

SLOW SELECTION ANALYSIS OF GENETIC TRAITS IN SYNCHRONIZED POPULATIONS

F. C. HOPPENSTEADT

The genetic structure defined by two autosomal loci (A, B) each having two alleles (A, a , and B, b , respectively) in a diploid population is described by ten genotypes. A simplified model of such traits results when we consider the population to be synchronized in the sense that all reproductions occur at once followed by the removal of parents. In this case, it is necessary to consider only the gamete types that participate in reproduction. There are four of these, and our notation for their frequencies after the n th reproduction time are given in the following table:

Gamete Type	AB	Ab	aB	ab
Frequency	$g_1(n)$	$g_2(n)$	$g_3(n)$	$g_4(n)$

Writing

$$\vec{g}(n) = \begin{bmatrix} g_1(n) \\ g_2(n) \\ g_3(n) \\ g_4(n) \end{bmatrix},$$

and assuming that the population is large and randomly mating, we can determine $\vec{g}(n + 1)$ in terms of $\vec{g}(n)$; namely

$$\vec{g}(n + 1) = \frac{\vec{g}(n) \circ // \vec{g}(n) - r\vec{k}(\vec{g}(n), \Lambda\vec{g}(n))}{(\vec{g}(n), // \vec{g}(n))}$$

where the following notation is used:

$//$ is a symmetric 4×4 -matrix of fitnesses of the genotypes, $//_{ij}$ gives the fitness of a $g_i g_j$ individual, (we assume $//_{23} = //_{14}$), r is the recombination rate (a fixed constant, $0 \leq r \leq 1/2$)

$$\vec{k} = \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix} \text{ and } \Lambda = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix},$$

Received by the editors on December 1, 1977.

and for any two vectors \vec{x} , \vec{y} ,

$$(\vec{x}, \vec{y}) = \sum_{i=1}^4 x_i y_i \text{ (Inner product)}$$

and

$$\vec{x} \circ \vec{y} = \begin{pmatrix} x_1 y_1 \\ x_2 y_2 \\ x_3 y_3 \\ x_4 y_4 \end{pmatrix} \text{ (Schur product).}$$

Note that $g_1(n) + g_2(n) + g_3(n) + g_4(n) = 1$.

A great deal of work has gone into determining the behavior of the gamete frequencies $\vec{g}(n)$ for increasing n under various conditions on the fitness matrix \mathscr{W} and recombination rate r . Work continues on the problem.

Slow Selection Analysis. When selection acts over a much longer time scale than reproduction, the model reduces to a form which can be analyzed without great effort. In the slow selection case, \mathscr{W} takes the form

$$\mathscr{W} = \mathscr{W} + \epsilon W \text{ (Slow Selection Assumption)}$$

where \mathscr{W} is a matrix having all components equal 1 ($\mathscr{W}_{ij} = 1$ for all i, j), ϵ is a small parameter ($0 < \epsilon \ll 1$) measuring selective intensity, and W is a matrix of relative fitnesses ($W = 0(1)$).

Using this form of \mathscr{W} and the facts that $\vec{g} \circ \mathscr{W} \vec{g} = \vec{g}$ and $(\vec{g}, \mathscr{W} \vec{g}) = 1$, we have that

$$\begin{aligned} \vec{g}(n+1) &= \vec{g}(n) - r\vec{k}(\vec{g}(n), \Lambda \vec{g}(n)) \\ &\quad + \epsilon[\vec{g}(n) \circ W \vec{g}(n) - (\vec{g}(n), W \vec{g}(n)) \\ &\quad (\vec{g}(n) - r\vec{k}(\vec{g}(n), \Lambda \vec{g}(n)))] + 0(\epsilon^2). \end{aligned}$$

The solutions of this equation can be analyzed by a multi-time method developed for difference equations in [1]. This will be applied in two cases. First, the case where there is little recombination of chromosomes (tight linkage) where $0 \leq r \ll \epsilon \ll 1$, and second, the case of loose linkage where $0 < \epsilon \ll 1$, $0 < r \leq 1/2$.

