Selection, Generalized Transmission and the Evolution of Modifier Genes. I. The Reduction Principle

Lee Altenberg* and Marcus W. Feldman†

**Department of Statistics, North Carolina State University, Raleigh, North Carolina 27695-8203 and †Department of Biological Sciences, Stanford University, Stanford, California 94305-5020*

Manuscript received March 19, 1987
Revised copy accepted July 16, 1987

ABSTRACT

Modifier gene models are used to explore the evolution of features of organisms, such as the genetic system, that are not directly involved in the determination of fitness. Recent work has shown that a general "reduction principle" holds in models of selectively neutral modifiers of recombination, mutation, and migration. Here we present a framework for models of modifier genes that shows these reduction results to be part of a more general theory, for which recombination and mutation are special cases.—The deterministic forces that affect the genetic composition of a population can be partitioned into two categories: selection and transmission. Selection includes differential viabilities, fertilities, and mating success. *Imperfect* transmission occurs as a result of such phenomena as recombination, mutation and migration, meiosis, gene conversion, and meiotic drive. Selectively neutral modifier genes affect transmission, and a neutral modifier gene can evolve only by generating association with selected genes whose transmission it affects.—We show that, in randomly mating populations at equilibrium, imperfect transmission of selected genes allows a variance in their marginal fitnesses to be maintained. This variance in the marginal fitnesses of selected genes is what drives the evolution of neutral modifier genes. Populations with a variance in marginal fitnesses at equilibrium are always subject to invasion by modifier genes that bring about perfect transmission of the selected genes. It is also found, within certain constraints, that for modifier genes producing what we call "linear variation" in the transmission processes, a new modifier allele can invade a population at equilibrium if it reduces the level of imperfect transmission acting on the selected genes, and will be expelled if it increases the level of imperfect transmission. Moreover, the strength of the induced selection on the modifier gene is shown to range up to the order of the departure of the genetic system from perfect transmission.

MODELS of modifier genes have been developed over the last 20 yr to extend the synthetic theory of evolution to features of organisms which fall outside of the classical purview of heritable variation for fitness. The biological phenomena encompassed in modifier gene models are diverse, and include features of the genetic system such as recombination and mutation, features of the mating system such as selfing, and ecological features such as migration rates.

A principle for the evolution of neutral modifier genes, the "reduction principle," has been developed by FELDMAN, CHRISTIANSEN and BROOKS (1980), LIBERMAN and FELDMAN (1986a,b; 1987) and FELDMAN and LIBERMAN (1986). In models of recombination, mutation and migration modification in which the modifying gene is multiallelic, LIBERMAN and FELDMAN have shown that there is a reduction principle that holds under a wide variety of conditions, and they suggest that complete reduction of each of these has the property of "evolutionary genetic stability" (ESHEL and FELDMAN 1982). The possibility that these parallel results reflect an underlying unity among modifier gene models is the motivation for the current work.

We introduce here a theoretical approach to the evolution of modifier genes that partitions the forces of evolution, other than drift, into two classes, selection and transmission. Transmission is the determination of offspring types by parental types. Imperfect transmission occurs whenever the gametes in offspring are different from those in their parents. Selection occurs when different types in the population differ in their relative contributions to the next generation. A large class of modifier gene models can be unified once they are seen to involve genetic variation for transmission processes. In contrast to genes that produce differences in fitness, neutral modifier genes produce differences in transmission processes. Modifiers of recombination, mutation, segregation distortion, and migration are in this class.

We investigate the initial increase properties of modifier alleles introduced into randomly mating diploid populations near genetic equilibrium for a set of

Dedicated to Luca Cavalli-Sforza on his 65th birthday.

polymorphic loci undergoing viability selection and generalized transmission. The main results can be summarized as follows. First, selection for or against the modifier requires that there be a variance in the equilibrium marginal fitnesses of the selected loci. This in turn requires that the selected loci be imperfectly transmitted. We obtain two results that generalize the reduction principle. First, we show that a population that maintains an equilibrium variance in the marginal fitnesses, due to imperfect transmission of loci under selection, can always be invaded by a new modifier allele that produces perfect transmission. Second, we show that when the modifier and major loci are tightly linked, a new modifier allele that produces "linear" variation in the transmission process can invade if and only if it reduces the level of imperfect transmission. We also obtain results on the possible strength of selection induced on a tightly linked neutral modifier, and show that it is on the order of the variation in equilibrium marginal fitnesses, which is bounded by the extent of imperfect transmission characteristic of the genetic system.

GENERALIZED TRANSMISSION

The term haplotype will be used to mean a gamete's genotype, or a gamete's contribution to a diploid genotype. *Perfect transmission* occurs when a diploid individual's two haplotypes, as originally derived from its parents, are transmitted unchanged to its gametes, and in equal proportions. For a multilocus parental haplotype to be perfectly transmitted, the whole haplotype must behave as a Mendelian unit. *Imperfect transmission* occurs when gamete haplotypes differ from the parental haplotypes as a consequence of mutation, recombination, gene conversion, translocation, etc.

To represent the transmission relations between parental and gamete haplotypes, we first enumerate all of the possible haplotypes, $\{i, j, k, \cdots\}$. The transmission can be represented by the set of probabilities $T(i \leftarrow j, k)$ that a gamete produced by a parent with haplotypes j and k has haplotype i;

$$T(i \leftarrow j, k) \geq 0, \quad \sum_i T(i \leftarrow j, k) = 1 \quad \text{for all } j, k. \tag{1}$$

For simplicity we restrict our analysis here to autosomal systems so that there is symmetry in T: $T(i \leftarrow j, k) = T(i \leftarrow k, j)$. If there are n possible haplotypes, then the genetic transmission system can be represented as an n by n^2 matrix:

$$\mathbf{T} = \|T(i \leftarrow j_1, j_2)\|^n_{i, j_1 j_2 = 1}. \tag{2}$$

This matrix is just the "segregation table" for all of the genotypes in the population. Matrices satisfying (1) are called biparental *transmission matrices.*

Varieties of transmission processes: Models of one locus without mutation or segregation distortion are examples of perfect transmission. The transmission probabilities with perfect transmission are

$$T_{id}(i \leftarrow j, k) = \tfrac{1}{2}(\delta_{ij} + \delta_{ik}),$$

where δ_{ij} is the Kronecker delta (δ) function,

$$\delta_{ij} = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}.$$

Different genetic systems may produce different levels of imperfect transmission during reproduction. A quantity that plays a central role in our analysis is the level of perfect transmission that is guaranteed by the genetic system for all possible parental genotypes. This upper bound (over all genotypes) on the probability of imperfect transmission we refer to as α. Any transmission matrix can be written in the form,

$$\mathbf{T} = (1 - \alpha)T_{id} + \alpha\mathbf{P}, \tag{3}$$

which involves a mixture of the perfect transmission matrix, $\mathbf{T}_{id}$ and a transmission matrix, $\mathbf{P}$. The value α is the minimum value for which the elements of the matrix

$$\mathbf{T} - (1 - \alpha)T_{id} \geq 0 \tag{4}$$

are non-negative, so that $1 - \alpha$ is a measure of the degree of perfect transmission present in T.

Tables 1, 2 and 3 show the transition probabilities for several standard models of mutation, segregation distortion, and recombination, and an example of a model of gene conversion at two loci. The value of α for each of these models is given.

EVOLUTION OF HAPLOTYPES UNDER SELECTION AND GENERALIZED TRANSMISSION

Here we formulate a general model of evolution for a diploid population undergoing random mating, viability selection, and transmission of haplotypes from zygote to gamete. Let $\mathbf{z}$ be the vector representing the frequencies of all gametic haplotypes in the population. Let

$$\mathbf{W} = \|w_{ij}\|^n_{i,j=1}$$

where w_{ij} is the viability of a genotype composed of haplotypes i and j. We assume sex symmetry so that $w_{i_1 i_2} = w_{i_2 i_1}$.

With a life cycle consisting of random union of gametes, viability selection, and generalized transmission, the recursion on the frequencies of gamete haplotypes is

$$\bar{w}z_i' = \sum_{j=1}^{n} \sum_{k=1}^{n} z_j z_k T(i \leftarrow j, k) w_{jk} \quad (i = 1, 2, \cdots, n) \tag{5a}$$

TABLE 1

Transmission table for mutation and segregation distortion

Parental genotypes	Frequencies of gamete haplotypes	
	A_1	A_2
A_1/A_1	$1 - m_1$	m_1
A_1/A_2	$\frac{1}{2}(1 - m_1 + m_2) + k/2$	$\frac{1}{2}(1 + m_1 - m_2) - k/2$
A_2/A_2	m_2	$1 - m_2$

With $m_1 = m_2 = k = 0$ we have perfect transmission.
With $k = 0$, $\alpha = \max(m_1, m_2)$ [see (3)].
With $m_1 = m_2 = 0$, $\alpha = |k|$ [see (3)].

TABLE 2

Transmission table for recombination

Parental genotypes	Frequencies of gamete haplotypes			
	A_1B_1	A_1B_2	A_2B_1	A_2B_2
A_1B_1/A_1B_1	1	0	0	0
A_1B_1/A_1B_2	$\frac{1}{2}$	$\frac{1}{2}$	0	0
A_1B_1/A_2B_1	$\frac{1}{2}$	0	$\frac{1}{2}$	0
A_1B_1/A_2B_2	$\frac{1}{2}(1 - R)$	$\frac{1}{2}R$	$\frac{1}{2}R$	$\frac{1}{2}(1 - R)$
A_1B_2/A_1B_2	0	1	0	0
A_1B_2/A_2B_1	$\frac{1}{2}R$	$\frac{1}{2}(1 - R)$	$\frac{1}{2}(1 - R)$	$\frac{1}{2}R$
A_1B_2/A_2B_2	0	$\frac{1}{2}$	0	$\frac{1}{2}$
A_2B_1/A_2B_1	0	0	1	0
A_2B_1/A_2B_2	0	0	$\frac{1}{2}$	$\frac{1}{2}$
A_2B_2/A_2B_2	0	0	0	1

Here $\alpha = R$.

or in matrix form

$$\mathbf{z}' = \frac{1}{\bar{w}} \| \sum_k z_k w_{jk} T(i \leftarrow j, k) \|_{i,j=1}^{n} \mathbf{z}, \qquad (5b)$$

where

$$\bar{w} = \sum_{ij} w_{ij} z_i z_j = \mathbf{z}^T \mathbf{W} \mathbf{z} \qquad (6)$$

is the mean fitness of the population.

