ADAPTIVE TOPOGRAPHIES

Dr. Bob Gardner

Department of Mathematical and Statistical Sciences

(DoMaSS)

East Tennessee State University

presented in

Independent Study - Mathematical Biology

Summer 2003

These notes are based on Chapter 4 of *Principles of Population Genetics*,
2nd edition, by D. Hartl and A. Clark, Sinauer Associates, 1989.

Adaptive Topographies

Note. Sewall Wright (1889-1988) introduced a graphical way to visualize the Fundamental Theorem of Natural Selection. He wanted to plot $\overline{w}$ as a function of allele frequencies. This gives a surface "above" the allele space called an *adaptive topography*. The Fundamental Theorem of Natural Selection says that populations will be pulled uphill on these surfaces through the force of selection. In the event of a one locus-two alleles model, the possible topographies were given in a previous presentation and were graphs of functions $\overline{w} = \overline{w}(p)$ defined on the interval $p \in [0, 1]$. In the case of one locus-three alleles, we can plot the surface above a DeFinetti diagram. For more than 3 alleles at a locus, though, we require more than three dimensions to graph the surface (and hence we cannot easily visualize the result).

An Example with 3 Alleles

Example. We take as an example, human β-globin data from West Africa. There are 3 common β-globin alleles: β^A, β^C, and β^S. Homozygous $\beta^S\beta^S$ individuals have sickle cell anemia. Fitness values were put on the 6 possible gentoypes by A. C. Allison [1] by calculating the ratio of observed to expected numbers (since not all ratios are 1, the population is not in Hardy-Weinberg equilibrium):

geotype	fitness
AA	0.991
SS	0.218
CC	1.446
AS	1.104
AC	0.982
SC	0.788

(here, for example, we write "AA" to represent genotype $\beta^A\beta^A$). Some trajectories of populations on the adpative topography are:
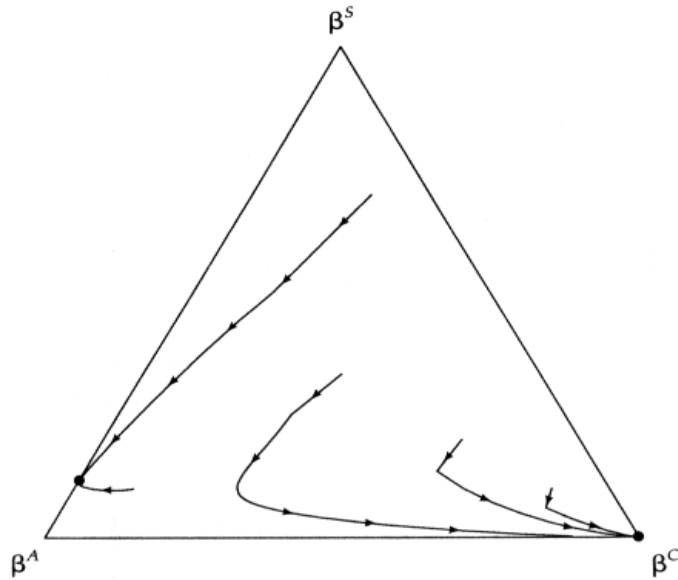
If a population starts with all A alleles (it is at the vertex of the triangle labelled β^A), and if the S allele is introduced (in small numbers), then the population will be drawn towards the polymorphic equilibirum where

$$p_s = \frac{w_{SS} - w_{AS}}{w_{AA} - 2w_{AS} + w_{SS}} \quad \text{and} \quad p_A = 1 - p_S.$$

For the fitness values above, this gives $p_S = 0.1209$, $p_A = 0.8791$, and $\overline{w} = 0.9033$. As can be seen in the figure, this point is "locally stable" and represents a local MAX of the surface.

Now suppose the C allele is introduced (in small quantity). Even though the population would have higher mean fitness if C could go to fixation, the dynamics of the situation will not allow it (the point $p_S = 0.1209$, $p_A = 0.8791$ is a stable equilibrium). However, if the C allele can be introduced at a sufficiently high frequency, then the

