Proc. Natl. Acad. Sci. USA
Vol. 76, No. 1, pp. 541–545, January 1979
Population Biology

Principles of polymorphism and epistasis for multilocus systems

(linkage/nonepistasis/symmetric selection/multilocus Hardy–Weinberg equilibrium/gametic and allelic polymorphisms)

SAMUEL KARLIN

Department of Mathematics, Stanford University, Stanford, California 94305

Contributed by Samuel Karlin, September 11, 1978

ABSTRACT **The nature of stable equilibrium configurations is described for general nonepistatic and generalized symmetric viability regimes in multilocus systems under conditions of tight and loose linkage. The influence of epistasis and symmetry can be better understood in terms of these standards. A dichotomy in the nature of stable polymorphisms emerges. More recombination, bisexuality, and multideme interactions facilitate the establishment of central type polymorphisms.**

This paper reports results providing some principles of multigene selection–linkage interactions. There are three principal components inherent to multilocus selection linkage structures: (*i*) The recombination process; (*ii*) the selection regime; (*iii*) the gamete frequency configurations and their description and classifications.

Some natural n loci recombination distributions (sets of recombination rates) are described in ref. 1. For n loci there can occur up to $2^{n-1} - 1$ basic recombination outcomes. Two important recombination distributions are those of *no recombination* $\mathcal{R}^{(0)}$ and *free recombination* $\mathcal{R}^{(f)}$. A mixture of the recombination distributions of $\mathcal{R}^{(0)}$ and $\mathcal{R}^{(f)}$ represents n loci divided into linkage groups or gene clusters.

The comparisons of two recombination distributions is unambiguous in the case of two loci because a single recombination rate is present. With more loci involving a vector of recombination rates it is not, *a priori*, clear how to order two recombination distributions. A problem of interest concerns the quantification of "natural" and "more recombination" relationships. Some approaches to this matter are elaborated in ref. 1. In this paper we focus on selection regimes and characterizations of their stable gamete configurations.

2. Parameterizations and classifications of multilocus selection regimes

Two important classes of n loci selection regimes that are substantially tractable comprise the *generalized nonepistatic selection mode* and the *generalized symmetric viability regime*. To assess the significance of epistasis it is vital to properly delimit the scope of nonepistasis. In the same way, to understand the consequences of asymmetry in selection expression it is essential to characterize forms and levels of symmetry. The generalized nonepistatic selection regime encompasses combinations of multiplicative and neutral viability effects distributed across loci. The generalized symmetric selection regime is an extension of the notions of symmetric over and underdominance at a single locus. In the latter model, fitness depends on which loci are homozygous or heterozygous and otherwise is not influenced by the allelic composition at each locus.

Generalized Nonepistatic Selection. The concept of nonepistasis involves selection coefficients at the separate loci. For purposes of illustration we begin with three loci. Let the intrinsic fitness matrix for the kth locus with m_k alleles be $W_k = \|w_{ij}^{(k)}\|_{i,j=1}^{m_k}$ specifying components of the fitnesses associated with the marginal genotypes $A_i^{(k)} A_j^{(k)}$. These matrices combine in eight ways (2^n for n loci) to generate the basic selection regime underlying generalized nonepistatic selection. We present these matrices in tabular form (Table 1) indicating the fitness associated with the genotype $A_i^{(1)} A_j^{(2)} A_k^{(3)} / A_m^{(1)} A_p^{(2)} A_q^{(3)}$. In this notation generalized nonepistasis connotes a fitness matrix combining linearly the types in Table 1 to the form