A central quantity involved in our analysis is the *marginal fitness* of a haplotype. The marginal fitness of haplotype i is $w_i = \Sigma_j w_{ij} z_j$. It represents the net effect of selection on a haplotype averaged over all the genotypes in which it appears. With perfect transmission, variation in marginal fitnesses is the sole force acting to change haplotype frequencies, in which case (5) becomes

$$\bar{w} z_i' = z_i \sum_j w_{ij} z_j, \quad \text{or,} \quad \bar{w} \mathbf{z}' = \text{diag}(\mathbf{W}\mathbf{z})\mathbf{z} \qquad (7)$$

(see EWENS 1979, pp. 40–41). The relative frequency of haplotype i changes in one generation by an amount proportional to $\Sigma_j w_{ij} z_j - \bar{w}$.

Employing the marginal fitnesses, the general recursion (5) can be written in the form

$$\mathbf{z}' = \mathbf{Y}\mathbf{D}\mathbf{z}, \qquad (8)$$

where $\mathbf{Y}$ is a column stochastic matrix, and $\mathbf{D}$ is a positive diagonal matrix, as follows:

$$\mathbf{Y} = \| \sum_k z_k \frac{w_{jk}}{w_j} T(i \leftarrow j, k) \|_{i,j=1}^{n},$$

and

$$\mathbf{D} = \text{diag}\left(\frac{w_i}{\bar{w}}\right) = \begin{vmatrix} \frac{w_1}{\bar{w}} & & 0 \\ & \ddots & \\ 0 & & \frac{w_n}{\bar{w}} \end{vmatrix}. \qquad (9)$$

$\mathbf{Y}$ is column stochastic since

$$\sum_{ik} z_k \frac{w_{jk}}{w_j} T(i \leftarrow j, k) = \sum_k \frac{z_k w_{jk}}{w_j} = 1.$$

EQUILIBRIUM PROPERTIES OF POPULATIONS UNDER SELECTION AND IMPERFECT TRANSMISSION

In this treatment we are concerned with the evolution of modifier genes introduced into populations near equilibrium. Here we derive the most salient

TABLE 3

Transmission table for unbiased gene conversion at two loci

Parental genotypes	Frequencies of gamete haplotypes			
	A_1B_1	A_1B_2	A_2B_1	A_2B_2
A_1B_1/A_1B_1	1	0	0	0
A_1B_1/A_1B_2	$\frac{1}{2}$	$\frac{1}{2}$	0	0
A_1B_1/A_2B_1	$\frac{1}{2}$	0	$\frac{1}{2}$	0
A_1B_1/A_2B_2	$\frac{1}{2}(1 - a - b + 2ab)$	$\frac{1}{2}(a + b - 2ab)$	$\frac{1}{2}(a + b - 2ab)$	$\frac{1}{2}(1 - a - b + 2ab)$
A_1B_2/A_1B_2	0	1	0	0
A_1B_2/A_2B_1	$\frac{1}{2}(a + b - 2ab)$	$\frac{1}{2}(1 - a - b + 2ab)$	$\frac{1}{2}(1 - a - b + 2ab)$	$\frac{1}{2}(a + b - 2ab)$
A_1B_2/A_2B_2	0	$\frac{1}{2}$	0	$\frac{1}{2}$
A_2B_1/A_2B_1	0	0	1	0
A_2B_1/A_2B_2	0	0	$\frac{1}{2}$	$\frac{1}{2}$
A_2B_2/A_2B_2	0	0	0	1

a = rate of gene conversion at locus A, b = rate of gene conversion at locus B. Here $\alpha = a + b - 2ab$.

property of populations at equilibrium, which is the relation between imperfect transmission and variation in marginal fitnesses.

Consider a population that is at an equilibrium for the frequencies of the haplotypes. Recursion (8) produces the equilibrium identity

$$\hat{\mathbf{z}} = \hat{\mathbf{Y}}\hat{\mathbf{D}}\hat{\mathbf{z}}. \tag{10}$$

With perfect transmission $\hat{\mathbf{z}} = \hat{\mathbf{D}}\hat{\mathbf{z}}$, so that

$$\hat{\mathbf{D}} = \operatorname{diag}\left(\frac{\hat{w}_i}{\hat{\bar{w}}}\right) = \mathbf{I},$$

and $\hat{w}_i = \hat{\bar{w}}$ for all haplotypes i such that $z_i > 0$. This is true regardless of any frequency dependence of the w's. Hence, we have the well known

Result 1: *With perfect transmission, all the haplotypes present in the population have equal marginal fitnesses at equilibrium.*

In the presence of imperfect transmission, there can be differences at equilibrium in the marginal fitnesses of the haplotypes present. Moreover, the potential magnitude of the variation in the equilibrium marginal fitnesses can be shown to depend on the maximum level of imperfect transmission, α, characteristic of the genetic system. On substituting the expressions (3) into the equilibrium equations (10), since $\mathbf{P} \geq 0$ we easily deduce that $\hat{w}_i/\hat{\bar{w}} \leq (1 - \alpha)^{-1}$ for all i.

Later we show that the strength of selection on the modifier gene depends quantitatively on the magnitude of the difference between the maximal marginal fitness in the population and the mean fitness. This quantity is a variant on the genetic load defined by CROW (1958), and we call it the "selection potential":

Definition 1: The *selection potential, V*, present in a population is defined as

$$V = \frac{\max_i(w_i)}{\bar{w}} - 1 = \frac{L}{1 - L}, \tag{11}$$

where L is the genetic load, $1 - (\bar{w}/\max_i(w_i))$, as defined by CROW (1958).

The upper bound to the equilibrium selection potential is

$$V \leq \frac{\alpha}{1 - \alpha}, \tag{12}$$

which increases without limit as α increases toward 1, and approaches 0 as α tends to 0. This definition of selection potential corresponds to the less common definition of genetic load used by EWENS (1979, p. 66).

GENERAL MODIFIER GENE MODEL

The general modifier gene model is a special case of formulation (5) with haplotypes undergoing both selection and imperfect transmission. The haplotype is partitioned so that one part, the set of selected loci, A, undergoes selection and generalized transmission, while the other part, the modifier locus M, controls the transmission of the first part. Thus haplotypes are indexed with two subscripts, so that

$$M_aA_i$$

represents the haplotype with modifier allele a and multilocus haplotype i for the set of loci under selection. Its frequency is z_{ai}. We assume that the modifier alleles have no effect on the fitness of their carriers. Thus, the fitness of genotype M_aA_j/M_bA_k will be w_{jk} for all a, b. We use the notation A_i for haplotypes that consist of alleles $G_{1i}G_{2i}G_{3i}\cdots$ with the subscripts 1, 2, 3, etc., indicating the selected gene locus. It will be assumed throughout that the location of modifier locus is external to the A haplotype. If G_1 is the selected locus nearest the modifier locus, we denote by r the probability of recombination between M and G_1.

We assume that the modifier alleles are perfectly transmitted regardless of the processes acting on the selected haplotypes, so that the only force acting on the modifier locus is the induced selection resulting from its effect on the transmission of the other loci. To take account of recombination between M and G_1 we introduce two new transmission matrices, $\mathbf{T}^*$ and $\tilde{\mathbf{T}}^*$, which describe the transmission of the complete system of modifier and selected genes and also keep track of each modifier allele during transmission.

$T^*(ai \leftarrow aj \mid bk)$ is the probability that haplotype M_aA_i is produced from haplotype M_aA_j in genotype M_aA_j/M_bA_k, given that no recombination has occurred between M and G_1. $\tilde{T}^*(ai \leftarrow ak \mid bj)$ is the probability that haplotype M_aA_i is produced from haplotype M_aA_k in genotype M_aA_j/M_bA_k given that recombination has occurred between M and G_1. Conditioning in this way on recombination between M and G_1 takes account of any possible interference between crossovers in the region between M and G_1 and the transmission of the A haplotype. In the absence of interference, $T^*(ai \leftarrow aj \mid bk) = \tilde{T}^*(ai \leftarrow aj \mid bk)$ for all a, b, i, j, k.

The recursion (8) for the general modifier model with random mating can then be rewritten as

$$z'_{bi} = \frac{1}{\bar{w}} \sum_{cjk} z_{bj}z_{ck}w_{jk}[(1 - r)T^*(bi \leftarrow bj \mid ck) + r\tilde{T}^*(bi \leftarrow bk \mid cj)]. \tag{13}$$

In terms of $\mathbf{T}^*$ and $\tilde{\mathbf{T}}^*$ the original $\mathbf{T}$ defined in (1) can be written

$$T(ai \leftarrow aj, bk) = \frac{1}{2}(1 - r)T^*(ai \leftarrow aj \mid bk) + \frac{r}{2}\tilde{T}^*(ai \leftarrow ak \mid bj) \quad \text{for} \quad a \neq b,$$

$$T(ai \leftarrow aj, ak)$$

$$= \frac{1}{2}(1 - r)[T^*(ai \leftarrow aj \mid ak) + T^*(ai \leftarrow ak \mid aj)]$$

$$+ \frac{r}{2}[\tilde{T}^*(ai \leftarrow ak \mid aj) + \tilde{T}^*(ai \leftarrow aj \mid ak)],$$

and $T(ai \leftarrow bj, ck) = 0$ if $b \neq a$ and $c \neq a$, where Σ_i $T(ai \leftarrow aj \mid bk) = 1$ for all a, b, j, k, and $T(ci \leftarrow aj \mid bk) = \tilde{T}(ci \leftarrow aj \mid bk) = 0$ if $c \neq a$.