population will be drawn to fixation in C (the point labelled β^C in the diagram). This shows that the "fate" of a population is dependent on its initial position in allele space. In fact, we could divide up the allele space into a "basin of attraction" for the $\beta^A \beta^S$ polymorphic equilibrium and a basin of attraction for the point β^C of fixation. This shows that isolated populations might reach different average fitnesses for a given trait. Some populations (fortunate enough to find themselves in the right basin of attraction) will be attracted to the highest (globally MAX) possible fitness value (w_{CC}), whereas other populations might be "stuck" at the lower $\beta^A \beta^S$ equilibrium. A topographic map of the adaptive topography is:
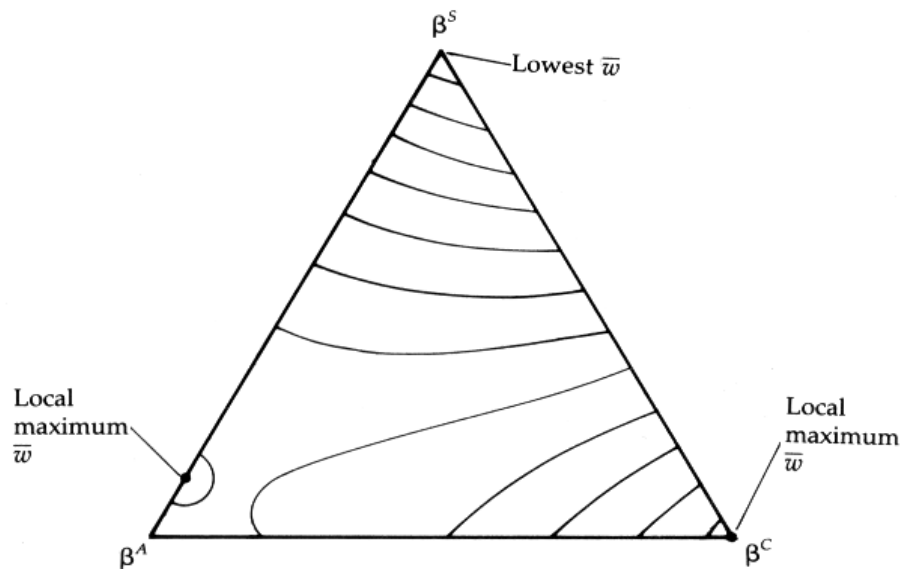
The Shifting Balance Theory

A type of selection called *interdeme selection* occurs between semi-isolated populations (demes) of the same species. If populations containing certain genotypes are more likely to become extinct and have their vacated habitats recolonized by migrants from other populations that are more persistent due to the particular genotypes that they contain, then the more successful populations can in some sense be considered as having a greater "fitness" than the less successful ones. Since this concept of population fitness is a characteristic of the entire population and not merely the average fitness of the genotypes within it ($\overline{w}$), interdeme selection is outside the realm of most conventional models of selection. Interdeme selection is one type of group selection [6].

Interdeme selection plays an essential role in the shifting balance theory of evolution (due also to Sewall Wright). In Wright's view, subdivision of a population into small, semi-isolated demes gives the best chance for the populations to explore the full range of their adaptive topography. Temporary reductions in fitness that would be prevented by selection in large populations become possible in small ones because of the random drift in allele frequencies that occurs in small populations. The lucky subpopulations that reach higher adaptive peaks on the fitness surface increase in size and send out more migrants than other subpopulations, and the favorable gene

combinations are gradually spread throughout the entire set of sub-populations by means of interdeme selection. The shifting balance process includes three distinct phases:

1. An *exploratory phase*, in which random genetic drift plays an important role in allowing small populations to explore their adaptive topography.

2. A phase of *mass selection*, in which favorable gene combinations createdby chance in phase 1 rapidly become incorporated into the genome of local populations by the action of natural selection.

3. A phase of *interdeme selection*, in which the more successful demes increase in size and rate of migration, and the excess migration shifts the allele frequencies of nearby populations until they also come under the control of the higher fitness peak. The favorable genotypes thereby become spread throughout the entire population in ever-widening concentric circles. Where the region of spread from two such centers overlaps, a new and still more favorable genotype may occur and itself become a center for interdeme selection. In this manner, the whole of the adaptive topography can be explored, and there is a continual shifting of control from one adaptive peak to a superior one.

The shifting balance theory has played an important role in evolutionary thinking, in part because of the prominent role assigned to random genetic drift in the initial phase of the process. How-

ever, as a comprehensive theory of evolution, many aspects of the theory remain to be tested. For the theory to work as envisaged, the interactions between alleles must often result in complex adaptive topographies within many peaks and valleys. The population must be split up into smaller demes, which must be small enough for random genetic drift to be improtant but large enough for mass selection to fix favorable combinations of alleles. While migration between demes is necessary, neighboring demes must be sufficiently isolated for genetic differentiation to occur, but sufficiently connected for favorable gene combinations to spread. Because of uncertainty about the applicability of these assumptions, the shifting balance process remains a picturesque metaphor that is still largley untested.