Table 1. Fitness matrix

Matrix	Entry	Description
$W(1,1,1)$	$w_{im}^{(1)} w_{jp}^{(2)} w_{kq}^{(3)}$	Multiplicative factors from all loci
$W(1,1,0)$	$w_{im}^{(1)} w_{jp}^{(2)}$	Multiplicative selection at loci 1, 2
$W(1,0,1)$	$w_{im}^{(1)} w_{kq}^{(3)}$	Multiplicative selection at loci 1, 3
$W(0,1,1)$	$w_{jp}^{(2)} w_{kq}^{(3)}$	Multiplicative selection at loci 2, 3
$W(1,0,0)$	$w_{im}^{(1)}$	Selection acting only at locus 1
$W(0,1,0)$	$w_{jp}^{(2)}$	Selection acting only at locus 2
$W(0,0,1)$	$w_{kq}^{(3)}$	Selection acting only at locus 3
$W(0,0,0)$	1	Neutral

$$\begin{aligned}\Gamma = {} & c(1,1,1)W(1,1,1) + c(1,1,0)W(1,1,0) \\ & + c(1,0,1)W(1,0,1) + c(0,1,1)W(0,1,1) \\ & + c(1,0,0)W(1,0,0) + c(0,1,0)W(0,1,0) \\ & + c(0,0,1)W(0,0,1) + c(0,0,0)W(0,0,0). \quad [2.1]\end{aligned}$$

The matrix $W(1,1,1)$ is the Kronecker product matrix $W_1 \otimes W_2 \otimes W_3$ and $W(1,1,0)$ is the Kronecker product $W_1 \otimes W_2 \otimes E_3$, in which E_3 is the matrix of order $m_3 \times m_3$ exhibiting only unit elements, $W(1,0,1) = W_1 \otimes E_2 \otimes W_3$, etc.

The classical multiplicative nonepistatic form has the specification $c(1,1,1) = 1$ with the other cs $= 0$. Classical additive nonepistasis follows the prescription $c(1,0,0) = c(0,1,0) = c(0,0,1) = 1$, and the other cs $= 0$. The choice $c(1,1,0) = c(0,1,1) = 1$ with the remaining cs $= 0$ entails multiplicative nonepistasis between the first two loci, but additive nonepistasis between the gene complex of the first two loci and the third locus.

The existence of epistasis for the three-locus model will be understood such that the array of fitness values cannot be represented in the form of Eq. **2.1.** The extent of epistasis may be measured by an appropriate "distance" of the fitness matrix to the class of Eq. **2.1.**

A generalized nonepistatic n-locus selection regime has the following structure: (*i*) each locus has an intrinsic fitness matrix that specifies the fitness values of the marginal genotypes; (*ii*) selection may or may not act at each locus; (*iii*) the fitness of an n-locus genotype is determined as multiplicative contributions of the marginal viabilities or neutral values of the constituent loci genotypes. Formally, we prescribe for each locus k an intrinsic fitness matrix W_k admitting m_k possible alleles. Let E_k be the neutral fitness matrix of order m_k of all unit entries. We propose for an n-*locus generalized nonepistasis regime* a fitness matrix of order $K = \prod_{k=1}^{n} m_k$ of the form

$$\Gamma = \sum_{\eta} c(\eta)\,(W_1^{(\eta 1)} \otimes W_2^{(\eta 2)} \otimes \ldots \otimes W_n^{(\eta n)}), \qquad [2.2]$$

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

Abbreviation: H-W, Hardy–Weinberg.

in which the sum is extended over *all* n-tuples $\eta = (\eta_1, \eta_2, \ldots, \eta_n)$, $\eta_i = 0$ or 1, subject to the *special* convention

$$W_k^{(1)} = W_k, W_k^{(0)} = E_k. \quad [2.3]$$

Accordingly, for each η, the fitness matrix $W(\eta) = (W_1^{(\eta_1)} \otimes W_2^{(\eta_2)} \otimes \ldots \otimes W_n^{(\eta_n)})$ prescribes a standard multiplicative nonepistatic fitness regime with selection forces operating at those component loci where $\eta_k = 1$ while the other loci act neutrally. The collection of fitness matrices, $\{W(\eta)\}$, joined as in Eq. 2.2 induce a nonepistatic regime based on the intrinsic fitness matrices $\{W_k\}$. The coefficients $c(\eta)$ contrast nonepistatic selection differentials and neutral effects distributed among groupings of loci.

