

# THE EFFECT OF EPISTASIS IN THE STANDARD TWO-LOCUS MODEL

By M. A. B. DEAKIN\*

[Manuscript received 10 March 1972]

## Abstract

Moran (1964) proved three negative results for the standard two-locus model of population genetics: (1) that mean fitness was not a function of the allelic frequencies alone; (2) that stationary states of the population do not, in general, correspond to stationary values of the mean fitness; (3) the mean fitness of a population may decrease from one generation to the next. Ewens (1969*a*, 1969*b*) showed that for an additive fitness (i.e. non-epistatic) model none of these results could apply. The present paper shows the crucial role of the non-epistatic hypothesis in the establishing of Ewens' results.

## I. INTRODUCTION

We consider the standard two-locus diploid model. Two loci each have two alleles. We term these  $A,a$  and  $B,b$  respectively. The fitnesses of the various geno-

TABLE 1  
FITNESS COEFFICIENTS IN THE GENERAL CASE

|      | $AA$     | $Aa$                           | $aa$     |
|------|----------|--------------------------------|----------|
| $BB$ | $w_{11}$ | $w_{14}$                       | $w_{33}$ |
| $Bb$ | $w_{12}$ | $w_{14} = w_{23}$<br>$= w = 1$ | $w_{34}$ |
| $bb$ | $w_{22}$ | $w_{24}$                       | $w_{44}$ |

types which may be formed are given in Table 1. We make the conventions that

$$w_{ij} = w_{ji}, \quad (1)$$

and

$$w_{14} = w_{23} = w. \quad (2)$$

We may, without loss of generality, set  $w = 1$ , as the only case in which this cannot be done ( $w = 0$ ) is trivial anyhow. This further convention will be adopted.

\* Mathematics Department, Papua New Guinea Institute of Technology, Lae, New Guinea. On leave from Monash University, Clayton, Victoria.

The four possible gametotypes  $AB$ ,  $Ab$ ,  $aB$ ,  $ab$  have frequencies  $x_1$ ,  $x_2$ ,  $x_3$ ,  $x_4$  respectively and

$$\sum_{i=1}^4 x_i = 1. \quad (3)$$

A quantity  $D$  is defined by

$$D = x_1x_4 - x_2x_3. \quad (4)$$

We further define

$$w_i = \sum_{j=1}^4 w_{ij}x_j, \quad (5)$$

and

$$\bar{w} = \sum_{i=1}^4 w_i x_i. \quad (6)$$

The coefficient of recombination between the loci is taken to be  $R$ .

The system follows the standard equations

$$\bar{w}\Delta x_i = x_i(w_i - \bar{w}) - \delta_i R D, \quad (7)$$

where  $\delta_i$  is defined as follows:

$$\delta_i = \begin{cases} 1 & \text{if } i = 1, 4 \\ -1 & \text{if } i = 2, 3 \end{cases} \quad (8)$$

Equilibrium is attained if

$$x_i(w_i - \bar{w}) = \delta_i R D. \quad (9)$$

## II. EPISTASIS

The quantities  $E_i$  defined by

$$E_i = \sum_{j=1}^4 \delta_j w_{ij} \quad (10)$$

were first introduced by Fisher (1918). They are termed the "coefficients of epistasis". We shall see below that they measure the extent to which the fitnesses are not additive over the loci. Set

$$\left. \begin{aligned} w'_{ii} &= w_{ii} - \delta_i E_i \\ w'_{ij} &= w_{ij}, \text{ if } i \neq j \end{aligned} \right\} \quad (11)$$

Consider now the equations

$$\left. \begin{aligned} u_1 + v_1 &= w'_{11} \\ u_2 + v_1 &= w'_{12} \\ u_3 + v_1 &= w'_{22} \\ u_1 + v_2 &= w'_{13} \\ u_2 + v_2 &= w'_{14} = 1 \\ u_3 + v_2 &= w'_{24} \\ u_1 + v_3 &= w'_{33} \\ u_2 + v_3 &= w'_{34} \\ u_3 + v_3 &= w'_{44} \end{aligned} \right\} \quad (12)$$