Two or More Loci, Linkage, and
Some Problems for the Fundemantal Theorem

Note. If we consider two loci, each with two possible alleles (say A, a at one locus and B, b at the other), then we can express the possible "states" of a population by plotting a point $(p_A, p_B) \in [0,1] \times [0,1]$. This then allows us to visualize the graph of $\overline{w}$ as a surface over this unit square. We have 9 possible genotypes, and hence:

genotype	fitness
$AABB$	w_1
$AABb$	w_2
$AAbb$	w_3
$AaBB$	w_4
$AaBb$	w_5
$Aabb$	w_6
$aaBB$	w_7
$aaBb$	w_8
$aabb$	w_9

Now <u>if we assume</u> that the A/a and B/b alleles are inherited <u>independently</u>, then we have:

genotype	frequency	fitness
$AABB$	$p_A^2 p_B^2$	w_1
$AABb$	$2p_A^2 p_B q_b$	w_2
$AAbb$	$p_A^2 q_b^2$	w_3
$AaBB$	$2p_A q_a p_B^2$	w_4
$AaBb$	$4p_A q_a p_B q_b$	w_5
$Aabb$	$2p_A q_a q_b^2$	w_6
$aaBB$	$q_a^2 p_B^2$	w_7
$aaBb$	$2q_a^2 p_B q_b$	w_8
$aabb$	$q_a^2 q_b^2$	w_9

where p_A is the frequency of allele A, $q_a = 1 - p_A$, and p_B is the frequency of allele B, $q_b = 1 - p_B$. Under these assumptions, we may calculate average fitness as:

$$\overline{w} = (p_A^2 p_B^2)w_1 + (2p_A^2 p_B q_b)w_2 + (p_A^2 q_b^2)w_3 + (2p_A q_a p_B^2)w_4+$$

$$(4p_A q_a p_B q_b)w_5 + (2p_A q_a q_b^2)w_6 + (q_a^2 p_B^2)w_7 + (2q_a^2 p_B q_b)w_8 + (q_a^2 q_b^2)w_9.$$
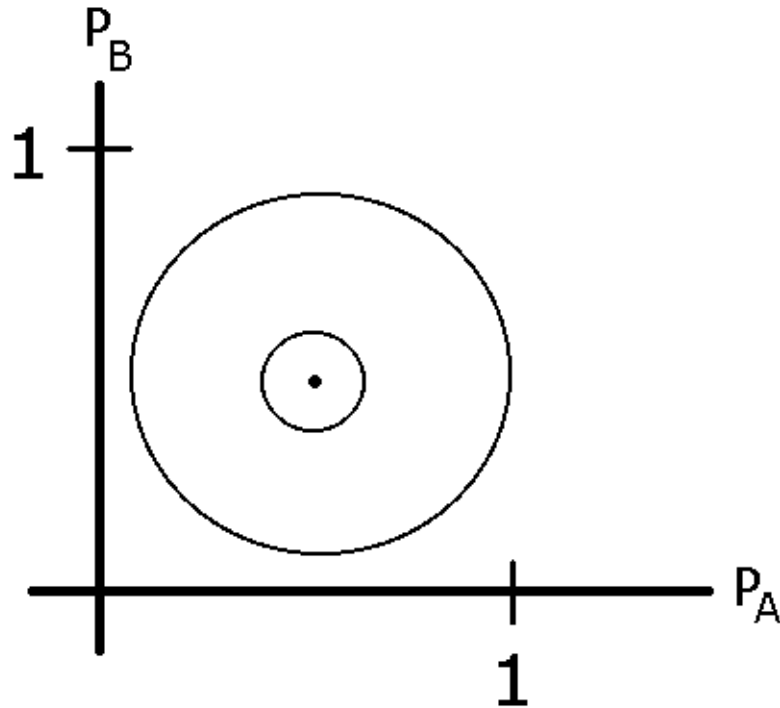
We can substitute $q_a = 1 - p_A$ and $q_b = 1 - p_B$ to get $\overline{w}$ in terms of p_A and p_B only.