Generalized Multilocus Hardy–Weinberg (H-W) Equilibria. Assume the existence of a polymorphic equilibrium $(\hat{x}_1^{(k)}, \hat{x}_2^{(k)}, \ldots, \hat{x}_{m_k}^{(k)}) = \hat{\mathbf{x}}_k$ for the one locus m_k-allele system having the intrinsic fitness matrix W_k. We may construct an n-locus gamete frequency array by multiplying the marginal frequencies of the constituent alleles. Accordingly, the frequency of the haplotype $A_i^{(1)} A_j^{(2)} \ldots A_l^{(n)}$ is $\hat{x}_i^{(1)} \hat{x}_j^{(2)} \ldots \hat{x}_l^{(n)}$.

This haplotype frequency array is an equilibrium of the fitness model of Eq. 2.2 independent of the recombination distribution and the coefficients $c(\eta)$. A population state composed by multiplying the marginal loci genotype frequencies is called a *multilocus H-W configuration*. In the two-locus case a H-W state signifies *linkage equilibrium*, and in the multilocus case zero measures of association of all orders.

Where $\hat{\mathbf{x}}_k$ is a polymorphic equilibrium for the separate loci fitness matrices W_k, $k = 1, 2, \ldots, n$, respectively, then the H-W equilibrium $\hat{\mathbf{x}} = \hat{\mathbf{x}}_1 \otimes \hat{\mathbf{x}}_2 \otimes \ldots \otimes \hat{\mathbf{x}}_n$ of components $\hat{x}_i^{(1)} \hat{x}_j^{(2)} \ldots \hat{x}_l^{(n)}$ is referred to as a *H-W polymorphic equilibrium*.

Generalized Symmetric Viability Regimes. These incorporate far-ranging extensions of the two- and three-locus, two-allele symmetric fitness models, (2, 3). It is helpful to exemplify first the case of three loci involving $\{A,a\}$, $\{B,b\}$, $\{C,c\}$ at the respective loci. The fitness values are as follows:

> α, in which all loci are homozygous; β_i, in which locus number i is heterozygous while the other two loci are homozygous; δ_i, in which locus i is homozygous while the other two loci are heterozygous; γ, in which all three loci are heterozygous. [2.4]

The construction of the n-locus model goes as follows. We single out for locus k the matrices I_k = *identity matrix of order* $m_k \times m_k$ and $J_k = \|u_{ij}\|_1^{m_k}$, $u_{ij} = 1$ for $i \neq j$ and 0 when $i = j$. The matrix J_k is a one-locus viability matrix in which all heterozygotes carry the same fitness value 1 while all homozygotes are lethal. The *central frequency state* $\mathbf{x}_k^* = (1/m_k, 1/m_k, \ldots, 1/m_k)$ is a stable and unstable equilibrium for J_k and I_k, respectively.

Parallel to Eq. 2.2 we form the extended fitness matrix

$$S = \sum_{\eta} \gamma(\eta)(J_1^{\eta_1} \otimes J_2^{\eta_2} \otimes \ldots \otimes J_n^{\eta_n}) \quad [2.5]$$

subject to the convention (different from Eq. 2.3) *that*

$$J_k^1 = J_k, J_k^0 = I_k \quad [2.6]$$

and $\gamma(\eta)$ are nonnegative weights. The summand in Eq. 2.5 of index η singles out the combination of loci positions identified by the unit components in η that yield a relative contribution $\gamma(\eta)$ to fitness. Where $\gamma(\eta)$ depends only on the number of heterozygous loci and not the location of these loci, so that $\gamma(\eta) = \gamma_r$ for $|\eta| = \sum_{i=1}^n \eta_i = r$, then Eq. 2.5 reduces to

$$R = \sum_{r=0}^{n} \gamma_r P_r \text{ with } P_r = \sum_{|\eta|=r} J_1^{\eta_1} \otimes J_2^{\eta_2} \otimes \ldots \otimes J_n^{\eta_n}. \quad [2.7]$$