We seek to solve these for the six quantities  $u_i$ ,  $v_i$ . This may be done as follows. First set  $u_2 = v_2 = \frac{1}{2}$ . Substitute progressively into the equations. Unique values

are obtained for  $u_1, u_3, v_1, v_3$ . The system of nine equations in six unknowns is not in fact overdetermined, as the four equations

$$\begin{aligned}u_1 + v_1 &= w'_{11} \\ u_3 + v_1 &= w'_{22} \\ u_1 + v_3 &= w'_{33} \\ u_3 + v_3 &= w'_{44}\end{aligned}$$

which are not required for the solution are satisfied automatically in view of the relations

$$\sum_{j=1}^4 \delta_j w'_{ij} = 0, \quad (13)$$

which follow from equations (10), (11).

The system in fact is actually underdetermined in that  $u_2, v_2$  are arbitrary provided that

$$u_2 + v_2 = 1.$$

The convention  $u_2 = v_2 = \frac{1}{2}$  is a convenient one to adopt.

The fact that equations (12) may be solved for the  $u_i, v_i$  demonstrates the complete equivalence of the quantities  $w'_{ij}$  and the additive fitnesses assumed by Ewens (1969*a*, 1969*b*). In particular if  $E_1 = E_2 = E_3 = E_4 = 0$ , fitness is additive. Further, if fitnesses are additive—i.e. as given in Table 2, all the quantities  $E_i$  are

TABLE 2  
FITNESS COEFFICIENTS IN THE NON-EPISTATIC CASE

|           | <i>AA</i>   | <i>Aa</i>   | <i>aa</i>   |
|-----------|-------------|-------------|-------------|
| <i>BB</i> | $u_1 + v_1$ | $u_1 + v_2$ | $u_1 + v_3$ |
| <i>Bb</i> | $u_2 + v_1$ | $u_2 + v_2$ | $u_2 + v_3$ |
| <i>bb</i> | $u_3 + v_1$ | $u_3 + v_2$ | $u_3 + v_3$ |

zero. Thus Ewens' additive fitness model may be described as non-epistatic. In the more general model, the quantities  $E_i$  measure the extent to which additivity is not achieved. In this case, equations (11) determine the values of the  $E_i$  uniquely.

### III. RESULTS OF MORAN AND EWENS

Moran (1964) has demonstrated that in general:

- (1)  $\bar{w}$  cannot be expressed as a function of  $P (=x_1+x_2)$ ,  $p (=x_1+x_3)$  alone;
- (2) the solutions of equations (9) do not, in general, occur at stationary values of  $\bar{w}$ ;
- (3) equations (7) need not lead to increase in the value of  $\bar{w}$  from one generation to the next.

These results show the inapplicability of Wright's (1932) concept of the "adaptive topography" to the general two locus model.

Ewens (1969) has shown that when fitnesses are assigned as in Table 2:

- (1)  $\bar{w} = w(P, p)$ ;
- (2)  $\bar{w}$  is stationary at equilibrium;
- (3)  $\bar{w}$  increases from one generation to the next.

Thus the adaptive topography concept is completely valid when epistasis is not present.

A number of attempts have been made to extend this result but with very little success. See, in this connection, the papers by Kimura (1966), Moran (1967), Ewens (1969b), and Arunachalam (1970).

The purpose of this paper is to demonstrate that except for a few trivial and biologically implausible cases, Ewens' results cannot apply to any epistatic model. Specifically we prove:

- (1) that mean fitness cannot be a function of allelic frequencies alone unless all the coefficients of epistasis are zero (Section V);
- (2) that a necessary condition for  $\bar{w}$  to be stationary at a polymorphic equilibrium is that an exact equality relation hold between the four coefficients of epistasis (Section VI);
- (3) if epistasis is present, there are necessarily situations from which  $\bar{w}$  will decrease unless:
  - (i)  $\bar{w}$  has no polymorphic maxima;
  - (ii) equations (7) have no stable polymorphic solution (Section VII).

These results are converses, in the mathematical sense of that term, of Ewens' results. They will be referred to as such in the sections to follow.