Example. It is common for heterozygotes to be more fit than homozygotes. As such, consider the following fitness values:

genotype	fitness
$AABB$	1
$AABb$	4
$AAbb$	1
$AaBB$	4
$AaBb$	5
$Aabb$	4
$aaBB$	1
$aaBb$	4
$aabb$	1

The associated adaptive topography has a single interior maximum.

Here are some level curves:

Note. We are interested in maintaining polymorphism through se-
lection. (This was a rather large debate in the history of population
genetics. When molecular techniques were first introduced in the
1960's, a great deal of diversity was found to be present in most
every population. The question became: "Is this diversity due to
selection, or due to the accumulation of neutral mutations?" This is
the heart of the "selection/neutrality" debate.) Therefore, we would
like to try to find adaptive topographies that are very . . . "lumpy."

That is, we want a surface with many local MAXs.

Example. Consider the following fitness values:

genotype	fitness
$AABB$	3
$AABb$	4
$AAbb$	1
$AaBB$	4
$AaBb$	3
$Aabb$	4
$aaBB$	1
$aaBb$	4
$aabb$	3

The associated adaptive topography has two interior maxima (at approximately $(p_A, p_B) = (0.21, 0.21)$ and $(p_A, p_B) = (0.79, 0.79)$) and a saddle point (at $(p_A, p_B) = 0.5, 0.5$)). Here are some level

curves:
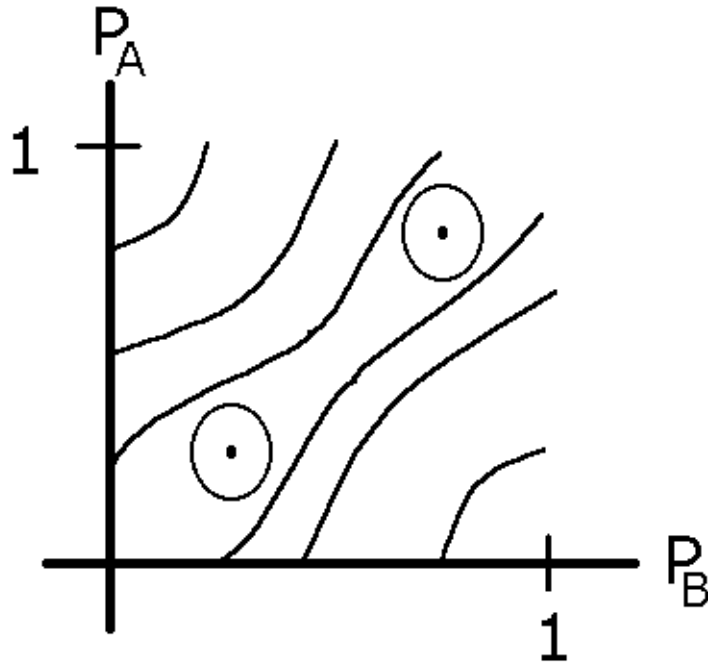
Note. However, this assumption that A/a and B/b are inherited independently may be questionable and, in particular, is biologically unrealistic. Hence, we need to discuss *linkage disequilibrium*. To do so, we need to study the genes making up gametes (sex cells).

Note. Let's consider two pairs of genes (autosomal — not located on a sex chromosome), say A/a and B/b. As above, let p_A, q_a, p_B, and q_b denote the frequencies of A, a, B, and b, respectively. In such a population, a random individual will form gametes (which

are monoploid — throughout we are discussing diploid organisms) as follows (assuming A/a and B/b are inherited independently):

gamete	frequency
AB	$p_A p_B$
Ab	$p_A q_b$
aB	$q_a p_B$
ab	$q_a q_b$

If these are in fact the frequencies of gametes, then we have "random association in the gametes" and the population is in *linkage equilibrium* for these genes.

Note. A population that is not in linkage equilibrium is said to be in *linkage disequilibrium*. We can denote gametic frequencies in general as:

gamete	frequency
AB	P_{11}
Ab	P_{12}
aB	P_{21}
ab	P_{22}

<u>If</u> the population is in linkage equilibrium, then

$$P_{11} = p_A p_B$$
$$P_{12} = p_A q_b$$
$$P_{21} = q_a p_B$$
$$P_{22} = q_a q_b$$

If the population is not in linkage equilibrium, then one or more of these equations will be violated. Notice, also, that $P_{11} + P_{12} + P_{21} + P_{22} = 1$.