1. Slow selection ($\epsilon \ll 1$) and tight linkage ($0 \leq r \ll \epsilon$). The gamete frequencies will be found in the form

$$\vec{g}(n) = \vec{u}(n, s, \epsilon) = \vec{u}_0(n, s) + \epsilon \vec{u}_1(n, s) + \dots$$

where $s = \epsilon n$ is a slow time variable, the one on which selection acts. The functions $\bar{u}_0, \bar{u}_1, \dots$, are assumed to be smooth functions of the variable s . Substituting this form into the equation for gamete frequencies gives

$$\begin{aligned} \bar{u}(n+1, s+\epsilon, \epsilon) \\ = \bar{u}(n, s, \epsilon) + \epsilon[\bar{u} \circ W\bar{u} - (\bar{u}, W\bar{u})\bar{u}] + O(\epsilon). \end{aligned}$$

The Taylor expansion of \bar{u} can be determined by successively differentiating this equation and setting $\epsilon = 0$ in the results:

$$\begin{aligned} \bar{u}_0(n+1, s) &= \bar{u}_0(n, s), \\ \bar{u}_1(n+1, s) &= \bar{u}_1(n, s) + [\bar{u}_0 \circ W\bar{u}_0 - (\bar{u}_0, W\bar{u}_0)\bar{u}_0] \frac{\partial \bar{u}_0}{\partial s}, \\ &\text{etc.} \end{aligned}$$

The first equation is easily solved: It states that \bar{u}_0 does not depend on n , and so

$$\bar{u}_0(n, s) = \bar{U}_0(s)$$

where \bar{U}_0 must be determined. The second equation becomes

$$\begin{aligned} \bar{u}_1(n+1, s) &= \bar{u}_1(n, s) \\ * \quad &+ [\bar{U}_0 \circ W\bar{U}_0 - (\bar{U}_0, W\bar{U}_0)\bar{U}_0] - \frac{d\bar{U}_0}{ds}. \end{aligned}$$

This can be solved for \bar{u}_1 as

$$\begin{aligned} \bar{u}_1(n, s) &= \bar{U}_1(s) \\ &+ n \left\{ \bar{U}_0 \circ W\bar{U}_0 - (\bar{U}_0, W\bar{U}_0)\bar{U}_0 - \frac{d\bar{U}_0}{ds} \right\}. \end{aligned}$$

In order for the expansion of \bar{u} to be useful, we should specify that \bar{u}_1 be bounded as a function of n . Using this, dividing both sides of the equation by n and passing to the limit $n = \infty$, we get

$$\frac{d\bar{U}_0}{ds} = \bar{U}_0 \circ W\bar{U}_0 - (\bar{U}_0, W\bar{U}_0)\bar{U}_0, \quad \bar{U}_0(0) = \bar{g}(0).$$

The function $V(\bar{x}) = (\bar{x}, W\bar{x})$ is a Liapunov function for this system. In fact,

$$\begin{aligned} (1/2) \frac{dV(\bar{U}_0)}{ds} &= (W\bar{U}_0, d\bar{U}_0/ds) \\ &= (W\bar{U}_0, \bar{U}_0 \circ W\bar{U}_0) - (W\bar{U}_0, \bar{U}_0)(\bar{U}_0, W\bar{U}_0) \end{aligned}$$

since $(\vec{U}_0, W\vec{U}_0) = (\vec{N}, \vec{U}_0 \circ W\vec{U}_0)$ where

$$\vec{N} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$

But this expression can be rewritten as

$$(W\vec{U}_0 - (WU_0, U_0)\vec{N}, U_0 \circ [WU_0 - (WU_0, U_0)\vec{N}]) \geq 0$$

since $(\vec{x}, \vec{U}_0 \circ \vec{x}) = \sum_{i=1}^4 U_{0,i} x_i^2 \geq 0$ for any vector \vec{x} .

A further calculation shows that

$$\frac{dV(\vec{U}_0)}{ds} > 0$$

except at rest points (where $dU_0/ds = 0$). This is carried out in [2]. The function V is called the fitness of the population and this calculation verifies Fisher's Fundamental Theorem of Natural Selection for two tightly linked loci, which states that a population's genetic structure changes in such a way as to increase its fitness.

