = \dfrac{i_3 i_4 + i_1 i_4 - \dfrac{(1-r)\ D}{(2N-2)(1-r) - 2N} - \dfrac{(i_1+i_3) r D}{(2N-2)(1-r) - 2N}}{(i_1+i_2)(i_3+i_4) + (i_1+i_3)(i_2+i_4)}$
B	$Q_B = \dfrac{i_1 i_3 + i_2 i_3 + \dfrac{(1-r)\ D}{(2N-2)(1-r) - 2N} + \dfrac{(i_1+i_2) r D}{(2N-2)(1-r) - 2N}}{(i_1+i_2)(i_3+i_4) + (i_1+i_3)(i_2+i_4)}$
b	$Q_b = \dfrac{i\ i_4 + i_1 i_4 - \dfrac{(1-r)\ D}{(2N-2)(1-r) - 2N} - \dfrac{(i_1+i_2) r D}{(2N-2)(1-r) - 2N}}{(i_1+i_2)(i_3+i_4) + (i_1+i_3)(i_2+i_4)}$

$(D = i_2 i_3 - i_1 i_4)$

function of the initial generation frequencies. For N large and $2Nr = \gamma$ moderate we obtain the approximate formulas

$$Q_A \sim \frac{(x_1+x_3)\,x_2 - \frac{\gamma}{2+\gamma}(x_1+x_3)\,\tilde{D} - \frac{1}{2+\gamma}\tilde{D}}{(x_1+x_2)\,(x_3+x_4) + (x_1+x_3)\,(x_2+x_4)}$$

$$Q_a \sim \frac{(x_1+x_3)\,x_4 + \frac{\gamma}{2+\gamma}(x_1+x_3)\,\tilde{D} + \frac{1}{2+\gamma}\tilde{D}}{(x_1+x_2)\,(x_3+x_4) + (x_1+x_3)\,(x_2+x_4)}$$

(12) $\qquad \tilde{D} = x_2 x_3 - x_1 x_4$

$$Q_B \sim \frac{(x_1+x_2)\,x_3 - \frac{\gamma}{2+\gamma}(x_1+x_2)\,\tilde{D} - \frac{1}{2+\gamma}\tilde{D}}{(x_1+x_2)\,(x_3+x_4) + (x_1+x_3)\,(x_2+x_4)}$$

$$Q_b \sim \frac{(x_1+x_2)\,x_4 + \frac{\gamma}{2+\gamma}(x_1+x_2)\,\tilde{D} + \frac{1}{2+\gamma}\tilde{D}}{(x_1+x_2)\,(x_3+x_4) + (x_1+x_3)\,(x_2+x_4)}$$

It is easy to verify that when γ is large (i.e., Nr large) the formulas simplify to

(13) $\quad Q_A \sim \frac{p_A p_B p_b}{p_A p_a + p_B p_b}, Q_a \sim \frac{p_a p_B p_b}{p_A p_a + p_B p_b}, Q_B \sim \frac{p_B p_A p_a}{p_A p_a + p_B p_b}, Q_b \sim \frac{p_b p_A p_a}{p_A p_a + p_B p_b}$

It is interesting to observe that the formulae in (13) of probabilities of fixation of some specified allele depend only on the initial allele frequencies and not on the initial gamete frequencies. We emphasize that this is only correct approximately provided Nr is sufficiently large. For moderate values of Nr but N large the formulas of (12) prevail and now the dependence on the initial gamete frequencies is significant.

Qualitative information concerning the monotone variation of, say, Q_A as a function of p_A, p_B, p_a, p_b can be readily extracted from examination of the formulas (13). We do not pursue this end here.

(e) When $r = 0$ (tight linkage i.e., no recombination) the formulas of Table 2 reduce to the following:

$$Q_A = \frac{i_1 i_2 + \frac{\Gamma}{2}}{C}, \qquad Q_a = \frac{i_3 i_4 + \frac{\Gamma}{2}}{C}$$

(14)

$$Q_B = \frac{i_1 i_3 + \frac{\Gamma}{2}}{C}; \qquad Q_b = \frac{i_2 i_4 + \frac{\Gamma}{2}}{C}$$

where $\Gamma = (i_2 i_3 + i_1 i_4)/2$ and $C = (i_1+i_2)(i_3+i_4) + (i_1+i_3)(i_2+i_4)$.