Note. Similar to a DeFinetti diagram, we select a point in the interior of a tetrahedron and the sum of the four perpendicular distances to the faces of the tetrahedron is a constant. Therefore, we could represent the "gametic state" $(P_{11}, P_{12}, P_{21}, P_{22})$ of a population by plotting a point in a tetrahedron. We have seen that a population in Hardy-Weinberg equilibrium has points in a DeFinetti diagram which lie on a parabola. Similarly, the points in a tetrahedron $(P_{11}, P_{12}, P_{21}, P_{22})$ which correspond to a population in linkage

equilibrium determines a surface (called the *Wright manifold*):
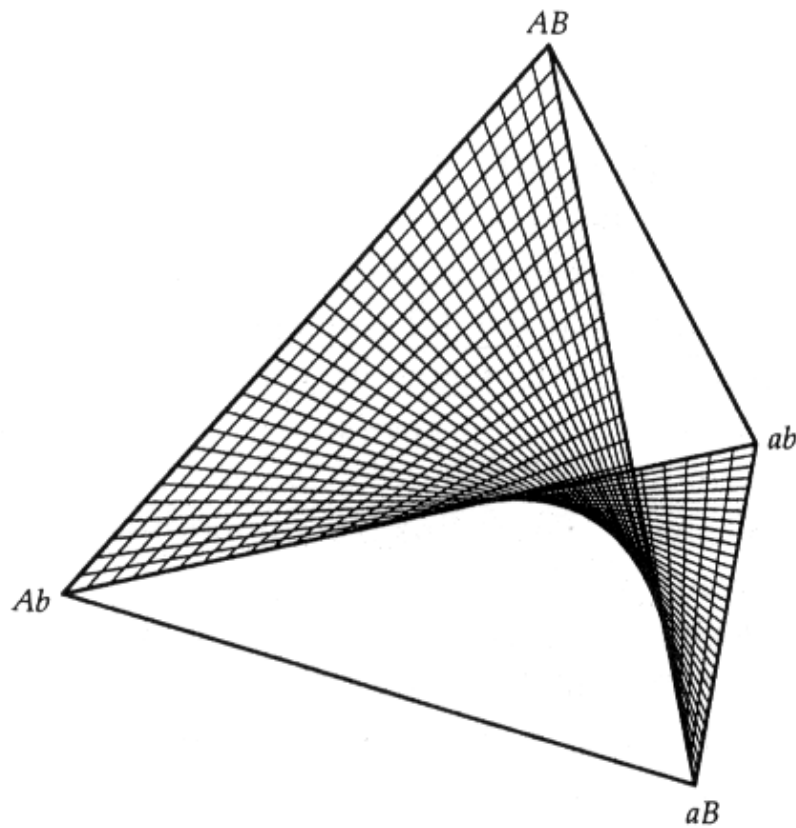
Points off of this manifold represent a situation of linkage disequilibrium.

Note. Unfortunately, linkage disequilibrium can have severe consequences for the dynamics of a population. In fact, due to the linkage disequilibrium, a population may be drawn to an equilibrium which does not yield a maximum of fitness (in apparent violation of the Fundamental Theorem of Natural Selection). A. Hastings [2,3] produced a set of parameters (for fitness and linkage disequilibrium) which produced an adaptive topography with four stable polymor-