The original T's are not used after this point, so for convenience the asterisks from the new sets of T's are suppressed. The new T's can be represented in matrix form as

$$\mathbf{T}_{ab} = \|T(ai \leftarrow aj_1 \mid bj_2)\|_{i,j_1,j_2=1}^{n}$$

and

$$\tilde{\mathbf{T}}_{ab} = \|\tilde{T}(ai \leftarrow aj_1 \mid bj_2)\|_{i,j_1,j_2=1}^{n}.$$

In the absence of interference $\mathbf{T}_{ab} = \tilde{\mathbf{T}}_{ab}$. It will be assumed that there are no position effects between M and A. That is, the linkage phase between M and A within the diploid genotype does not affect the transmission of A. This is expressed as:

$$T(ai \leftarrow aj \mid bk) = T(bi \leftarrow bj \mid ak) \quad \text{and}$$

$$\tilde{T}(ai \leftarrow aj \mid bk) = \tilde{T}(bi \leftarrow bj \mid ak),$$

or in matrix form, $\mathbf{T}_{ab} = \mathbf{T}_{ba}$ and $\tilde{\mathbf{T}}_{ab} = \tilde{\mathbf{T}}_{ba}$. This excludes the possibility that the modifier controls transmission processes that occur during the haploid phase of the life cycle.

Segregation distortion acting on the modifier locus is excluded here because this would constitute imperfect transmission of the modifier. This entails Σ_i $T(ai \leftarrow aj \mid bk) = 1$, Σ_i $\tilde{T}(ai \leftarrow aj \mid bk) = 1$ for all a, b, j, k. This restriction is actually necessary to ensure that $\mathbf{T}$ and $\tilde{\mathbf{T}}$ satisfy condition (1) in the original definition of transmission matrices.

If the *selected* haplotypes are perfectly transmitted in an organism of modifier genotype M_a/M_b, then

$$T_{id}(ai \leftarrow aj \mid bk) = \tilde{T}_{id}(ai \leftarrow aj \mid bk) = \delta_{ij} = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}, \tag{14}$$

Only if $r = 0$, however, will this constitute perfect transmission of the entire haplotype of both modifier and selected loci.

ANALYSIS OF THE RECURSIONS

In this section we investigate the fate of new modifier alleles introduced into the population in the neighborhood of a stable equilibrium.

The equilibrium identity: We introduce the following notation:

$$\hat{\mathbf{z}}_b = \begin{vmatrix} \hat{z}_{b1} \\ \hat{z}_{b2} \\ \vdots \end{vmatrix},$$

$$\boldsymbol{\Omega}_b = \left\| \sum_{ck} \hat{z}_{ck} \frac{w_{jk}}{\hat{w}_j} [(1 - r)T(bi \leftarrow bj \mid ck) + r\tilde{T}(bi \leftarrow bk \mid cj)] \right\|_{i,j=1}^{n}, \tag{15}$$

$$\hat{w}_i = \sum_{bj} \hat{z}_{bj} w_{ij},$$

and

$$\hat{\mathbf{D}} = \operatorname{diag}\left(\frac{\hat{w}_i}{\bar{\hat{w}}}\right),$$

as in (9). $\boldsymbol{\Omega}_b$ are column stochastic matrices.

At any equilibrium for (13), the following identity must be satisfied:

$$\hat{z}_{bi} = \frac{1}{\bar{\hat{w}}} \sum_{cjk} \hat{z}_{bj}\hat{z}_{ck}w_{jk}[(1 - r)T(bi \leftarrow bj \mid ck) + r\tilde{T}(bi \leftarrow bk \mid cj)]. \tag{16}$$

This can be written in a vector form analogous to (8):

$$\hat{\mathbf{z}}_b = \boldsymbol{\Omega}_b \hat{\mathbf{D}} \hat{\mathbf{z}}_b. \tag{17}$$

The "external stability" of equilibria: To investigate the long-term evolutionary fate of the modifier locus, we need to know those characteristics of new modifier alleles that determine whether they will increase in the population when introduced near an equilibrium. Let a new modifier allele, M_a, be introduced into the population near $\hat{\mathbf{z}}$, with ϵ_{ai} the frequency of haplotype M_aA_i. From (13) the vector form of the linearized recursion for the ϵ_{ai}'s, ignoring terms of order ϵ_{ai}^2, is

$$\boldsymbol{\epsilon}' = \boldsymbol{\Omega}_a \hat{\mathbf{D}} \boldsymbol{\epsilon}. \tag{18}$$

Assuming that the polymorphic equilibrium $\hat{\mathbf{z}}$ is locally stable to perturbations in the frequencies of the haplotypes already present, the initial increase of the new modifier allele is determined by the spectral radius of the stability matrix,

$$\rho(\boldsymbol{\Omega}_a \hat{\mathbf{D}}). \tag{19}$$

If $\rho(\boldsymbol{\Omega}_a \hat{\mathbf{D}}) > 1$, the new modifying allele increases. If $\rho(\boldsymbol{\Omega}_a \hat{\mathbf{D}}) < 1$, it decreases and if $\rho(\boldsymbol{\Omega}_a \hat{\mathbf{D}}) = 1$, the modifying allele cannot change at a geometric rate. We call these the "external stability" properties of the equilibrium.

REDUCTION PRINCIPLE

Since $\boldsymbol{\Omega}_a$ is a column stochastic matrix, the spectral radius $\rho(\boldsymbol{\Omega}_a \hat{\mathbf{D}})$ can be different from 1 only if $\hat{\mathbf{D}} \neq \mathbf{I}$; i.e., only if there is a nonzero variance in the marginal

fitnesses of the selected haplotypes at equilibrium. Thus, when a population is near an equilibrium, selection cannot act on a newly introduced modifier allele to change its frequency at a geometric rate unless there is an equilibrium selection potential.

In this section we present the main results about the initial increase properties of a new modifier allele introduced to the randomly mating population at a polymorphic equilibrium. These results concern *external stability* of the equilibria with respect to changes in the direction of the modifier frequencies. Although the assumption of the existence of the equilibria may entail constraints on the selection regimes and transmission probabilities, in none of the results do these constraints or closed form solutions of the equilibrium frequencies appear. The only generic assumptions made are as follows: (1) There is polymorphism in the selected haplotypes, and (2) new modifier alleles do not cause the selected haplotypes present at equilibrium to be transformed into haplotypes not present.

Our results regarding the evolution of reduced levels of imperfect transmission fall into two cases according to the effect of the new modifier allele on transmission:

Case 1. Imperfect transmission eliminated at the selected loci

Result 2: a) *A population at equilibrium with a nonzero selection potential can always be invaded by a new modifier allele that eliminates imperfect transmission at the selected loci, for any amount of linkage of the modifier to these loci.*

b) *In particular, for a modifier locus that is absolutely linked to the selected loci (which are absolutely linked in the presence of the new modifier allele), the spectral radius of the external local stability matrix equals one plus the equilibrium selection potential, i.e.,*

$$\rho(\mathbf{\Omega}_a\hat{\mathbf{D}}) = 1 + \hat{V} = \max_i \left(\frac{\hat{w}_i}{\hat{\bar{w}}}\right).$$

c) *Moreover, for all $r \leq \frac{1}{2}$, the spectral radius of the external local stability matrix is always greater than one by at least the equilibrium marginal fitness variance, i.e.,*

$$\rho(\mathbf{\Omega}_a\hat{\mathbf{D}}) \geq 1 + \text{var}\left(\frac{\hat{w}_i}{\hat{\bar{w}}}\right).$$

d) *For a given set of equilibrium fitnesses and haplotype frequencies, the selective advantage of a new modifier allele that eliminates imperfect transmission decreases with looser linkage to the selected haplotypes.*

e) *The selection potential is always greater than the population variance in the marginal fitnesses.*

The proofs of this and subsequent results are given in the APPENDIX.

Remarks: (i) Result e) does not require that the population be at equilibrium, but holds for populations in the transient phase of their dynamics as well.

(ii) In principle, because the frequency vector $\mathbf{v}$ and the fitness matrix $\mathbf{W}$ are the only facts about the polymorphic equilibria that are relevant to the fate of the new modifier allele, Result 2 can be shown to hold no matter what linkage arrangements obtain for those modifier alleles and selected loci segregating at the polymorphism. Moreover, the modifier alleles present at the polymorphism may even control their own linkage to the selected loci without changing this result. The *new* modifier allele, however, must fit the assumptions given earlier (that it be external to the selected haplotype and have a single recombination rate r with the nearest selected locus) for the proof of Result 2 to hold.

Case 2. Linear variation in the transmission

So far we have considered only modifier alleles that produce large reductions in the imperfect transmission acting on a set of selected loci. Now we consider modifiers that can change the level of imperfect transmission over a continuous range of values. Here, the way changes in $\mathbf{T}$ reflect variation at the modifier locus is critical in the evolution of modification.

The simplest kind of variation in transmission constitutes a basic "building block" of variation for more complex systems. Suppose that each selected haplotype, A_j, has a certain probability, m, of being "hit" by some transforming process, and given that it is hit, it is transformed into various other selected haplotypes with different probabilities. These probabilities, $T_1(i \leftarrow j|k)$, include possible dependence on both haplotypes A_j and A_k in the diploid genotype. If the effect of the modifier gene is to scale the "hit" rate, m, up or down equally for all haplotypes, it will be said to produce "linear" variation in the transmission.

Definition 2: For *linear variation*, the transmission matrix for modifier genotype M_a/M_b can be represented as a weighted average,

$$\mathbf{T}(m_{ab}) = (1 - m_{ab})\mathbf{T}_{id} + m_{ab}\mathbf{T}_1,$$

where $\mathbf{T}_1$ is a transmission matrix that is independent of the allelic configuration at the M locus, and the modifier's sole effect to determine the parameters m_{ab}. Examples are discussed after Result 3.

We consider here the initial increase properties of a new modifier allele that produces linear variation in the transmission. Attention is restricted to the introduction of new modifier alleles into populations that are fixed at the modifier locus, but all of the initial increase results extend to a class of modifier polymorphisms discussed in LIBERMAN and FELDMAN (1986a,b; 1987).