Notice in this case that the quantities in (14) involve only ratios of quadratics in the variables i_1, i_2, i_3, i_4 (cf. paragraph (c) above) in contrast with Table 2 where the numerators involve cubics.

IV. *Rate of Loss of an Allele.* (a) The standard analysis of finite Markov chains informs us that the probability π_n that all four gametes are present in the nth generation is of the order

(15) $\qquad \pi_n \sim B(\bar{\imath})\,\mu^n$

for some μ $(0 < \mu < 1)$ where $B(\bar{\imath})$ is a constant depending on the initial population numbers but is independent of n. Thus π_n decreases at a geometric rate by a factor μ per generation. We refer to the constant $1-\mu$ as the *rate of loss of an allele.* The explicit value of μ can be determined. It is

$$\mu = \left(1 - \frac{1}{2N}\right) x_0 \tag{16}$$

where x_0 is the largest positive root of the cubic equation

$$-x^3 + x^2E + xF + G = 0 \tag{17}$$

where the constants E, F and G are explicitly the burdensome expressions

$$\begin{aligned}(2N^2)E &= [2N(1-r) + (2r-3)]\,(1-r)\,(N-1) \\ &\quad + (N-1)\,(1-2r)\,(2N-r) + 2N^2(1+r) + (r-4N)\,r\,, \\ (2N^3)F &= (1-r)\,(N-1)\,\{[(2r-3) + 2N(1-r)]\,[r - N(1+r)] \\ &\quad + (r-2N)N + (2N-3)\,(1-r)\,[(3r-1) - (1-2r)N]\}\,, \\ (2N^3)G &= (1-r)^3\,(2N-3)\,(N-1)^2.\end{aligned}$$

For any specification of $2N$ and r, (17) is a concrete cubic equation. Thus for no linkage i.e., $r = 1/2$ the cubic (17) becomes

$$-x^3 + x^2\left(\frac{14N^2 - 14N + 5}{8N^2}\right) + x\,\frac{(N-1)\,(-14N^2 + 18N - 7)}{16N^3} + \frac{(2N-3)\,(N-1)^3}{16N^3} = 0$$

which does not possess rational roots.

It can be proved that as r varies from 0 to 1, μ decreases from

$$\left(1 - \frac{1}{2N}\right) \quad \text{to} \quad \left(1 - \frac{1}{2N}\right)^2$$

Numerical calculation of μ have been made in Table 3, below for various choices of $2N$ and r.

It is of interest to compare the values of μ with the corresponding rate $\mu_{q,p}$ of the loss of q out of p alleles $(0 \leq q < p-1)$ in a one locus p allele population of $2N$ haploid individuals reproducing by binomial sampling. In other words, the probability that the population in the nth generation includes at least $p-q$ alleles is of the order $\sim C[\mu_{q,p}]^n$ where $0 < \mu_{q,p} < 1$ and C is a constant depending on the initial population structure but not on n. We found (KARLIN and MCGREGOR 1965a) the value to be

$$\mu_{q,p} = (2N)!/(2N)^{p-q}.$$

(In the cited paper we actually determined $\mu_{q,p}$ for a p allele haploid model with general fertility distribution for the number of offspring per mating.)

We have the inequalities

$$\left(1 - \frac{1}{2N}\right) = \mu_{2,4} > \mu > \mu_{1,4} = \left(1 - \frac{1}{2N}\right)\left(1 - \frac{2}{2N}\right) \tag{18}$$

holding independent of the value of r provided only that r is positive. Thus the probability of maintaining all four gametes in the nth generation of the two locus model decreases to zero at a slower rate than the probability of having at least three alleles of a 4 allele haploid population present in the nth generation. However, the rate at which two alleles are lost from a 4 allele haploid population

exceeds the value of μ. This could be argued on intuitive grounds as well. In the two locus model the instant that one allele fixes, then two gametes are lost simultaneously and therefore indeed $(1 - 1/2N) = \mu_{2,4} > \mu$. Of course for $r = 0$, $\mu = \mu_{0,2} = \mu_{2,4}$ as expected.