Thus, the fitness regime R confers fitness γ_r if a genotype involves r heterozygous loci independent of their locations. The model 2.5 exhibits the *central* H-W *equilibrium state* $\mathbf{c}^* = \mathbf{x}_1^* \otimes \mathbf{x}_2^* \otimes \ldots \otimes \mathbf{x}_n^*$ *in which all gametes share equal frequency.*

3. Stability conditions for central type equilibria

Under the assumption that each marginal fitness matrix W_k, $k = 1, 2, \ldots, n$ is "overdominant," (in the sense that the one-locus viability matrix W_k possesses an interior stable polymorphism), the generalized nonepistatic selection regime admits a *unique multilocus H-W polymorphism*. For the generalized symmetric selection regime there exists the *central polymorphism* $\mathbf{c}^*$. The stability properties of these "central" equilibria are summarized in Table 2.

Analytic conditions for stability of the H-W polymorphic equilibrium under a generalized nonepistatic selection regime and any specification of the recombination distribution are given in refs. 4 and 6.

It is useful to record some analytic conditions for stability of the central polymorphism in the case of no and free recombi-

Table 2. Existence and stability of the "central" equilibrium

Generalized nonepistatic selection regime, 2.2	Generalized symmetric selection regime, 2.5
The *existence* of a H-W polymorphic equilibrium occurs for any level of recombination provided each intrinsic fitness matrix admits a polymorphism (4).	The central polymorphism $\mathbf{c}^*$ exists for any level of recombination (5).
The H-W polymorphism in the presence of absolute linkage is never stable.	The central equilibrium for some ranges of the selection coefficients can be stable for absolute linkage; cf. Eq. 3.1.
The H-W polymorphism for tight recombination is generally *not* stable; the exception is pure additive nonepistasis (4).	When the central polymorphism is stable for absolute linkage, then it is stable for any level of recombination.
The H-W polymorphism under free recombination or loose linkage is *always* stable and apparently uniquely stable (4).	When increased heterozygosity enhances fitness, then the central polymorphism is stable for moderate to loose linkage (5) and uniquely stable for free recombination; see Eq. 3.4.
If for a given level of recombination the H-W or central equilibrium is *stable*, then it is stable for "more recombination" (4).	
Bisexuality compared to a corresponding monoecious model facilitates the stability of the H-W and central polymorphism (6).	
Effects of migration. Stability of the H-W or central polymorphism in each deme implies stability in the system with any form of migration. An increased level of migration entails more opportunities for stability of the "central" polymorphism (7).	
An increased number of loci generally depresses the stability of the H-W polymorphism (4).	An increased number of loci facilitates the stability of the central equilibrium (8).
Increased allelism generally diminishes the stability opportunities of the H-W polymorphism.	The effects of increased allelism depend on the selection values and the degree of recombination (8).
Increasing the neutral component in Eq. 2.2 does not necessarily decelerate the rate of convergence to the H-W polymorphism (4).	The introduction of more neutral effects generally leaves unchanged the property of stability as against instability of the central equilibrium (8).

nation. We find the *central equilibrium* $\mathbf{c}^*$ *for the selection regime* **2.5** *stable under absolute linkage provided*

$$\sum_{\eta} \gamma(\eta) \Big(\prod_{\nu=1}^{\eta} (m_\nu - 1)^{\eta_\nu} \Big) \prod_{\mu=1}^{n} \left(\frac{-1}{(m_\mu - 1)} \right)^{\eta_\mu \theta_\mu} < 0,$$

for all $\theta_i = 0$ *or* 1 *and* θ_i *not all zero.* Where the number of alleles per locus is constant $m_\nu = m$ and the fitness value of a genotype depends on the number of heterozygous loci (model **2.7**), the stability conditions reduce to the inequalities

$$\sum_{k,j=0}^{n} \gamma_k (m-1)^{k-j} (-1)^j \binom{l}{j} \binom{n-l}{k-j} < 0, \; l = 1, 2, \ldots, n. \qquad [3.1]$$