Consider a population at a stable equilibrium satisfying (16) where the modifier locus is fixed for one allele, M_1, which produces transmission matrices

$$\mathbf{T}_1 = \|T(1i \leftarrow 1j_1 \,|\, 1j_2)\|_{i,j_1,j_2=1}^{n}$$

and

$$\tilde{\mathbf{T}}_1 = \|\tilde{T}(1i \leftarrow 1j_2 \,|\, 1j_1)\|_{i,j_1,j_2=1}^{n}.$$

Denote the total frequency of each selected haplotype by

$$v_i = \sum_b z_{bi}.$$

In this case $\hat{z}_{1k} = \hat{v}_k > 0$ for all k. At equilibrium, from (16) the selected haplotype frequencies must satisfy the identity

$$\hat{\mathbf{v}} = \Omega_1 \hat{\mathbf{D}}\hat{\mathbf{v}} = [(1 - r)\hat{\mathbf{Y}}_1 + r\hat{\tilde{\mathbf{Y}}}_1]\mathbf{D}\hat{\mathbf{v}}, \tag{20}$$

where

$$\hat{\mathbf{Y}}_1 = \|\sum_k \hat{v}_k \frac{w_{jk}}{\hat{w}_j} T(1i \leftarrow 1j \,|\, 1k)\|_{i,j=1}^{n}, \tag{21}$$

$$\hat{\tilde{\mathbf{Y}}}_1 = \|\sum_k \hat{v}_k \frac{w_{jk}}{\hat{w}_j} \tilde{T}(1i \leftarrow 1k \,|\, 1j)\|_{i,j=1}^{n}, \tag{22}$$

and $\hat{\mathbf{D}}$ is the diagonal matrix of relative marginal fitnesses defined in (15). $\hat{\mathbf{Y}}_1$ and $\hat{\tilde{\mathbf{Y}}}$ are both column stochastic.

Reduction principle for linear variation in transmission

A new modifier allele M_2, yielding transmission matrices $\mathbf{T}_2$ and $\tilde{\mathbf{T}}_2$ as a heterozygote with M_1 is introduced into the population. A general way to represent linear variation between $\mathbf{T}_1$ and $\mathbf{T}_2$, and between $\tilde{\mathbf{T}}_1$ and $\tilde{\mathbf{T}}_2$ uses (14), and Definition 2, as follows:

$$T(2i \leftarrow 2j \,|\, 1k) = (1 - m_{21})\delta_{ji} + m_{21}T(1i \leftarrow 1j \,|\, 1k) \tag{23a}$$

is the relation between $\mathbf{T}_1$ and $\mathbf{T}_2$, and

$$\tilde{T}(2i \leftarrow 2k \,|\, 1j) = (1 - m_{21})\delta_{ki} + m_{21}\tilde{T}(1i \leftarrow 1k \,|\, 1j) \tag{23b}$$

is the relation between $\tilde{\mathbf{T}}_1$ and $\tilde{\mathbf{T}}_2$. Here m_{21} represents the parameter for M_2/M_1.

Remark: Recall from the characterization of linear variation that m_{11} and m_{21} are interpreted as being the overall probability that the selected haplotypes are "hit" by a transforming process. The above definitions of $\mathbf{T}_2$ and $\tilde{\mathbf{T}}_2$ in terms of $\mathbf{T}_1$ and $\tilde{\mathbf{T}}_1$ require that m_{11} be scaled to 1. Thus m_{21} is the "hit" rate for modifier genotype M_2/M_1 relative to modifier genotype M_1/M_1. We will refer to m_{21} as the value of the "modified parameter" for modifier genotype M_2/M_1.

Define

$$\hat{\mathbf{Y}}_2 = \|\sum_k \hat{v}_k \frac{w_{jk}}{\hat{w}_j} T(2i \leftarrow 2j \,|\, 1k)\|_{i,j=1}^{n}$$

and

$$\hat{\tilde{\mathbf{Y}}}_2 = \|\sum_k \hat{v}_k \frac{w_{jk}}{\hat{w}_j} \tilde{T}(2i \leftarrow 2k \,|\, 1j)\|_{i,j=1}^{n}.$$

Therefore,

$$\begin{aligned} \hat{\mathbf{Y}}_2 &= (1 - m_{21})\mathbf{I} + m_{21}\hat{\mathbf{Y}}_1 \quad \text{and} \\ \hat{\tilde{\mathbf{Y}}}_2 &= (1 - m_{21})\mathbf{Q} + m_{21}\hat{\tilde{\mathbf{Y}}}_1, \end{aligned} \tag{24}$$

where

$$\mathbf{Q} = \|\hat{v}_i \frac{w_{ij}}{\hat{w}_j}\|_{i,j=1}^{n} = \operatorname{diag}(\hat{\mathbf{v}})\ \mathbf{W}\ \operatorname{diag}(\mathbf{W}\hat{\mathbf{v}})^{-1}. \tag{25}$$

From (18), the recursion on the frequencies of haplotypes containing the new modifier allele is

$$\epsilon' = \Omega_2 \hat{\mathbf{D}}\epsilon = [(1 - r)\hat{\mathbf{Y}}_2 + r\hat{\tilde{\mathbf{Y}}}_2]\hat{\mathbf{D}}\epsilon. \tag{26}$$

M_2 will increase when rare if and only if

$$\rho(\Omega_2\hat{\mathbf{D}}) = \rho([(1 - r)\hat{\mathbf{Y}}_2 + r\hat{\tilde{\mathbf{Y}}}_2]\hat{\mathbf{D}}) > 1.$$

Result 3: *A new allele at a modifier locus that is absolutely linked to a set of selected loci is introduced into a population near an equilibrium in which the modifier is fixed, with a nonzero selection potential for the selected haplotypes. If the new modifier allele produces linear variation in the transmission of the selected loci, then if* Ω_2 *is irreducible (reducible)*

a) *the new modifier allele will increase (not decrease) in frequency at a geometric rate if it brings the transmission closer to perfect transmission, i.e.,* $m_{21}/m_{11} < 1$ *and*

b) *it will decrease (not increase) in frequency at a geometric rate if it takes the transmission further away from perfect transmission, i.e.,* $m_{21}/m_{11} > 1$.

Moreover,

c) *the asymptotic rate of change in the frequency of the modifier allele is an increasing function of the change it makes in the transmission probabilities.*

Remarks: (i) The result is proven for $r = 0$ but will hold also for sufficiently small positive values of r (KARLIN and MCGREGOR 1972) when $\mathbf{Y}_2$ is irreducible.

(ii) The proof of this result employs Theorem 5.2 of KARLIN (1982) and is included in the APPENDIX. The variation in the transmission for which this Theorem applies must be linear and this is the condition in the analysis of mutation modification by LIBERMAN and FELDMAN (1986b) that yields reduction. In their model, the equilibrium relation with the modifier fixed on M_1 is

$$\hat{\mathbf{v}} = \begin{vmatrix} 1 - \mu_{11} & \nu_{11} \\ \mu_{11} & 1 - \nu_{11} \end{vmatrix} \begin{vmatrix} \frac{\hat{w}_1}{\hat{\bar{w}}} & 0 \\ 0 & \frac{\hat{w}_2}{\hat{\bar{w}}} \end{vmatrix} \hat{\mathbf{v}},$$

and, setting $r = 0$, the recursion for the new modifier allele M_2 is

$$\epsilon' = \begin{vmatrix} 1-\mu_{12} & \nu_{12} \\ \mu_{12} & 1-\nu_{12} \end{vmatrix} \begin{vmatrix} \frac{\hat{w}_1}{\hat{\bar{w}}} & 0 \\ 0 & \frac{\hat{w}_2}{\hat{\bar{w}}} \end{vmatrix} \epsilon.$$

Set positive constants $m_{11}, f_{11}, f_{12}, f_{21}, f_{22}$ such that

$$\begin{vmatrix} 1-\mu_{11} & \nu_{11} \\ \mu_{11} & 1-\nu_{11} \end{vmatrix} = (1-m_{11}) \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} + m_{11} \begin{vmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{vmatrix},$$

where $f_{11} + f_{21} = f_{12} + f_{22} = 1$; then linear variation requires that μ_{12} and ν_{12} satisfy

$$\begin{vmatrix} 1-\mu_{12} & \nu_{12} \\ \mu_{12} & 1-\nu_{12} \end{vmatrix} = (1-m_{12}) \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} + m_{12} \begin{vmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{vmatrix},$$

which entails $\nu_{11}/\mu_{11} = \nu_{12}/\mu_{12}$, *i.e.*, there exists b such that $\nu_{ij} = b\mu_{ij}$ for all i, j. This is the assumption under which Liberman and Feldman obtain the reduction result, namely $\mathbf{N} = b\mathbf{M}$. The definition of linear variation here includes the previous models of recombination modification by Nei (1967), Feldman (1972), Feldman and Balkau (1973) as well as the migration and mutation modification models of Balkau and Feldman (1973), Karlin and McGregor (1974) and Teague (1977).

Linear variation embodies the fourth condition given in Liberman and Feldman (1986a) for the reduction principle to hold for modifiers of recombination, mutation, and migration, namely, that besides viability selection, the only evolutionary force in the system is that feature of the genetic system subject to genetic modification. This requirement is illustrated in the models of recombination modification in Feldman, Christiansen and Brooks (1980). In their Model I, which produces the reduction result, with $r = 0$ the stability recursion on the new modifier allele is

$$\epsilon' = (1 - r_2)\hat{\mathbf{D}}\epsilon + r_2 \frac{1}{\bar{w}} \begin{vmatrix} \hat{v}_1 w_{11} + \hat{v}_3 w_{13} & \hat{v}_1 w_{12} + \hat{v}_3 w_{23} & 0 & 0 \\ \hat{v}_2 w_{12} + \hat{v}_4 w_{14} & \hat{v}_2 w_{22} + \hat{v}_4 w_{24} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \hat{v}_1 w_{13} + \hat{v}_3 w_{33} & \hat{v}_1 w_{14} + \hat{v}_3 w_{34} \\ 0 & 0 & \hat{v}_2 w_{23} + \hat{v}_4 w_{34} & \hat{v}_2 w_{24} + \hat{v}_4 w_{44} \end{vmatrix} \epsilon$$

where r_2 is the recombination rate produced by the new modifier heterozygote. This again fits the form for linear variation, and Result 3 can be seen to apply.