A quantitative comparison of μ and $\mu_{2,4}$ and $\mu_{1,4}$ is available from (18) and (16).

(a) We can use the quantity μ to determine how many generations may pass in which the probability of the event that all gametes are represented in the population remains greater than α (e.g., $\alpha = .1$, $\alpha = .05$ or $\alpha = .01$). The solution of this problem is obtained simply by determining n to satisfy $\pi_n = \alpha$. By (15) this is approximately equivalent to $C(i)\mu^n = \alpha$ and therefore $n \sim \log \alpha / \log \mu$ for α sufficiently small since the effect of the initial population vector (expressed through $C(\bar{\imath})$) is negligible compared to $\log\alpha/\log \mu$. The above computations play a role in appraising the influence of linkage on evolutionary change.

(b) The final remark of this section concerns the variance of the linkage deviation function $D^{(n)}$ as a function of time. It can be proved that $\mathrm{Var}(D^{(n)})$ tends to zero at the asymptotic rate

$$\mathrm{Var}(D^{(n)}) \sim \gamma(\bar{\imath})\ \mu^n \tag{19}$$

where $\gamma(\bar{\imath})$ is a positive constant depending on the initial conditions. It is not difficult to establish the relationship

$$\mu > \left(1 - \frac{1}{2N}\right)(1-r) \qquad \text{for all } r > 0 \tag{20}$$

Comparing (20) and (7) we see that $\mathcal{E}(D^{(n)})$ tends to zero at a faster rate than $\mathrm{Var}(D^n)$. Hence while both may be close to zero after the effect of the initial conditions wears off, the variance of $D^{(n)}$ is of larger order than its expectation. We infer from this that the actual value of $D^{(n)}$ (as opposed to $\mathcal{E}(D^n)$) in a relative sense is significantly different from zero although equally likely of positive or negative sign.

The contrast in the approach of $\mathcal{E}(D^{(n)})$ and $\mathrm{Var}(D^n)$ to zero invalidates the assumption frequently made (HILL and ROBERTSON 1966, LATTER 1966) that the initial gamete frequencies are in linkage equilibrium, i.e., have zero linkage deviation.

The Case of a General Progeny Distribution. The model constructed in the

TABLE 3

Values of μ (rate of loss of first allele from the population)

$r =$	.0	.01	.05	.10	.20	.30	.40	.45	.49	.50	.70	.90
$2N$ 6	.833	.825	.795	.764	.721	.697	.683	.679	.676	.676	.668	.667
$2N$ 10	.900	.891	.863	.842	.823	.815	.813	.812	.811	.811	.810	.810
$2N$ 16	.938	.928	.903	.889	.879	.877	.876	.876	.875	.875	.875	.875
$2N$ 20	.950	.941	.920	.909	.904	.903	.903	.903	.903	.903	.903	.903
$2N$ 30	.967	.958	.942	.937	.935	.934	.934	.934	.934	.934	.934	.934
$2N$ 50	.980	.971	.962	.960	.960	.960	.960	.960	.960	.960	.960	.960
$2N$ 100	.990	.983	.980	.980	.980	.980	.980	.980	.980	.980	.980	.980

Introduction can be extended to allow the possibility of a general progeny distribution for the number of offspring contributed per mating.

We denote the state space by Δ. A finite state M.C. defined on $\Delta = \{\bar{\imath} = (i_1, i_2, i_3, i_4) \mid i_v \text{ integers} \geq 0 \text{ and } \sum i_v = 2N\}$ and governing the fluctuations of the gamete populations under the influence of general fertility and the recombination fraction r is set up in the framework of the direct product branching process as follows (see KARLIN and MCGREGOR 1964a). Let $f(s) = \sum_{k=0}^{\infty} a_k s^k$ denote the probability generating function (p.g.f.) of the number of offspring resulting from a union of two gametes. Thus, for example, a union of AB and Ab produces a random number of offspring with p.g.f. $f(s)$ where each gamete from this union is either AB or Ab with probability 1/2 each. Let $g(x_1, x_2, x_3, x_4)$ denote the joint p.g.f. of the numbers of AB, Ab, aB and ab gametes respectively created after one generation of reproduction by random union of gametes, where the parent population consists of i_1, i_2, i_3, i_4, AB, Ab, aB and ab gametes respectively. Considering all the possibilities we obtain