With free recombination, the stability of the central polymorphism for the model **2.5** *requires the single condition*

$$\sum_{k=0}^{n-1} (m-1)^k \binom{n-1}{k} (\gamma_k - \gamma_{k+1}) < 0. \qquad [3.2]$$

Manifestly, if $\gamma_0 < \gamma_1 < \ldots < \gamma_n$ *then the central polymorphism is stable under free recombination.*

By virtue of Eq. **3.2** we can infer that the influence of multiple allelism generally diminishes the amount of recombination needed to make the central equilibrium stable. There are exceptions for a fitness pattern satisying $\gamma_0 < \gamma_1 < \ldots < \gamma_{k-1} < \gamma_k > \gamma_{k+1} > \ldots > \gamma_n$, *for some* k < n. *In this situation stability of the central polymorphism is only possible if the number of alleles* m *per locus is not too large. On the other hand, if* $\gamma_0 > \gamma_1 > \ldots > \gamma_k < \gamma_{k+1} < \ldots < \gamma_n$ *for some* k < n, *then increasing the extent of allelism facilitates the maintenance of the central polymorphism.*

4. The symmetric model for tight linkage

The stable realizations under small recombination rates can be quite different from loose linkage. The following concepts of one-locus multiallelic models are germane in the analysis of multilocus systems for the case of tight linkage.

Allelic against Gametic Polymorphism. An equilibrium under absolute linkage is called an *allelic polymorphism* if every allele at each locus occurs with positive frequency. A *gametic polymorphism* constitutes an equilibrium state involving *every* gamete with positive frequency. The stable equilibrium structures for absolute linkage can be delineated as follows: (*i*) *every stable equilibrium is an allelic polymorphism;* (*ii*) *at least one (but not every) stable equilibrium is an allelic polymorphism;* (*iii*) *all stable equilibria entail fixation at one or more loci.*

The theory of small parameters (9) tells us that for small recombination rates a stable allelic polymorphism becomes a stable gamete polymorphism approximately near the corresponding equilibrium state existing for absolute linkage. With category *ii* the equilibrium possibilities involve both stable polymorphisms and population states having one or more loci fixed and the evolution then depends finely on initial conditions, founder effects, or other environmental factors.

To illustrate, we restrict attention to the case of two alleles per locus, and to the aggregate heterozygosity fitness structure of Eq. **2.7.** A case of interest concerns the nature of the stable gamete frequency realizations under the assumption

$$\gamma_0 < \gamma_1 < \ldots < \gamma_n \qquad [4.1]$$

in which *increasing heterozygosity enhances the fitness value.* In this situation under *free recombination the central polymorphism is stable;* cf. Eq. **3.2.** However, the evolutionary consequences for tight linkage are more sensitive to the order and form of increase of $\{\gamma_k\}$. Generally, the property, Eq. **4.1,** does not ensure stability of the central equilibrium without recombination. Because more aggregate heterozygosity improves fitness we would expect more opportunities of polymorphism at all levels of recombination. In this vein the following result prescribes the nature of polymorphism for tight linkage.

Result I: If the fitness values in Model **2.7** increase for genotypes with more heterozygous loci, then every stable equilibrium for sufficiently tight linkage is a gametic polymorphism. Equivalently, without positive recombination, every stable equilibrium constitutes an allelic polymorphism.

The numerical analysis of Table 3 was performed for 200 selection determinations of the fitness matrix **2.7,** each $\{\gamma_k\}$ independently uniformly distributed on [0,1].