#### IV. MODIFICATION OF THE BASIC EQUATIONS

We insert equations (11) into equations (5), (6). There results:

$$w_i = w'_i + \delta_i E_i x_i, \quad (14)$$

where

$$w'_i = \sum_{j=1}^4 w'_{ij} x_j, \quad (15)$$

and

$$\bar{w} = \bar{w}' + \mathcal{E} \quad (16)$$

where

$$\bar{w}' = \sum_{i=1}^4 w'_i x_i, \quad (17)$$

and

$$\mathcal{E} = \sum_{i=1}^4 \delta_i E_i x_i^2 \quad (18)$$

With this notation, the governing equations of the system become

$$(\bar{w}' + \mathcal{E}) \Delta x_i = x_i (w'_i - \bar{w}') + x_i (\delta_i E_i x_i - \mathcal{E}) - \delta_i R D. \quad (19)$$

Equilibrium is achieved when

$$x_i (w'_i - \bar{w}') + x_i (\delta_i E_i x_i - \mathcal{E}) = \delta_i R D. \quad (20)$$

## V. THE CONVERSE OF EWENS' FIRST RESULT

We show that the condition

$$E_1 = E_2 = E_3 = E_4 = 0 \quad (21)$$

is necessary if  $w$  is to be a function of  $P, p$  alone. Ewens (1969*a*, 1969*b*) has already demonstrated the sufficiency of condition (21).

$\bar{w}$  ( $= \bar{w}' + \mathcal{E}$ ) is a function of three independent variables as the four quantities  $x_i$  are connected by equation (3). We may choose these variables to be  $P, p$  and  $x_1 + x_4$ . For  $\bar{w}$  to depend upon  $P, p$  alone, we must show  $w$  to be independent of  $x_1 + x_4$ . Ewens has already shown that  $\bar{w}'$  is independent of  $x_1 + x_4$ . Thus we need examine only the quantity  $\mathcal{E}$ .

It may readily be shown that

$$\begin{aligned} \mathcal{E} &= E_1\{1 - P - p - (x_1 + x_4)\}^2 - E_2\{1 + P - p - (x_1 + x_4)\}^2 \\ &= -E_3\{1 - P + p - (x_1 + x_4)\}^2 + E_4\{1 - P - p + (x_1 + x_4)\}^2 \\ &= f(P, p) - 2(x_1 + x_4)\{E_1(1 - P - p) - E_2(1 + P - p) \\ &\quad - E_3(1 - P + p) - E_4(1 - P - p)\} + (x_1 + x_4)^2(E_1 - E_2 - E_3 + E_4). \end{aligned}$$

The right-hand side of this equation can only be independent of  $x_1 + x_4$  if

$$E_1(1 - P - p) - E_2(1 + P - p) - E_3(1 - P + p) - E_4(1 - P - p)$$

is identically zero, and if

$$E_1 - E_2 - E_3 + E_4 = 0.$$

The first condition reduces to  $E_1 - E_4 = 0$ ,  $E_2 = 0$ ,  $E_3 = 0$ , and equation (21) readily follows.

## VI. THE CONVERSE OF EWENS' SECOND RESULT

We now consider the question of whether  $\bar{w}$  can attain a stationary value at equilibrium.

In the trivial cases  $R = 0$ , and  $w = 0$ , this will necessarily be achieved as a corollary of a result due to Scheuer and Mandel (1959). It is also possible that equilibria of the two-locus model may coincide with stationary values of  $\bar{w}$  in cases in which one allele becomes extinct (see Section VII below). We here show that fully polymorphic equilibria cannot, in practice, be stationary points of  $\bar{w}$  otherwise.

It may readily be shown that if mean fitness is to be stationary, we must satisfy the equations

$$\bar{w}_i = \bar{w} \quad (22)$$

at a polymorphic equilibrium. By equations (9) therefore we must have at equilibrium  $D = 0$  if fitness is to be stationary. Thus, by equations (20), we require

$$x_i(w'_i - \bar{w}') + x_i(\delta_i E_i x_i - \mathcal{E}) = 0 \quad (23)$$

at equilibrium.