$$
\begin{aligned}
(21)\quad & g(x_1, x_2, x_3, x_4) \\
& = f^{i_1^2}(x_1)\, f^{2i_1i_2}\left(\frac{x_1+x_2}{2}\right) f^{2i_1i_3}\left(\frac{x_1+x_3}{2}\right) f^{2i_1i_4}\left((1-r)\,\frac{(x_1+x_4)}{2} + \frac{r}{2}\,(x_2+x_3)\right) \\
& \times f^{i_2^2}(x_2)\, f^{2i_2i_3}\left((1-r)\,\frac{(x_2+x_3)}{2} + \frac{r(x_1+x_4)}{2}\right) f^{2i_2i_4}\left(\frac{x_2+x_4}{2}\right) \\
& \times f^{i_3^2}(x_3)\, f^{2i_3i_4}\left(\frac{x_3+x_4}{2}\right) f^{i_4^2}(x_4)\ .
\end{aligned}
$$

(Notice that a union of AB and ab can produce offspring of all kinds due to the possibility of recombination, and similarly for a union of Ab and aB.) The x's appearing in (21) are variables of the generating function and are not to be confused with the use of x's as gamete frequencies in (8).

A frequency model is induced by considering only those realizations of the process (21), which yield $2N$ gametes. In other words, we condition the outcome so that the total number of offspring is $2N$. Equivalently we postulate that the fluctuations of the gamete populations is governed by a Markov chain on the state space Δ with transition probability matrix

$$
(22)\qquad P_{i,j} = \frac{\text{coefficient } x_1^{j_1} x_2^{j_2} x_3^{j_3} x_4^{j_4} \text{ in } g(x_1, x_2, x_3, x_4)}{\text{coefficient } z^{2N} \text{ in } g(z, z, z, z) = f^{4N^2}(z)}
$$

for $i = (i_1, i_2, i_3, i_4) \ \varepsilon\ \Delta$ and $j = (j_1, j_2, j_3, j_4) \ \varepsilon\ \Delta$.

When the offspring distribution is Poisson, i.e., $f(x) = e^{a(x-1)}$ the transition probability matrix (22) reduces to the case of multinomial sampling with explicit transition probability law given in (5). Actually, the multinomial sampling model of the Introduction proposed to study fluctuations of gene frequency

necessarily implies that the offspring distribution per parent is Poisson (see KARLIN and MCGREGOR 1964a, 1968 for further discussion of this point). KOJIMA and KELLEHER (1962) pointed out the inadequacy of the Poisson assumption and the superior fit for the progeny distribution (especially relevant to data on human populations) of the negative binomial family of probability generating functions.

We now describe a series of results partly extending those of Section 2. The previous formulas need to be appropriately modified to include the non-Poisson character of the progeny distribution function.

(I) *Probability of Fixation.* Let the initial population vector be $\bar{\imath} = (i_1, i_2, i_3, i_4)$ and let $D = i_2 i_3 - i_1 i_4$,

(23)

Gamete	Probability of fixation in the gamete in question
AB	$\frac{i_1}{2N} + \frac{r}{(2N)^2(1-\beta)} D$
Ab	$\frac{i_2}{2N} - \frac{r}{(2N)^2(1-\beta)} D$
aB	$\frac{i_3}{2N} - \frac{r}{(2N)^2(1-\beta)} D$
ab	$\frac{i_4}{2N} + \frac{r}{(2N)^2(1-\beta)} D$

where

$$\beta = 4N^2(1-r)\,\lambda_{21} + \frac{(\lambda_{22}-\lambda_{21})(1-2r)}{2},$$

$$\lambda_{21} = \frac{\text{coeff. } z^{2N-2} \text{ in } f^{4N^2-2}(z)\,[f'(z)]^2}{\text{coeff. } z^{2N} \text{ in } f^{4N}(z)},$$

and

$$\lambda_{22} = \frac{\text{coeff. } z^{2N-2} \text{ in } f^{4N^2-1}(z)\, f''(z)}{\text{coeff. } z^{2N} \text{ in } f^{4N^2}(z)}.$$

In comparing (23) and (4) we observe that the two sets of formulas differ only in the coefficient of rD. Thus the general progeny distribution per mating will affect the probabilities of fixation only if the initial linkage deviation value D is non-zero and then the fertility component contributes a scale factor multiplying D.