Thus, with a random set of fitness values in Eq. **2.7,** the *chance of at least one polymorphism for tight linkage increases with more loci involved.*

Symmetric Type Equilibria. To a large extent the relevant equilibrium frequency states inherit much of the symmetry (with respect to allelic and loci substitutions) of the viability structure. The central polymorphism $\mathbf{c}^*$ is symmetric in the strongest sense. Another familiar frequency configuration pertains to *complementary gamete types* such as *ABC-abc, ABc-abC,* etc., in a three-locus, two-allele context. For *n* loci with two alleles per locus, a pair of gametes is called *complementary* if each allele occurs in one of the *two gametes.* The third class of natural symmetric equilibrium arrays in the *n*-locus, two-allele model, consist of half of the gametes, the *half central symmetric equilibria.* (The three-locus representatives consist of the gamete arrays *ABC-Abc-aBc-abC* and *abc-aBC-AbC-ABc.*) A half central equilibrium array can be characterized as follows. This equilibrium involves half the number of possible gametes $(2^n - 1)$ of equal frequencies in which every combination of l, $l = 1, 2, \ldots, n-1$ alleles, one per locus, occurs in 2^{n-l-1} gametes of this array. An example of a half central symmetric equilibrium with four loci consists of the gametes {*ABCD, aBCd, ABcd, aBcD, AbCd, abCD, AbcD, abcd*} each with frequency 1/8. There exist further classes of *mixed symmetric type* equilibria that can be stable under absolute linkage, (8). Also, nonsymmetric stable equilibrium types can occur with $n \geq 3$ loci for a restricted range of the fitness space.

The precise conditions for local stability of the indicated symmetric equilibrium types for the selection regime (**2.5** *n* loci, $m = 2$ alleles per locus) are as follows:

Result II: The central equilibrium is stable provided **3.1** holds. The half central equilibria are stable provided $\Sigma_{k=0}^{n} \gamma_k (-1)^k \binom{n}{k} > 0$ and $\Sigma_{k=\text{even}} \gamma_k \Sigma_{j=0}^{n} (-1)^j \binom{l}{j} \binom{n-l}{k-j} < 0$, $l = 1, 2, \ldots, n-1$, hold. The complementary equilibria are stable under the conditions $\gamma_n > \gamma_0$ and $\gamma_n + \gamma_0 > \gamma_{n-k} + \gamma_k$ for $k = 1, 2, \ldots, n-1$.

Stability of the Symmetric Equilibria in Terms of Properties of a Fitness Function $\varphi(x)$. Having in mind the prospect of a multilocus theory that is independent of the number of loci involved ("a large number of loci"), we consider fitness as a function of the proportion of heterozygous loci in the manner that in the fitness regime of Eq. **2.7** we take $\gamma_k = \varphi(k/n)$, $k = 1, 2, \ldots, n$ in which the *fitness function* $\varphi(x)$ is defined for $0 \leq x \leq 1$. We are mostly interested in the connection between

Table 3. Numerical analysis

The nature of stable polymorphisms for tight linkage	No. of loci			
	2	3	4	5
Category (*i*) involving only allelic or gametic polymorphisms	107	104	91	82
Category (*ii*) in which allelic polymorphisms and stable boundary configurations coexist	107	122	124	124
Category (*iii*) having some loci nonsegregating	93	78	76	76

properties of the fitness function and stability of the symmetric type equilibria. Important choices of the function φ are

$$\varphi(x) = e^{\lambda x},\ \lambda > 0;\ \varphi(x) = (x + c)^{\alpha},\ \alpha > 0,\ c \geq 0;$$
$$\varphi(x) = (x + a)^{\alpha}(1 - x)^{\beta},\ \alpha, \beta > 0. \quad [4.2]$$

The first case of **4.2** means that the fitness pattern is *both symmetric and multiplicative* with respect to loci and allele viability effects. The third fitness form of Eq. **4.2** has the optimum phenotype at an intermediate level of heterozygosity.