Strength of selection on modifiers

Previous treatments of modifier gene evolution have generally not examined the magnitude of the selective forces that act on modifier genes, one exception being the work on inversions by Charlesworth and Charlesworth (1973). Wright (1964) stressed that pleiotropic effects of modifier genes could cause them to have intrinsic fitness differences. It has generally been felt that such selection on the modifier locus would overwhelm any selection due to its effects on transmission (*e.g.*, Karlin and McGregor 1974). Here we show that the selection on modifier genes due to their effects on transmission can be quite strong.

From Result 2b, in the case of an absolutely linked modifier allele that eliminates all imperfect transmission, its induced marginal selective advantage is equal to the equilibrium *selection potential*. This advantage decreases with looser linkage of the modifier but is always greater than the *population variance* in the equilibrium marginal fitnesses (Results 2c and 2d).

For one special case of transmission, we can derive an estimate of the amount of selection on a new modifying allele that yields linear variation in the transmission. It will be a function of how far the new value of the modified parameter deviates from the parameter value of the population at the original equilibrium. The transmission for which this estimate can be obtained is the "house of cards" mutation distribution defined by Kingman (1980). This is simply a mutation process which has no memory; all haplotypes mutate at the same frequency, and the probability that the mutant is a given haplotype does not depend on what the original haplotype was. There are several familiar analogs to "house of cards" (*i.e.*, memoryless) distributions for models of migration, including the Wright island model, the Levene model, and the Deakin model (see Karlin, 1982).

With linear variation in the transmission, and no interference between the mutation process and recombination between M and A, the general form for the transmission matrix with a memoryless distribution is

$$T(ai \leftarrow aj \mid bk) = (1 - m_{ab})\delta_{ji} + m_{ab}p_i. \qquad (27)$$

where m_{ab} is the overall mutation rate and p_i is the probability of producing selected haplotype i given that there is a mutation, and $\mathbf{T} = \tilde{\mathbf{T}}$. The modifier, therefore, may change the overall rate but not the relative distribution of mutations.

Suppose that the population is at an equilibrium, with a modified parameter of m_{11}. Substitution of (27) into (21) yields

$$\hat{\mathbf{Y}}_1 = (1 - m_{11})\mathbf{I} + m_{11}\mathbf{P}, \quad \text{and} \quad \hat{\tilde{\mathbf{Y}}}_1 = \hat{\mathbf{Y}}_1\mathbf{Q}, \qquad (28)$$

where $\mathbf{P} = \text{diag}(\mathbf{p})\mathbf{U}$, $\mathbf{p}^T = (p_1, p_2, \cdots, p_n)$, and $\mathbf{U}$ is

the n by n matrix of ones. The equilibrium value $\hat{\mathbf{v}}$ solves the identity

$$\hat{\mathbf{v}} = [(1 - m_{11})\mathbf{I} + m_{11}\mathbf{P}]\,\mathbf{D}\hat{\mathbf{v}} = (1 - m_{11})\mathbf{D}\hat{\mathbf{v}} + m_{11}\mathbf{p} \tag{29}$$

for any r, which emerges from (20) using the identity $\mathbf{Q}\hat{\mathbf{D}}\hat{\mathbf{v}} = \hat{\mathbf{D}}\hat{\mathbf{v}}$.

A new modifier allele, M_2, is now introduced into the population with m_{21} as the value of the modified parameter in the M_2/M_1 heterozygotes. From (26) the linearized recursion on the frequencies of the haplotypes bearing M_2 is

$$\epsilon' = \hat{\mathbf{Y}}_2[(1 - \mathrm{r})\mathbf{I} + r\mathbf{Q}]\hat{\mathbf{D}}\epsilon$$

where m_{21} has been substituted for m_{11} in (28) to produce $\hat{\mathbf{Y}}_2$. Since there is assumed to be no interference, $\hat{\bar{\mathbf{Y}}}_2 = \hat{\mathbf{Y}}_2\mathbf{Q}$. Then we have the following

Result 4: *Consider the case of linear variation in transmission with a memoryless distribution specified by (28), with $p_i > 0$ for all i. For a tightly linked modifier, the spectral radius of the stability matrix for a new modifying allele with parameter m_{21} as a heterozygote, with m_{21} close to $m_{11} > 0$, is approximately:*

$$\rho(\hat{\mathbf{Y}}_2\hat{\mathbf{D}}) \cong 1 + (m_{11} - m_{21}) \cdot (\hat{\bar{w}}m_{11}^2 \sum_i p_i^{-1}\hat{v}_i^2\hat{w}_i)^{-1} \sum_i p_i^{-1}\hat{v}_i^2(\hat{w}_i - \hat{\bar{w}})^2.$$

Remarks: The spectral radius is the actual value for the asymptotic relative marginal fitness of the new modifier allele, *i.e.*,

$$\frac{\hat{w}_{M_2}}{\hat{\bar{w}}} = \rho(\hat{\mathbf{Y}}_2\hat{\mathbf{D}}).$$

Consistent with Result 3, a new modifier allele can increase if and only if it reduces the amount of mutation. The term

$$\sum_i p_i^{-1}\hat{v}_i^2(\hat{w}_i - \hat{\bar{w}})^2 \tag{30}$$

is of the same order as the marginal fitness variance and is zero if and only if the variance of the marginal fitnesses is zero. Therefore, for small modifications of the transmission process the selection for or against the new modifier allele will be on the order of the equilibrium fitness variance in the population times the deviation of the value of its modified parameter from the equilibrium value before its introduction.

The selective force on the modifier may be interpreted as a consequence of the selection on the major genes rather than imperfect transmission. By substituting from an alternative form of the equilibrium identity (29), namely

$$\frac{m_{11}}{1 - m_{11}}\hat{\bar{w}}(\hat{v}_i - p_i) = \hat{v}_i(\hat{w}_i - \hat{\bar{w}}),$$

(30) becomes

$$\hat{\bar{w}}^2 m_{11}^2(1 - m_{11})^{-2} \sum_i p_i^{-1}(\hat{v}_i - p_i)^2. \tag{31}$$

If no selection were acting then under mutation alone the equilibrium haplotype frequencies would all be

$$\hat{v}_i = p_i.$$

Thus the squared term is the deviation of the selected haplotype frequencies from what they would be under pure mutation. The effect of adding selection to a system of pure mutation is to create selection for decreased mutation.

Selection on inversions

A new chromosomal inversion is formally the same as a new modifier allele that eliminates recombination between the loci flanking the inversion, and is tightly linked to them. If we consider the two flanking loci in isolation, a rate ξ of recombination between them, prior to the introduction of the inversion, permits a maximum selection potential of $\xi/(1 - \xi)$ when there is a two-locus polymorphic equilibrium [see (12)]. Result 2b can therefore be used to give an upper bound for the selective force on an inversion in terms of the map length alone. The expression for the selective force given by CHARLESWORTH and CHARLESWORTH (1973) requires in addition an estimate of the linkage disequilibrium.

Result 5: *The relative amount of induced selective advantage on a chromosomal inversion can range up to*

$$\frac{\xi}{1 - \xi},$$

where ξ is the map length of the inversion, in units of crossover frequency.

Remark: In the examples of CHARLESWORTH and CHARLESWORTH (1973), the strength of selection on the inversion is seen to approach $\xi/(1 - \xi)$ as epistasis and ξ become smaller. From the case in Table 1 of KARLIN and CARMELLI (1975), where the initial increase of an appropriate new inversion violates the "mean fitness principle" (KARLIN and MCGREGOR 1974), the strength of selection on inversions bordered by two loci that are 2.5 to 3 centimorgans apart can be seen to be more than 98.8% of $\xi/(1 - \xi)$.

DISCUSSION

General considerations: We have shown in this paper that many genetic models in the literature that incorporate complex transmission, selection, recombination, mutation, and modification, are special cases of recursion (5). A basic finding from this general framework is the intimate connection between imperfect transmission, equilibrium selection potential (*i.e.*,

variation in marginal fitnesses), and the evolution of transmission under the control of modifier genes. The "selection potential" we have defined enters quantitatively into the strength of selection acting on modifier genes by providing an upper bound for the potential relative rate of increase in frequency of any haplotype in the population due to selection. With perfect transmission, this value vanishes at any equilibrium. In the presence of imperfect transmission, equilibria may be reached where the selection potential is kept above zero by an amount dependent on the level of imperfect transmission. This balance occurs because haplotypes that are depleted by selection are replenished from other haplotypes by imperfect transmission, and conversely, haplotypes whose numbers are augmented by selection are depleted by imperfect transmission.

Selective forces on modifier genes: When the modifier locus is perfectly transmitted, as we have assumed throughout, the only force changing the frequencies of its alleles is selection, manifested as differences in the alleles' marginal fitnesses. When the modifier locus is selectively neutral, the marginal fitness of each modifier allele is simply the average of the fitnesses of the selected haplotypes with which it is associated. Summing over i in (13) shows that

$$x_b' = \frac{1}{\bar{w}} \sum_{cjk} z_{bj} z_{ck} w_{jk} = \frac{1}{\bar{w}} x_b w_{M_b}$$

where x_b is the total frequency of modifier allele M_b, and

$$w_{M_b} = \sum_j z_{bj} w_j / x_b$$

is its marginal fitness.

These marginal fitnesses would clearly be the same for all of the modifier alleles if they were randomly associated with the selected haplotypes. Evolution of modifier genes occurs because, through their effects on transmission, they may be able to create linkage disequilibrium between themselves and the selected haplotypes. Those modifier alleles that occur more frequently with the fitter types in the population increase in frequency. This is known as "hitchhiking," and is the fundamental mechanism by which neutral, perfectly transmitted modifier genes evolve.