It is interesting to consider the values of β, λ_{21}, and λ_{22} for some typical distributions (Table 4). Notice that the Poisson distribution provides a value of β intermediate between that for the binomial family and the negative binomial family.

(II) *Rate of Approach to Fixation.* We can prove that the rate of approach to fixation is of the order (see (10)), $1-\lambda$ where

$$\lambda = 4N^2\,\lambda_{21} + (\lambda_{22} - \lambda_{21})/2 \tag{24}$$

TABLE 4

Values of β, λ_{21} and λ_{22} for some typical distributions

Progeny Distribution	Probability Generating Functions	λ_{21}	λ_{22}	β	for N large $\beta\sim$
Poisson	$f(x)=e^{a(x-1)}$, $a>0$	$\frac{2N-1}{(2N)\,4N^2}$	$\frac{2N-1}{2N(4N^2)}$	$(1-\frac{1}{2N})(1-r)$	$(1-\frac{1}{2N})(1-r)$
Negative Binomial	$f(x)=\frac{(1-p)^\alpha}{(1-px)^\alpha}$ $0<\alpha, 0<p<1$	$\alpha\frac{2N-1}{2N(4N^2\alpha+1)}$	$\frac{(\alpha+1)(2N-1)}{2N(4N^2\alpha+1)}$	$\frac{\alpha\,2N(2N-1)}{4N^2\alpha+1}(1-r)$ $+\frac{(2N-1)}{(4N)(4N^2\alpha+1)}(1-2r)$	$\frac{\alpha(2N)(2N-1)}{4N^2\alpha+1}(1-r)$
Binomial	$(1-p+ps)^\gamma$, $\gamma>0$ integer, $0<p<1$	$\frac{\gamma(2N-1)}{2N(4N^2\gamma-1)}$	$\frac{(\gamma-1)(2N-1)}{2N(4N^2\gamma-1)}$	$\frac{\gamma(2N-1)\,2N}{(4N^2\gamma-1)}(1-r)$ $-\frac{(2N-1)}{\gamma\,2N(4N^2\gamma-1)}(1-2r)$	$\frac{\gamma(2N)(2N-1)}{(4N^2\gamma-1)}(1-r)$

and again this value coincides with the rate of approach of fixation for the one locus two allele random union of gamete model with general fertility probability generating function $f(x)$.

(c) *Mean Change of Disequilibrium Function.* The expected change of the linkage deviation function can be computed exactly. We find that

$$\mathcal{E}(D^{(1)}) = \beta D^{(0)}$$

and generally

$$\mathcal{E}(D^{(n)}) = \beta^n D^0$$

Inspection of Table 4 reveals that the mean value of $D^{(n)}$ tends to zero at a faster rate when the progeny distribution is binomial than when it is Poisson, even though the expected number of offspring is the same. Furthermore the rate of approach of the mean of $D^{(n)}$ to zero in the Poisson case is faster than for the case where the progeny distribution is a negative binomial law.

We have not been able to determine the rate of loss of the first allele from the population in the case of a general fertility probability generating function. The formulae for the probabilities of which locus fixes first given in (11) still hold under the more general conditions.

DISCUSSION

KIMURA (1963) has obtained the probability of fixation in a given gamete for the case of random union of zygotes; that is where recombination occurs only in homologous chromosomes from one individual. The fact that our model, the straightforward generalization of the classical Wright-Fisher model gives slightly different results is merely a manifestation of the fact that in the presence of recombination the stochastic process of random union of gametes is not equivalent to the stochastic process of random union of zygotes.

From Table 1, the probability of fixation in a particular gamete for the multinomial sampling model depends on the initial population only through the frequency of that gamete and the initial linkage deviation $D^{(0)}$. A comparison of the expectation and variance of $D^{(n)}$ has been made and the variance of $D^{(n)}$ can be shown to converge to zero at a slower rate than its expectation. These two facts taken together, indicate that the assumption made by HILL and ROBERTSON (1966) and LATTER (1966), that $D^{(0)} = 0$, is a very stringent one. In fact one of the most interesting problems arising in populations such as have been considered here is that of the increase of an initially rare gamete. In this case it is obvious that one can have D initially near its maximum ($\frac{1}{4}$).