2. The effects of slow selection ($0 < \epsilon \ll 1$) and loose linkage ($0 < r \leq 1/2$), on a population initially at linkage equilibrium ($(\Lambda g_0, g_0) = 0(\epsilon)$).

The function

$$L = (\vec{g}, \Lambda \vec{g})$$

is called the *linkage disequilibrium function*. When $L \equiv 0$, the gamete frequencies are the same as when the two loci act independently. This function satisfies the equation

$$\begin{aligned} L(n+1) &= (1-2r)L(n) + 2\epsilon[(r-1)(\vec{u}, W\vec{u})]L(n) \\ &\quad + 2\epsilon(\Lambda \vec{u}, \vec{u} \circ W\vec{u}). \end{aligned}$$

It follows that $L(0, 0, \epsilon) = 0(\epsilon)$ implies $L = 0(\epsilon)$ for all $n = 0(1/\epsilon)$. Therefore, we write

$$L(n) = 2\epsilon D(n),$$

and then D satisfies

$$\begin{aligned} D(n+1, s+\epsilon, \epsilon) &= (1-2r)D(n, s, \epsilon) \\ &\quad + (\Lambda \vec{u}, \vec{u} \circ W\vec{u}) + 0(\epsilon). \end{aligned}$$

It follows that

$$D(n, s, \epsilon) = (\Lambda \vec{u}_0, \vec{u}_0 \circ W \vec{u}_0) / 2r + 0(\epsilon) + 0[(1 - 2r)^n].$$

Returning to the original problem, we have

$$\begin{aligned} \vec{u}(n+1, s) &= \vec{u}(n, s) + \epsilon [\vec{u}(n, s) \circ W \vec{u}(n, s) \\ &\quad - (\vec{u}(n, s), W \vec{u}(n, s)) \vec{u}(n, s) - 2rk D_n] + 0(\epsilon^2) \end{aligned}$$

Using the same method as before, we now have

$$\frac{d\vec{U}_0}{ds} = \vec{U}_0 \circ W \vec{U}_0 - (\vec{U}_0, W \vec{U}_0) \vec{U}_0 - k(\Lambda \vec{U}_0, \vec{U}_0 \circ W \vec{U}_0).$$

$V(U_0)$ again can be shown to be increasing along nonstatic solutions. Now, however, we can only verify the Fundamental Theorem after an initial transient period in which the gamete frequencies equilibrate to $\vec{u}_0 + 0(\epsilon)$.

REMARK. The same calculations carry over to some cases where W varies with n and \vec{g}_n . For example, in the tight linkage case equation (*) must be replaced by

$$\begin{aligned} \vec{u}_1(n, s) &= \vec{U}_1(s) \\ &+ \sum_{k=0}^{n-1} [\vec{U}_0 \circ W(k, \vec{U}_0) \vec{U}_0 - (\vec{U}_0, W(k, \vec{U}_0) \vec{U}_0)] - n \frac{d\vec{U}_0}{ds}, \end{aligned}$$

and the equation for \vec{U}_0 becomes

$$\frac{d\vec{U}_0}{ds} = \vec{U}_0 \circ \bar{W}(\vec{U}_0) \vec{U}_0 - (\vec{U}_0, \bar{W}(\vec{U}_0) \vec{U}_0)$$

where

$$\bar{W}(\vec{U}_0) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{l=0}^{n-1} W(l, \vec{U}_0).$$

Additional hypotheses are needed to ensure that this limit exists and defines a smooth function. For example, if W is a periodic or almost periodic function of n , then \bar{W} exists.

BIBLIOGRAPHY

1. F. C. Hoppensteadt and W. L. Miranker, *Multi-time methods for systems of difference equations*, Studies in Applied Math. **56** (1977) 273-289.
2. F. C. Hoppensteadt, *A slow selection analysis of two locus, two allele traits*, Theor. Pop. Biol. **9** (1976), 68-81.

MATHEMATICS DEPARTMENT, UNIVERSITY OF UTAH, SALT LAKE CITY, UTAH 84112