The family of fitness regimes engendered by the (power) fitness function $\varphi(x) = (x + c)^{\alpha}$ (α and c positive) offers a broad class of selection contrasts. The marginal change in fitness due to an additional heterozygous locus is increasing for $\alpha > 1$ but decreasing for $0 < \alpha < 1$. The selection regime with smaller α induces a stronger epistatic selection interaction in comparison to multiplicative selection which comes out for $\alpha \to \infty$. A natural inquiry concerns what kind of regular pattern of fitness function allows more opportunities for a stable central equilibrium with tight linkage. The following result obtains.

Result III: Any convex fitness function cannot maintain the central equilibrium stable with tight linkage. On the other hand, if $0 < \alpha \leq 1$, and $\varphi(x) = (x + c)^{\alpha}$, $c \geq 0$, the central equilibrium is stable for any level of recombination.

That some extent of monotonicity of the fitness function $\varphi(x)$ is consonant with enhanced stability of a central polymorphism is anticipated. What is surprising is that for tight linkage "*increasing concavity*" is associated with a stable polymorphism, while "*increasing convexity*" is inconsistent with this attribute. Further analysis of the consequences attendant to the fitness functions **4.2** will be published elsewhere.

5. The generalized nonepistatic selection regime with tight linkage

We deal with n loci, two alleles per locus having the marginal fitness values at locus k, α_k = fitness (A_kA_k), β_k = fitness (a_ka_k), and the heterozygote fitness (A_ka_k) normalized to 1. We have proved (10), for the multiplicative nonepistatic selection regime.

Result IV: The only stable equilibria when $\alpha_k = \beta_k$, $k = 1, \ldots, n$ for absolute linkage and pure multiplicative selection consists of complementary gamete pairs.

In the following we contrast the fact of *Result IV* with the nature of the stable equilibrium possibilities for a pure multiplicative viability regime, which is *symmetric across loci*, but asymmetric with respect to *intralocus homozygote types*. Each locus is assumed overdominant, but one homozygous form confers greater fitness relative to the other homozygote: namely, α = fitness $(A_iA_i) > \beta$ = fitness (a_ia_i), $i = 1, 2, \ldots, n$. We refer to A_i as a "*good*" allele relative to a_i. For ease of exposition, we display the results for the three-locus case in Table 4. Taking account of the loci symmetry, we describe sets of gametes by integer arrays in which we specify the numbers of good alleles per gamete. Thus, {3,0} stands for the complementary pair {*ABC*, *abc*}; [{2,1} − (comp)] signifies any of the complementary gamete pairs {*ABc*, *abA*}, {*AbC*, *aBc*}; and {*aBC*, *Abc*}; and class [{3,2,1} − (comp)] can refer to {*ABC*, *ABc*, *abC*}, {*ABC*, *AbC*, *aBc*}, . . . , in which {2,1} of {3,2,1} consists of complementary gametes. Unambiguously, {2,2,2} = {*ABc*, *AbC*, *aBC*}, {3,2,2,2} = {*ABC*, *ABc*, *AbC*, *aBC*}.

Table 4. Stable equilibrium arrays depending on the parameters α and β, $0 \leq \beta < \alpha \leq 1$

Conditions	Stable configuration types
$f(\alpha,\beta) > 0$	All complementary pairings are stable
$f(\alpha,\beta) < 0, g(\alpha,\beta) > 0, h(\alpha,\beta) < 0$	[3,0], [3,2,1 − comp]
$g(\alpha,\beta) < 0, k(\alpha,\beta) > 0, f(\alpha,\beta) < 0$	[3,0], [3,2,1 − comp], [2,2,2]
$Q(\alpha) > 0, g(\alpha,\beta) > 0, h(\alpha,\beta) > 0$	[3,0], [3,2,1 − comp], [2,2,2,1]
$Q(\alpha) > 0, k(\alpha,\beta) < 0$	[3,0], [3,2,1 − comp], [3,2,2,2]