Again from Table 1, the probabilities of fixation depend on r (if D is not initially zero) but not through the quantity Nr. Although the simulation studies made by HILL and ROBERTSON and LATTER are slightly more general, in that they include selection, we expect that if the selection pressures are small enough (say less than $o(1/N)$) then continuity would dictate that our results should almost agree with those of the above authors. However, both these studies claim that the probabilities of fixation depend on r through Nr, in obvious disagreement with our findings. In both simulation studies mentioned above, use has been made of

the function $u(p_0)$, the chance of eventual fixation of a gene with initial frequency p_0. However, this function is calculated on the bases of the one locus diffusion approximation, and the validity of its use in a two locus situation with nonzero recombination fraction seems questionable.

A number of important questions remain unanswered. We hope to attempt a simulation study including selection and based on the models and results obtained above. There remains the interesting problem of the rates and probabilities of loss of an allele from the population in the general fertility case. We also hope to attack the analogous problems with three or more loci. It should be mentioned that the above treatment is easily modified to take account of any linear evolutionary pressures, e.g., migration and mutation, cf. KARLIN and MCGREGOR (1965a). The analogous problems for 2 loci and any number of alleles also present no difficulties. It seems feasible that a rigorous approach to the corresponding problems of diffusion approximation can also be based on the findings of this paper, perhaps along the lines of KARLIN and MCGREGOR (1965a).

SUMMARY

A stochastic treatment of a two locus random mating population model taking account of recombination is given. This seems to be the first time exact results have been obtained for such quantities as the probability of fixation in a given allele and, the rate of loss of a given allele from the population. The probability of fixation in a given gamete in the case of general fertility is also given for the first time, although KIMURA (1963) had previously obtained this for a different model in the special case of a Poisson progeny distribution function. The rate of decrease of the expectation and variance of the linkage disequilibrium function are also determined and their rather surprising consequences discussed. The results are compared with those from recent simulation studies of HILL and ROBERTSON (1966) and the work of KIMURA (1963).

LITERATURE CITED

BODMER, W. F., M. FELDMAN, and KARLIN, 1968 *Theoretical population genetics.* (manuscript in preparation).

BODMER, W. F., and J. FELSENSTEIN, 1967 Linkage and selection: theoretical analysis of the deterministic two locus random mating model. Genetics **57**: 237–265.

BODMER, W. F., and P. A. PARSONS 1962 Linkage and recombination in evolution. Advan. Genet. **11**: 1–100.

EWENS, W. J., 1966 Linkage and the evolution of dominance. Heredity **21**: 363–370. —— 1967 The probability of fixation of a mutant; the two locus case. (in press).

HILL, W. G., and A. ROBERTSON, 1966 The effect of linkage on limits to artificial selection. Genet. Res. **8**: 269–294.

JAIN, S. K., and R. W. ALLARD, 1965 The nature and stability of equilibria under optimizing selection. Proc. Natl. Acad. Sci. U.S. **54**: 1436–1443. —— 1966 The effects of linkage, epistasis and inbreeding on population changes under selection. Genetics **53**: 633–659.

KARLIN, S., 1966 *A First Course in Stochastic Processes.* Academic Press, New York and London.

KARLIN, S., and J. MCGREGOR, 1964a Direct product branching processes and related Markoff chains, Proc. Natl. Acad. Sci. U.S. **51**: 598–602. —— 1964b On some stochastic models in genetics, 245–279, *Stochastic Models in Medicine and Biology* edited by JOHN GURLAND. University of Wisconsin Press, Madison. —— 1965a Direct product branching processes and related induced Markoff chains I. Calculations of rates of approach to homozygosity. Bernoulli, Bayes, Laplace Anniversary Volume, Springer-Verlag, pp. 11–145. ——1965b The number of mutant forms maintained in a population, Proc. 5th Berkeley Symposium on Prob. and Stat., University of California Press (in press). —— 1968 The role of the Poisson progeny distribution in population genetics models. Mathematical Biosciences (in press).