The dynamics of hitchhiking in populations that are in a transient phase of their evolution differ greatly from the dynamics in populations that are near equilibrium (CHARLESWORTH, 1976; STROBECK, MAYNARD SMITH and CHARLESWORTH, 1976). Hitchhiking is usually thought of as a process whereby an allele at one locus increases in frequency by being linked to an allele at another locus that is increasing in frequency due to selection (EWENS 1979, p. 205). It might be expected that at an equilibrium, because there are no changes in the frequencies, there could be no hitchhiking effects at all. But to the contrary, Result 2a demonstrates that in fact hitchhiking can occur in the absence of linkage between the modifier locus and the selected loci, and when the alleles of the selected loci are very close to equilibrium.

This paradox is resolved by noting that when imperfect transmission permits a positive selective potential at equilibrium, there is a constant net "flow," by imperfect transmission, from the fitter haplotypes to the less fit. By altering this "flow," a modifier allele can come to be in disequilibrium with the selected haplotypes, and may therefore acquire a marginal fitness different from the mean of the population. For Result 2a, the linkage disequilibrium between the modifier and the selected loci that causes hitchhiking is generated by selection every generation, and even free recombination can no more than halve this disequilibrium by the next phase of selection. It should be noted that any continuous dependence of $\mathbf{T}$ or $\mathbf{W}$ on allele frequencies does not enter into $\Omega_a\hat{\mathbf{D}}$, so provided that $\rho(\Omega_a\hat{\mathbf{D}}) \neq 1$ neither the magnitude nor direction of selection on a new modifier allele are affected by frequency dependence.

What do the estimates from Results 2 and 4 for the strength of selection on a new modifier allele suggest for the order of magnitude of selection on modifiers in nature? The selection potential is the maximum possible rate of increase for modifier alleles, as well as selected alleles, and is exploited to its fullest by a modifier that stops all imperfect transmission and is tightly linked to the selected haplotype.

As the guaranteed level of perfect transmission of the haplotypes decreases, *i.e.*, as the value of α increases, the upper bound on the equilibrium selection potential, $\alpha/(1-\alpha)$, increases without limit (as α goes to one). The typical values of α in nature depend on the particular imperfect transmission process. In the case of mutation, even though per-locus mutation rates are quite small, per-haplotype rates can be quite high, of order 1. Similarly, per-chromosome recombination rates can be on this order (VON WETTSTEIN, RASMUSSEN and HOLM 1984, p. 399). Therefore, the selection potentials that may be typical of natural populations at equilibrium, which are the potential strengths of selection that can be induced on neutral modifier genes, may be quite strong. So we may consider what happens when selection acts directly on the modifier locus, which may be a biologically more reasonable assumption given the ubiquity of pleiotropy (WRIGHT 1964). We see that the induced selection on a modifier due to its effect on transmission can overwhelm considerable direct selection acting on the modifier. If direct selection gives modifier genotype M_1/M_2 a selective value f_{12}, and interacts multiplicatively with selection on the major loci, then the fitness of genotype M_1A_i/M_2A_j will be $f_{12}w_{ij}$, and the rate of change in the frequency of modifier allele M_2 when rare will be $f_{12}\rho(\Omega_2\hat{\mathbf{D}}) - 1$.

Reduction principle: In Result 2, the condition for the invasion of a new modifier allele was that it eliminate imperfect transmission acting on the selected loci provided a nonzero selection potential is maintained at equilibrium by imperfect transmission. This is seen to be true for all viability regimes, and all polymorphisms at the modifier and selected loci. Result 3, for the case of linear variation, is more restricted in scope, applying only for tightly linked modifier genes that produce linear variation in the transmission, with the population fixed on the modifier gene. In this case, a new modifier allele introduced into a population bearing a selection potential induced by imperfect transmission can increase if and only if it takes the transmission closer to perfect transmission, and is excluded if it takes the transmission further away. We conjecture that our results for linear variation will hold for arbitrary levels of linkage between the modifier and selected genes.

Result 3 has several implications. In previous treatments of recombination modification, only two loci under selection with two alleles each were allowed, and it was assumed that both double heterozygotes at the selected loci had the same fitness, and had the same crossover frequencies for any given modifier genotype. None of these assumptions enters Result 3: there may be multiple alleles at each selected locus and there may be multiple loci under selection. If there are more than two selected loci, however, for the variation of the transmission to be linear, there must be complete interference between crossover events among the loci, and the modifier must simultaneously control every crossover event for the set of loci. Different fitnesses are allowed for the different linkage phases in the multiple heterozygotes at the selected loci (*i.e.*, position effects on fitness are allowed). For any given modifier genotype, different recombination rates for the different double heterozygotes at the selected loci are also allowed. For multiple loci under selection, the crossover frequency distribution among the loci need not be equal for the different heterozygotes. All that is required is that the modifier gene affect the crossover rates linearly.

Most of the basic models of modification in the literature treat linear variation, including models of mutation modification (LIBERMAN and FELDMAN 1986b; HOLSINGER and FELDMAN 1983b; KARLIN and MCGREGOR 1974), recombination modification (FELDMAN, CHRISTIANSEN and BROOKS 1980; TEAGUE 1976; CHARLESWORTH 1976) and migration modification (BALKAU and FELDMAN 1973; KARLIN and MCGREGOR 1974; TEAGUE 1977; LIBERMAN and FELDMAN 1988). Several factors cause a model to deviate from linear variation, including nonrandom mating, and the involvement of more than one imperfect transmission process during the life cycle. Also, the modifier may not simply vary the overall chance that a selected haplotype is "hit" by an imperfect transmission event, but may change the *distribution* of gametic haplotypes produced by a given parental genotype, and this also produces deviations from linear variation.

Examples of nonlinear variation include models of recombination modification in the presence of mutation (FELDMAN, CHRISTIANSEN and BROOKS 1980), models such as recombination modification in the presence of migration (CHARLESWORTH and CHARLESWORTH 1979b), the model of CHRISTIANSEN and FELDMAN (1975) where a modifier controls either recombination or migration in a population undergoing recombination and migration, and the model of a modifier acting on recombination between multiple loci by CHARLESWORTH and CHARLESWORTH (1979a). In the next section we discuss some examples of nonlinear variation that lead to increases in the modified parameter.

Evolutionary genetic stability of perfect transmission: ESHEL and FELDMAN (1982) developed "evolutionary genetic stability" of a phenotype as a first order criterion for establishing that the long-term tendency of evolution is toward this phenotype. In their papers on the general reduction principle in the cases of mutation, recombination, and migration modification, LIBERMAN and FELDMAN (1986a,b; 1988) conjecture that for the models they consider, the phenotypes of zero mutation, zero recombination, and zero migration have evolutionary genetic stability given four conditions:

1. The selection of the major locus (or loci) is at the level of differential genotypic viabilities.
2. These viabilities do not change over time.
3. There is random mating.
4. Besides viability selection, the only evolutionary force operating is that feature of the genetic system subject to genetic modification.

In the cases analyzed here, we would expect that after its initial increase a new modifying allele would change in frequency until a new equilibrium is approached at which the population average parameter value is less than that at the initial equilibrium. In other words, perfect transmission should be the state towards which evolution proceeds in these cases. The proof of this result, analogous to that of KARLIN and LESSARD (1984), appears to be significantly more difficult than that of our results here.

Because perfect transmission is not the norm in nature, with the ubiquity of mutation and the prevalence of different kinds of recombination, it is important to ask whether there are ever conditions under which a new modifier allele that increases the level of imperfect transmission in the population can invade. In fact, there are cases where it can. When selection fluctuates, recombination rates (CHARLESWORTH 1976) or mutation rates (GILLESPIE 1981a) may in-

crease. When the population is of finite size and random genetic drift is included in the model, GILLESPIE (1981b) has shown that mutation rates may increase and the same has been shown for recombination by FELSENSTEIN (1974) and FELSENSTEIN and YOKOYAMA (1976). When one gene controls the segregation pattern at another locus, which is also subject to viability selection, the recombination between the pair is not necessarily reduced (PROUT, BUNGAARD and BRYANT 1973; THOMSON and FELDMAN 1974). Another example is provided by the case of recombination modification analyzed by FELDMAN, CHRISTIANSEN and BROOKS (1980), where the loci whose linkage is controlled by the modifier also undergo mutation. In this case there are conditions where a new modifier allele can invade the population if it *increases* the frequency of recombination between the loci.

How is this latter result possible in light of Result 3 on modifiers producing linear variation in transmission? The answer is that in the presence of mutation, the variation produced by a modifier of recombination is not linear. In this example, condition 4 in the conjecture of LIBERMAN and FELDMAN (1986a,b), above, is violated. In other cases where increases in imperfect transmission can evolve in populations near equilibrium, the variation in the transmission again does not fit the form for linear variation. This is the case, for example, when recombination can evolve to increase in the presence of migration (CHARLESWORTH and CHARLESWORTH, 1979b), or selfing (CHARLESWORTH, CHARLESWORTH and STROBECK 1979; HOLSINGER and FELDMAN 1983a), or mutation can increase in the presence of fertility selection (HOLSINGER, FELDMAN and ALTENBERG 1986) or selfing (HOLSINGER and FELDMAN 1983b). The generalized reduction results we obtain here show that it is the nature of the *variation* in the transmission process that determines the direction of its evolution. Factors influencing the functional form of this variation should therefore be the focus for developing further understanding of the evolution of genetic transmission.

We wish to thank FREDDY CHRISTIANSEN, URI LIBERMAN, LUIGI CAVALLI-SFORZA, JOSEPH OLIGER, SHRIPAD TULJAPURKAR, KENT HOLSINGER and JOEL PECK for their helpful comments, and RICHARD LEWONTIN for the idea (in the context of cultural evolution) that selection and transmission form a useful dichotomy.

This research was supported in part by research grants GM 28016, GM 10452, GM 07181 and FD 2589-9414L from the National Institutes of Health.

LITERATURE CITED

BALKAU, B. J., and M. W. FELDMAN, 1973 Selection for migration modification. Genetics **74:** 171–174.

CHARLESWORTH, B, 1976 Recombination modification in a fluctuating environment. Genetics **83:** 181–195.