Equations (23) may be written

$$w'_i + \delta_i E_i x_i = \bar{w}' + \mathcal{E}, \quad (24)$$

which are best solved for the equilibrium values of  $x_i$  by rewriting them as

$$w'_i + \delta_i E_i x_i - \lambda = 0, \quad (25)$$

and using the auxiliary equation (3). This gives a system of five linear equations in the five unknowns  $x_i$ ,  $\lambda$ .  $\lambda$  is interpretable as the equilibrium value of  $\bar{w}$ .

The system composed of equations (25), (3) may be solved by the application of Cramer's rule. In general it possesses a unique solution. In this case and if this solution satisfies the constraints  $x_i > 0$  for all  $i$ , then a stationary value of  $\bar{w}$  may be attained. However, the equilibrium values must also satisfy the equation  $D = 0$ , which may be translated into a (complicated) equality which must hold among the fitness coefficients. This possibility may be neglected as implausible due to the necessity of satisfying a *precise* equality. The assumption that such equalities do not hold has been termed by Hardin (1960) the "axiom of inequality". It does not hold in certain cases where we have good reason for postulating an equality. In the present instance, the case  $E_1 = E_2 = E_3 = E_4 = 0$  provides a case in which equality holds. We have no reason, however, to postulate equality relations between the quantities  $E_i$  in the case in which they do not vanish, and thus the axiom will be taken to apply when epistasis is present. Where the solution involves a negative value of one of the  $x_i$ , a stationary value of  $\bar{w}$  cannot be attained even in theory.

The axiom of inequality may be used to exclude cases in which equations (25) have no solution or infinitely many solutions.

The use of the axiom of inequality to arrive at this result does not preclude the possibility of certain special cases providing examples of equilibrium at stationary values of  $\bar{w}$ . Indeed, one such is known. It is a ready corollary of work by Moran (1967) that in the case where the fitnesses are multiplicative, there must exist one equilibrium for which  $\bar{w}$  is stationary. We may expect that similar cases may occasionally be found, but they must necessarily be extremely rare.

## VII. THE CONVERSE OF EWENS' THIRD RESULT

It is common to represent gamete frequencies as points in a tetrahedron (see Karlin and Feldman 1970). Consider the point  $P$  of this tetrahedron for which  $\bar{w}$  is maximized. This point is either (a) an interior point, (b) a point on a face of the tetrahedron, (c) a point on an edge of the tetrahedron, or (d) a corner point.

In case (a),  $P$  will not in general be an equilibrium point by the previous section and in case (b),  $P$  cannot be an equilibrium as equations (9) cannot hold. If a population structure is initially represented by  $P$  therefore, it will alter in such a way that the point representing it moves away from  $P$  towards a position of true equilibrium. This initial motion necessarily entails decrease of  $\bar{w}$ .

In case (c), the matter is more complicated. Edge equilibria entail the extinction of one of the alleles. Suppose therefore that  $\bar{w}$  is maximized on an edge. Then  $P$ , the point of maximization is necessarily an edge equilibrium at which one allele ( $a$ , say) is

absent. Let the equilibrium values of  $x_1, x_2$  be  $p, q$  respectively, and perturb the equilibrium by altering  $x_2$  to  $q-x$  and letting  $x_3$  become  $x$ , where  $x$  is small. Set

$$W_i = w_{i1}p + w_{i2}q, \quad (26)$$

$$W = W_1p + W_2q, \quad (27)$$

$$F = w_{23} - w_{22} - w_{13} + w_{12}. \quad (28)$$

Under these conventions, the equations of motion become

$$\left. \begin{aligned} \bar{W}\Delta x_1 &= xp(\bar{W} - W_3 - Fq) + Rqx \\ \bar{W}\Delta x_2 &= xq(\bar{W} - W_3 + Fp) - Rqx \\ \bar{W}\Delta x_3 &= x(W_3 - \bar{W}) - Rqx \\ \bar{W}\Delta x_4 &= Rqx \end{aligned} \right\} \quad (29)$$

The allele  $a$  will increase in frequency if  $\Delta(x_3 + x_4)$  is positive, i.e. if  $W_3 > \bar{W}$ . However, this condition is inconsistent with the assumption that  $\bar{W}$  was a maximum value of  $\bar{w}$ . Thus a stable equilibrium can maximize  $\bar{w}$  if and only if one allele becomes extinct.

Case (d) is a special case of case (c) in which two alleles become extinct. The analysis is the same.