KARLIN, S., J. MCGREGOR, and W. F. BODMER, 1965 The rate of production of recombinants between linked genes in finite populations. Proc. 5th Berkeley Symposium on Prob. and Stat. University of California Press, (in press).

KEMPTHORNE, O., 1957 *An Introduction to Genetic Statistics.* Wiley, New York.

KIMURA, M., 1955 Solution of a process of random genetic drift with a continuous model. Proc. Natl. Acad. Sci. U.S. **41**: 144–150. —— 1956 A model of genetic system which leads to closer linkage by natural selection. Evolution **10**: 278–287. —— 1963 A probability method for treating inbreeding schemes especially with linked genes. Biometrics **19**: 1–17. —— 1965 Attainment of quasi linkage equilibrium when gene frequencies are changing by natural selection. Genetics **52**: 875–890.

KIMURA, M., and J. F. CROW, 1963 The measurement of effective population number. Evolution **17**: 279–288.

KOJIMA, K., 1965 The evolutionary dynamics of two-gene systems, pp. 197–220. *Computers in Biomedical Research*, Vol. 1, edited by Ralph Stacy and Bruce Waxman, Academic Press.

KOJIMA, K., and T. M. KELLEHER, 1961 Changes of mean fitness in random mating population when epistasis and linkage are present. Genetics **16**: 527–540. —— 1962 Survival of mutant genes. Am. Naturalist **96**: 329–346.

KOJIMA, K., and H. E. SCHAFER, 1964 Accumulation of epistatic gene complexes. Evolution **18**: 127–129.

LATTER, B. D. H., 1966 The interaction between effective population size and linkage intensity under artificial selection. Genet. Res. **7**: 313–323.

LEWONTIN, R. C., 1964a The interaction of selection and linkage. I. General considerations; heterotic models. Genetics **49**: 49–67. —— 1964b The interaction of selection and linkage II, Optimum models. Genetics **50**: 757–782. —— 1964c The role of linkage in natural selection. *Proc. 11th Intl. Cong. Genet.* **3**: 517–525.

LEWONTIN, R. C., and K. KOJIMA, 1960 The evolutionary dynamics of complex polymorphisms. Evolution **14**: 458–472.

LI, C. C., 1955 *Population Genetics.* University of Chicago Press.

MALÉCOT, G., 1948 *Les mathématiques de l'hérédité.* Masson, Paris. —— 1951 Un traitement stochastique des problèmes lineaires (mutation, linkage, migration) en génétique de population. Ann. Univ. Lyon, Sciences, Section A **14**: 79–000. —— 1959a Les modèles stochastiques en génétique de population. Pub. ISUP, Vol. VIII, fasc. 3. —— 1959b La génétique de populations. II. Modèles stochastiques. Publ. Inst. Statist. Paris **8** F3, 173–210.

MORAN, P. A. P., 1962 *The Statistitical Processes of Evolutionary Theory.* Clarendon Press, Oxford.

MORAN, P. A. P., and G. A. WATTERSON, 1959 The genetic effects of family structure in natural populations. Austral. J. Biol. Sci. **12**: 1–15.

PARSONS, P. A., 1963a Complex polymorphisms where the coupling and repulsion double heterozygote viabilities differ. Heredity **18**: 369–374. —— 1963b Polymorphisms and the balanced polygenic combination. Evolution **17**: 564–574.

SINGH, M., and R. LEWONTIN, 1966 Stable equilibria under optimizing selection. Proc. Natl. Acad. Sci. U.S. **56:** 1345–1348.

WATTERSON, G. A., 1959a Non-random mating and its effect on the rate of approach to homozygosity. Ann. Human Genet. **23:** 204–220. —— 1959b A new genetic population model and its rate of approach to homozygosity. Ann. Human Genet. **23:** 221–232.

WRIGHT, S., 1931 Evolution in Mendelian populations. Genetics **16:** 97–159. —— 1933 Inbreeding and recombination. Proc. Natl. Acad. Sci. U.S. **19:** 420–433. —— 1952 The genetics of quantitative variability. pp. 5–41. *Quantitative Inheritance* London, Her Majesty's Stationary Office.