CHARLESWORTH, B., and D. CHARLESWORTH, 1973 Selection of new inversions in multi-locus genetic systems. Genet. Res., **21:** 167–183.

CHARLESWORTH, D., and B. CHARLESWORTH, 1979a Selection on recombination in a multi-locus system. Genetics **91:** 575–580.

CHARLESWORTH, D., and B. CHARLESWORTH, 1979b Selection on recombination in clines. Genetics **91:** 581–589.

CHARLESWORTH, D., B. CHARLESWORTH and C. STROBECK, 1979 Selection for recombination in partially self-fertilizing populations. Genetics **93:** 237–244.

CHRISTIANSEN, F. B., and M. W. FELDMAN, 1975 Subdivided populations: a review of one- and two-locus deterministic theory. Theor. Popul. Biol. **7:** 13–38.

CROW, J. F., 1958 Some possibilities for measuring selection intensities in man. Hum. Biol. **30:** 1–13.

ESHEL, I., and M. W. FELDMAN, 1982 On evolutionary genetic stability of the sex ratio. Theor. Popul. Biol. **21:** 430–439.

EWENS, W. J., 1979 *Mathematical Population Genetics.* Springer-Verlag, Berlin.

FELDMAN, M. W., 1972 Selection for linkage modification. I. Random mating populations. Theor. Popul. Biol. **3:** 324–346.

FELDMAN, M. W., and B. BALKAU, 1973 Selection for linkage modification. II. A recombination balance for neutral modifiers. Genetics **74:** 713–726.

FELDMAN, M. W., and U. LIBERMAN, 1986 An evolutionary reduction principle for genetic modifiers. Proc. Natl. Acad. Sci. USA **83:** 4824–4827.

FELDMAN, M. W., F. B. CHRISTIANSEN and L. D. BROOKS, 1980 Evolution of recombination in a constant environment. Proc. Natl. Acad. Sci. USA **77:** 4838–4841.

FELSENSTEIN, J., 1974 The evolutionary advantage of recombination. Genetics **78:** 737–756.

FELSENSTEIN, J., and S. YOKOYAMA, 1976 The evolutionary advantage of recombination. II. Individual selection for recombination. Genetics **83:** 845–859.

GILLESPIE, J. H., 1981a Mutation modification in a random environment. Evolution **35:** 468–476.

GILLESPIE, J. H., 1981b Evolution of the mutation rate at a heterotic locus. Proc. Natl. Acad. Sci. USA **78:** 2452–2454.

HOLSINGER, K. E., and M. W. FELDMAN, 1983a Linkage modification with mixed random mating and selfing: a numerical study. Genetics **103:** 323–333.

HOLSINGER, K. E., and M. W. FELDMAN, 1983b Modifiers of mutation rate: Evolutionary optimum with complete selfing. Proc. Natl. Acad. Sci. USA **80:** 6732–6734.

HOLSINGER, K., M. W. FELDMAN and L. ALTENBERG, 1986 Selection for increased mutation rates with fertility differences between matings. Genetics **112:** 909–922.

KARLIN, S., 1982 Classifications of selection-migration structures and conditions for a protected polymorphism. Evol. Biol. **14:** 61–204.

KARLIN, S., and D. CARMELLI, 1975. Numerical studies on two loci selection models with general viabilities. Theor. Popul. Biol. **7:** 399–421.

KARLIN, S., and S. LESSARD, 1984 On the optimal sex ratio: a stability analysis based on a characterization for one locus multiallele viability models. J. Math. Biol. **20:** 15–38.

KARLIN, S., and J. MCGREGOR, 1972 Polymorphisms for genetic and ecological systems with weak coupling. Theor. Popul. Biol. **3:** 210–238.

KARLIN, S., and J. MCGREGOR, 1974 Towards a theory of the evolution of modifier genes. Theor. Popul. Biol. **5:** 59–103.

KINGMAN, J. F. C., 1980 *Mathematics of Genetic Diversity.* Society for Industrial and Applied Mathematics, Philadelphia.

LIBERMAN, U., and M. W. FELDMAN, 1986a A general reduction principle for genetic modifiers of recombination. Theor. Popul. Biol. **30:** 341–371.

LIBERMAN, U., and M. W. FELDMAN, 1986b Modifiers of mutation

rate: A general reduction principle. Theor. Popul. Biol. **30:** 125–142.

LIBERMAN, U., and M. W. FELDMAN, 1988 The reduction principle for genetic modifiers of the migration rate. In: *Mathematical Evolutionary Theory*, Edited by M. W. FELDMAN. Princeton University Press, Princeton, N. J. (to appear)

NEI, M., 1967 Modification of linkage intensity by natural selection. Genetics **57:** 625–641.

PROUT, T., J. BUNDGAARD and S. BRYANT, 1973 Population genetics of modifiers of meiotic drive. I. The solution of a special case and some general implications. Theor. Popul. Biol. **4:** 446–465.

STROBECK, C., J. MAYNARD SMITH and B. CHARLESWORTH, 1976 The effects of hitchhiking on a gene for recombination. Genetics **82:** 547–558.

TEAGUE, R., 1976 A result on the selection of recombination altering mechanisms. J. Theor. Biol. **59:** 25–32.

TEAGUE, R., 1977 A model of migration modification. Theor. Popul. Biol. **12:** 86–94.

THOMSON, G. J., and M. W. FELDMAN, 1974 population genetic of modifiers of drive. II. Linkage modification in the segregation distortion system. Theor. Popul. Biol. **5:** 155–162.

VON WETTSTEIN, D., S. W. RASMUSSEN and P. B. HOLM, 1984 The synaptonemal complex in genetic segregation. Annu. Rev. Genet. **18:** 331–413.

WRIGHT, S., 1964 Pleiotropy in the evolution of structural reduction and of dominance. Am. Nat. **98:** 65–69.

Communicating editor: M. TURELLI

APPENDIX: PROOFS OF THE RESULTS

Proof of Result 2

Here the new modifier allele M_a yields perfect transmission, hence

$$T(ai \leftarrow aj\,|\,bk) = \tilde{T}(ai \leftarrow aj\,|\,bk) = \delta_{ij} \quad \text{for all} \quad i, j, b, k.$$

Then (18) reduces to

$$\epsilon'_{ai} = \frac{1}{\hat{\bar{w}}}\,[(1 - r)\epsilon_{ai}\hat{w}_i + r\hat{v}_i \sum_j \epsilon_{aj} w_{ij}],$$

or, in vector form,

$$\epsilon' = \Omega_a \hat{\mathbf{D}}\epsilon = [(1 - r)\mathbf{I} + r\mathbf{Q}]\hat{\mathbf{D}}\epsilon, \tag{32}$$

where $\mathbf{D}$ is defined as before in (10) and $\mathbf{Q} = \|\hat{v}_i \frac{w_{ij}}{\hat{w}_j}\|_{i,j=1}^{n} = \mathbf{D}_1\mathbf{W}\mathbf{D}_2$, with $\mathbf{D}_1 = \text{diag}(\hat{\mathbf{v}})$, and $\mathbf{D}_2 = \text{diag}(\hat{w}_i^{-1})$.

Result 2a: We must show that the spectral radius

$$\rho(\Omega_a\hat{\mathbf{D}}) = \rho\{[(1 - r)\mathbf{I} + r\mathbf{Q}]\hat{\mathbf{D}}\} > 1, \quad \text{for all} \quad 0 \le r \le \tfrac{1}{2}.$$

Observe first that all eigenvalues of $\mathbf{Q}$ are real since both matrices $\mathbf{D}_1\mathbf{W}\mathbf{D}_2$ and $(\mathbf{D}_2\mathbf{D}_1)^{1/2}\,\mathbf{W}(\mathbf{D}_1\mathbf{D}_2)^{1/2}$ have the same eigenvalues and the latter, being symmetric, has all real eigenvalues. Since $\mathbf{Q}$ is stochastic its spectral radius is 1 (all eigenvalues of $\mathbf{Q}$ lie between -1 and 1) and therefore the eigenvalues of $(1 - r)\mathbf{I} + r\mathbf{Q}$ are confined to the positive interval $0 \le \lambda \le 1$ provided $0 \le r \le \tfrac{1}{2}$. The symmetric matrix

$$(\mathbf{D}_1\mathbf{D}_2)^{1/2}\,[(1 - r)(\mathbf{D}_1\mathbf{D}_2)^{-1} + r\mathbf{W}](\mathbf{D}_1\mathbf{D}_2)^{1/2}$$

is non-negative and symmetric and it has the same non-negative eigenvalues as the matrix $(1 - r)\mathbf{I} + r\mathbf{Q}$. Hence the former matrix is positive semidefinite.

This enables us to use Theorem 5.1 Corollary F.2 of KARLIN (1982), which states that for a positive diagonal matrix $\mathbf{D} = \text{diag}(D_i)$ and a matrix $\mathbf{M}$ that is symmetrizable to a positive semidefinite matrix (*i.e.*, $\mathbf{M} = \mathbf{E}_1\mathbf{P}\mathbf{E}_2$ where $\mathbf{E}_1$ and $\mathbf{E}_2$ are diagonal positive matrices and $\mathbf{P}$ is positive semidefinite), with $\mathbf{M}\xi = \xi$, $\mathbf{e}^T\mathbf{M} = \mathbf{e}^T$, and $\mathbf{e}^T\xi = 1$, we have

$$\rho(\mathbf{MD}) \ge \sum_i \xi_i D_i.$$

In the present case ξ is $\hat{\mathbf{D}}\hat{\mathbf{v}}$, since

$$\sum_j Q_{ij}(\hat{\mathbf{D}}\hat{\mathbf{v}})_j = \sum_j \hat{v}_i \frac{w_{ij}}{\hat{w}_j}\left(\frac{\hat{w}_j}{\hat{\bar{w}}}\,\hat{v}_j\right) = \hat{v}_i \sum_j \hat{v}_j \frac{w_{ij}}{\hat{\bar{w}}} = \hat{v}_i \frac{\hat{w}_i}{\hat{\bar{w}}} = (\hat{\mathbf{D}}\hat{\mathbf{v}})_i.$$

Thus

$$[(1 - r)\mathbf{I} + r\mathbf{Q}]\hat{\mathbf{D}}\hat{\mathbf{v}} = \hat{\mathbf{D}}\hat{\mathbf{v}},$$

and therefore

$$\rho(\Omega_a\hat{\mathbf{D}}) = \rho([(1 - r)\mathbf{I} + r\mathbf{Q}]\mathbf{D}) \ge \sum_i \hat{v}_i \frac{\hat{w}_i}{\hat{\bar{w}}}\frac{\hat{w}_i}{\hat{\bar{w}}} = \frac{1}{\hat{\bar{w}}^2}\sum_i \hat{v}_i(\hat{w}_i^2 - \hat{\bar{w}}^2) + 1 = 1 + \text{var}\left(\frac{\hat{w}_i}{\hat{\bar{w}}}\right) \ge 1. \tag{33}$$

Thus, when not all $\hat{w}_i$ are equal,

$$\rho(\Omega_a\hat{\mathbf{D}}) = \rho([(1 - r)\mathbf{I} + r\mathbf{Q}]\hat{\mathbf{D}}) > 1,$$

and the new modifier increases when rare.