Consider now cases in which  $\bar{w}$  has edge maxima only. If a stable equilibrium of equations (7) lies in the interior of the tetrahedron, then contour surfaces of  $\bar{w}$  may be drawn in the neighbourhood of the point  $Q$  representing this equilibrium. For all points in some neighbourhood of  $Q$  the trajectory representing the solutions of (7) will pass through  $Q$ . Since, by the previous section,  $Q$  is not a stationary point of  $\bar{w}$ ,  $\bar{w}$  must decrease along some of these trajectories.

Thus, we find that if  $\bar{w}$  has stationary points in the interior of the tetrahedron or if equations (7) have equilibria there, there necessarily exist conditions under which  $\bar{w}$  may decrease. However, if no polymorphic equilibria of equations (7) exist and if  $\bar{w}$  has no internal maxima, this conclusion need not follow. This exceptional case is one in which one allele or other necessarily becomes extinct.

## VIII. CONCLUSIONS

The three situations demonstrated by Moran (1964) have been shown by Ewens (1969*a*, 1969*b*) to occur if epistasis is absent. The present paper shows that:

- (1) If epistasis is present, gamete frequencies can never be expressed in terms of allelic frequencies only;
- (2) If epistasis is present,  $\bar{w}$  can achieve a polymorphic stationary value at equilibrium only in a special and biologically implausible case;
- (3) If epistasis is present, there are always population structures for which  $\bar{w}$  decreases under natural selection, except possibly in the case where one of the alleles becomes extinct.

These results are converses, in the mathematical sense, of the three results proved by Ewens (1969*a*). Result (1) alone suffices to show that the concept of an adaptive topography in the sense of Wright (1932) can never be applied to an

epistatic model. Ewens (1969a) has shown that this concept necessarily applied in the non-epistatic case. A similar dichotomy applies, with the exceptions noted above, to the maximization of mean fitness.

It is a common statement that recombination prevents the application of the fitness maximization and adaptive topography concepts. This is true in the sense that for a general model in which  $R \neq 0$ , neither concept can be applied. However, even if  $R = 0$  both concepts apply for non-epistatic cases. As soon, however, as epistasis is allowed, both concepts become inapplicable.

It thus follows that the failure of both the adaptive topography and fitness maximization concepts in the standard two-locus model is more properly described as being due to epistasis then to recombination. It would be of interest to know if this conclusion applies more generally.

#### IX. REFERENCES

- ARUNCHALAM, V. (1970).—Fundamental theorem of natural selection in two loci. *Ann. Hum. Genet.* **34**, 195–9.
- EWENS, W. J. (1969a).—Mean fitness increases when fitnesses are additive. *Nature, Lond.* **221**, 1076.
- EWENS, W. J. (1969b).—A generalised fundamental theorem of natural selection. *Genetics, Princeton* **63**, 531–7.
- FISCHER, R. A. (1918).—The correlation between relatives on the supposition of Mendelian inheritance. *Trans. R. Soc. Edin.* **52**, 399–433.
- HARDIN, D. (1960).—The competitive exclusion principle. *Science, N.Y.* **131**, 1291–7.
- KARLIN, S., and FELDMAN, M. (1970).—Linkage and selection: two locus symmetric viability model. *Theoret. Pop. Biol.* **1**, 39–71.
- KIMURA, M. (1966).—Two loci polymorphism as a stationary point. *Ann. Rep. Nat. Inst. Genet. Japan*, **17**, 65–7.
- MORAN, P. A. P. (1964).—On the nonexistence of adaptive topographies. *Ann. Hum. Genet.* **7**, 383–93.
- MORAN, P. A. P. (1967).—Unsolved problems in evolutionary theory. Proc. 5th Berkeley Symp. Math. Stat. Prob. Vol. 4. pp. 457–80. (University of California Press: Berkeley.)
- SCHEUER, P. A. G., and MANDEL, S. P. H. (1959).—An inequality in population genetics. *Heredity* **31**, 519–24.
- WRIGHT, S. (1932).—The roles of mutation, inbreeding and selection in evolution. Proc. 6th Int. Conf. Genet., Ithaca, N.Y. pp. 356–66.