phic equilibria. However, the equilibria do not correspond to extrema of fitness. The surface is:
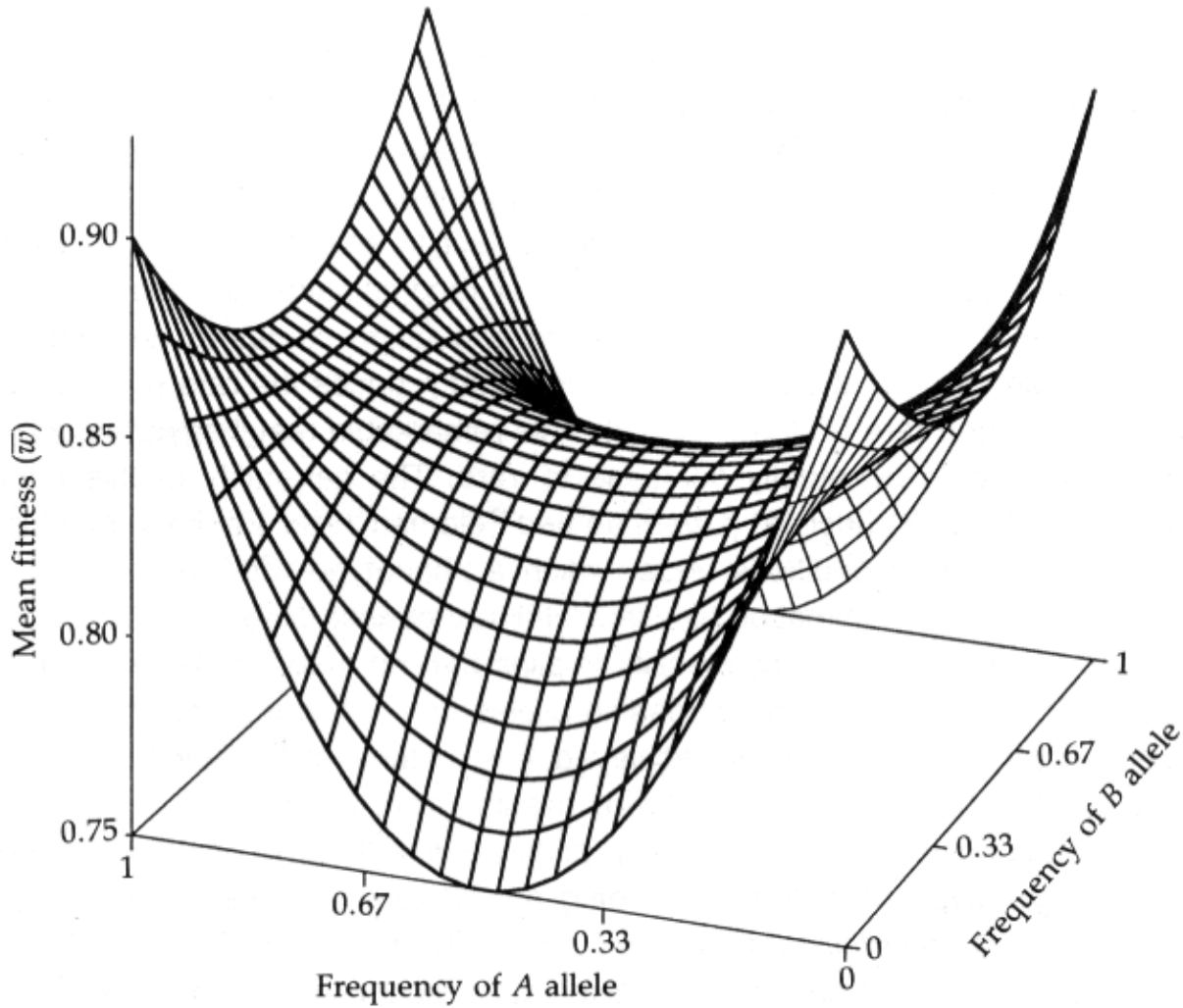
Note. The above topography with equilibria which do not correspond to extrema of fitness, shows us that we must be careful in applying the Fundamental Theorem of Natural Selection. Controversy over this theorem dates back many years (Wright and Fisher were frequently in conflict, in particular over Fisher's objections to

Wright's assumption of random mating in his shifting balance theory). W. J. Ewens has studied the Fundamental Theorem is some details [4,5]. The abstract of one of his works [4] reads:

Fisher's "Fundamental Theorem of Natural Selection" has long caused controversy in population genetics theory. Viewed as a statement about the increase, or rate of increase, of mean fitness over time, it encountes difficulties with cases arising in a multi-locus system for which mean fitness can decrease. An interpretation of the theorem is put forward here which implies that it is correct as a mathematical statement, but of less biological value than was claimed by Fisher.

References

1. A. C. Allison, Polymorphism and Natural Selection in Human Populations, *Cold Harbor Symposium on Quantitative Biology*, **29**, (1964) 137–149.

2. A. Hastings, Linkage Disequilibrium, Selection and Recombination of Three Loci, *Genetics* **106** (1984), 153-164.

3. A. Hastings, Four Simultaneously Stable Polymorphic Equilibria in Two-Locus Two-Allele Model, *Genetics* **109** (1985), 255–261.

4. W. J. Ewens, An Interpretation and Proof of the Fundamental Theorem of Natural Selection, *Theoretical Population Biology* **36** (1989), 167–180.

5. W. J. Ewens, An Optimizing Principle of Natural Selection in Evolutionary Population Genetics, *Theoretical Population Biology* **42** (1992), 333–346.

6. D. S. Wilson, The Group Selection Controversy: History and Current Status, *Annual Reviews of Ecological Systems* **14** (1983), 159–188.