Result 2b: if $r = 0$ then $\Omega_a = \mathbf{I}$, and so $\rho(\Omega_a\hat{\mathbf{D}}) = \rho(\hat{\mathbf{D}}) = \max_i(w_i/\hat{\bar{w}})$.

Result 2c: See (33).

Result 2d: This is a direct result from Theorem 5.2 of KARLIN (1982), which states that $\rho([(1 - \alpha)\mathbf{I} + \alpha\mathbf{M}]\mathbf{D})$ is a decreasing function of α, where $\mathbf{M}$ is a stochastic matrix and $\mathbf{D}$ a positive diagonal matrix.

Result 2e: This is an incidental implication of Results 2b and 2d and Definition 1, which give

$$\rho(\hat{\mathbf{D}}) = \max_i\left(\frac{\hat{w}_i}{\hat{\bar{w}}}\right) \ge 1 + \text{var}\left(\frac{\hat{w}_i}{\hat{\bar{w}}}\right).$$

Proof of Result 3

The recursion on the frequency of M_2 is, from (15) and (18)

$$\epsilon' = [(1 - m_{21})((1 - r)\mathbf{I} + r\mathbf{Q}) + m_{21}((1 - r)\hat{\mathbf{Y}}_1 + r\hat{\bar{\mathbf{Y}}}_1)]\hat{\mathbf{D}}\epsilon, \tag{34}$$

where $\mathbf{Q}$, $\hat{\mathbf{Y}}_1$, $\hat{\bar{\mathbf{Y}}}_1$ and $\hat{\mathbf{D}}$ are defined as in (25), (21), (22) and (15).

We know by the Perron-Frobenius theorem that since the strictly positive eigenvector $\hat{\mathbf{v}}$ in (20) has eigenvalue 1, $\rho([(1 - r)\hat{\mathbf{Y}}_1 + r\hat{\bar{\mathbf{Y}}}_1]\hat{\mathbf{D}}) = 1$. At $r = 0$, equation (34) reduces to

$$\epsilon' = [(1 - m_{21})\mathbf{I} + m_{21}\hat{\mathbf{Y}}_1]\hat{\mathbf{D}}\epsilon = \Omega_2\hat{\mathbf{D}}\epsilon.$$

We now use Theorem 5.2 of KARLIN (1982), which states:

> Let $\mathbf{M}$ be an irreducible stochastic matrix. Consider the family of stochastic matrices
>
> $$\mathbf{M}_\alpha = (1 - \alpha)\mathbf{I} + \alpha\mathbf{M}.$$
>
> Then for any positive diagonal matrix $\mathbf{D}$ ($\mathbf{D} \ne c\mathbf{I}$, $c > 0$), $\rho(\alpha) = \rho(\mathbf{M}_\alpha\mathbf{D})$ is strictly decreasing as α increases.

Therefore, $\rho(\Omega_2\hat{\mathbf{D}})$ is non-increasing in m_{21}. If $m_{21} = 1$, then $\rho(\Omega_2\hat{\mathbf{D}}) = 1$, so when $r = 0$, $\rho(\Omega_2\hat{\mathbf{D}}) \ge 1$ if $m_{21} < 1$ and $\rho(\Omega_2\hat{\mathbf{D}}) \le 1$ if $m_{21} > 1$.

Implicit in these inequalities on m_{21} is the scaling $m_{11} = 1$. With respect to the general situation, the previous inequalities on m_{21} are replaced by inequalities in m_{21}/m_{11}. From the theory of small parameters (KARLIN and MCGREGOR 1972), when $\hat{\mathbf{Y}}_1$ is irreducible, Result 3 holds also for some range of $r > 0$. The general case $0 < r < \tfrac{1}{2}$ is conjectured to hold also.

Proof of Result 4

Assume first that $p_i > 0$ for all i, and that $r = 0$. We make use of a result due to KARLIN (1982, p. 173) to the effect that if the matrix $\mathbf{S}$ has the representation

$$\mathbf{S} = \mathbf{E} + \mathbf{R}, \tag{35}$$

where $\mathbf{E} = \text{diag}(e_1, e_2, \cdots, e_n)$ and $\mathbf{R} = \| u_i v_j \|$ with $u_i v_j \geq 0$ for all i and j and $e_i < 1$ for all i then the spectral radius $\rho(\mathbf{S})$ of $\mathbf{S}$ is that value of λ for which

$$1 + \sum_{i=1}^{n} u_i v_i/(e_i - \lambda) = 0. \tag{36}$$

In the present case the matrix of external stability is $[(1 - m_{21})\mathbf{I} + m_{21}\mathbf{P}]\hat{\mathbf{D}}$, and comparing this to (35) we may use

$$\mathbf{E} = (1 - m_{21})\text{diag}(\hat{d}_i) \tag{37}$$

and

$$\mathbf{R} = m_{21} \| p_i \hat{d}_j \| \tag{38}$$

where $\hat{d}_i = \hat{w}_i/\hat{\bar{w}}$. Note that from (29)

$$\hat{v}_i[1 - (1 - m_{11})\hat{w}_i/\hat{\bar{w}}] = m_{11} p_i$$

so that $1 - (1 - m_{11})\hat{w}_i/\hat{\bar{w}} > 0$. For m_{21} sufficiently close to m_{11} obviously $(1 - m_{21})\hat{w}_i/\hat{\bar{w}} < 1$ also. Hence by KARLIN's result quoted above the required spectral radius λ_o is the unique solution of

$$U(\lambda) = 1 + m_{21} \sum_i p_i \hat{d}_i/[(1 - m_{21})\frac{\hat{w}_i}{\hat{\bar{w}}} - \lambda] = 0. \tag{39}$$

Clearly when $m_{21} = m_{11}$, $\lambda_o = 1$ and for m_{21} close to m_{11}, λ_o is close to 1. For λ_o near 1 we have, by Taylor's theorem,

$$U(1) = U(\lambda_o) + U'(1)(1 - \lambda_o) + O(1 - \lambda_o)^2$$

with $U(\lambda_o) = 0$. Hence, to this order of approximation,

$$\lambda_o = 1 - U(1)/U'(1). \tag{40}$$

We then write

$$\begin{aligned}(1 - m_{21})\frac{\hat{w}_i}{\hat{\bar{w}}} - 1 &= [(1 - m_{11})\frac{\hat{w}_i}{\hat{\bar{w}}} - 1] + (m_{11} - m_{21})\frac{\hat{w}_i}{\hat{\bar{w}}} \\ &= -\frac{m_{11} p_i}{\hat{v}_i} + (m_{11} - m_{21})\frac{\hat{w}_i}{\hat{\bar{w}}}\end{aligned} \tag{41}$$

by (29). We may then expand (39) at $\lambda = 1$ and neglect terms $O(m_{11} - m_{21})^2$ to obtain

$$U(1) = \frac{m_{11} - m_{21}}{m_{11}}\left\{1 - \frac{m_{21}}{m_{11}\hat{\bar{w}}^2}\sum_i \hat{v}_i^2 \hat{w}_i^2/p_i\right\}. \tag{42}$$

In the same way,

$$U'(1) = \frac{m_{21}}{m_{11}^2 \hat{\bar{w}}}\sum_i \frac{\hat{w}_i \hat{v}_i^2}{p_i} + O(m_{11} - m_{21}). \tag{43}$$

On substituting (42) and (43) into (40) we conclude that, except for terms $O(m_{11} - m_{21})^2$,

$$\lambda_o \cong 1 + \frac{m_{11} - m_{21}}{\hat{\bar{w}}\sum_i \hat{w}_i \hat{v}_i^2/p_i}\left[\sum_i \frac{\hat{v}_i^2 \hat{w}_i^2}{p_i} - \hat{\bar{w}}^2\right]. \tag{44}$$

From (29) $\hat{v}_i \hat{w}_i - p_i \hat{\bar{w}} = m_{11}^{-1}\hat{v}_i(\hat{w}_i - \hat{\bar{w}})$, so that

$$\sum_i \frac{\hat{v}_i^2 \hat{w}_i^2}{p_i} = m_{11}^{-2}\sum_i \frac{\hat{v}_i^2}{p_i}(\hat{w}_i - \hat{\bar{w}})^2 + \hat{\bar{w}}^2, \tag{45}$$

which upon substitution into (44) produces the conclusion of the stated Result 4.

It should be noted that if $p_h = 0$ for some h then with $r = 0$, for any m_{21} with $0 \leq m_{21} < 1$ it can be shown that

$$\rho(\hat{\mathbf{Y}}_2\hat{\mathbf{D}}) = 1 + \frac{m_{11} - m_{21}}{1 - m_{11}} = \frac{1 - m_{21}}{1 - m_{11}} = 1 + \left(1 - \frac{m_{21}}{m_{11}}\right)\hat{V},$$

where $\hat{V}$ is the selection potential at this equilibrium.