$f(\alpha,\beta) = 1 - \alpha^2 - \alpha(1 - \beta)(1 + \alpha\beta)$, $g(\alpha,\beta) = (2\alpha - \alpha^2)\beta + 1 - 2\alpha$
$h(\alpha,\beta) = (3 + \sqrt{5})\alpha - 1 - \sqrt{5} - 2\beta\alpha$, $k(\alpha,\beta) = \alpha\beta + 2 - 3\alpha$, $Q(\alpha) = \alpha^2 + 5\alpha - 5$.

It is clear that where α and β differ sharply or are close to 1, an abundance of noncomplementary stable equilibrium types are relevant, contrary to the dictum of ref. 11 predicting exclusive high complementary equilibrium realizations. When *asymmetry* in both allelic and loci effects occurs even for multiplicative nonepistasis, the most common stable equilibrium involves, for tight recombination, an intermediate number of gamete types (in excess of two), while for moderate to loose linkage a central H-W equilibrium takes over; cf. *Section 3*. With two loci some aspects of the influence of asymmetry in viability expression are expounded in ref. 10.

We summarize the results of *Sections 4* and *5* concerning tight linkage in Table 5.

6. Summary and discussion

We need a standard in order to understand epistasis and asymmetry. We described two classes of nonepistatic and symmetric multilocus selection regimes and reported extensive results on the attendant stable equilibrium configurations. Tables 2 and 5 summarize the salient qualitative properties of the equilibrium realizations under conditions of loose and tight linkage, respectively.

Some general principles on polymorphism (not necessarily universally valid, but of wide scope) are suggested by our results, which we now discuss.

Principle I: If there exists a stable polymorphic equilibrium in the presence of no recombination, then there exists a stable polymorphism for any positive recombination.

In line with *Principle I*, an essential dichotomy in the nature of stable polymorphisms emerges: (*i*) The selection interactions are predominant, establishing the polymorphism, while the

Table 5. Nature of the stable equilibrium types for absolute linkage (and consequences for tight linkage)

Generalized nonepistatic selection, Eq. 2.2	Generalized symmetric selection, Eq. 2.7
With general marginal selection components, no symmetric equilibrium states exist.	A whole spectrum of symmetric equilibrium types always exists, highlighting complementary, half-central, central, and mixed symmetric forms.
With marginal strong overdominance, *every stable* equilibrium constitutes an *allelic polymorphism.*	Where fitness increases with increasing heterozygosity, then *only polymorphisms are stable* for sufficiently tight linkage.
Stability of *complementary equilibrium arrays* is likely only where fitness is invariant under allelic substitutions while unlikely with loci or allelic asymmetry in fitness expression.	Stability of *complementary equilibrium arrays* is likely under conditions of convexity of the fitness function (see *Result III*) (5, 9).
The H-W equilibrium is essentially never stable; see Table 2 (4).	The central equilibrium is only stable for a strongly concave fitness function (see *Result III*).
Stable equilibria generally involve a few to an intermediate number of gametes.	Mostly mixed symmetric types constitute the stable equilibrium arrays involving an intermediate number of gametes.

recombination mechanism exerts minor effects. (*ii*) Selection forces alone without recombination would lead to the elimination of certain gamete types; but with some recombination the full complement of gamete types are sustained. The characteristics of the polymorphic outcomes of *i* and *ii* markedly differ. For situation *i*, usually a single rather central, globally stable polymorphism is predominant for each set of recombination rates. In case *ii*, a multiplicity of stable polymorphisms are feasible for tight linkage of the form that the gamete arrays divide into two groups such that the frequencies in one group are rare, while the gamete frequencies in the other group are more discernible. Such polymorphic states generally manifest second or higher order moderate to strong linkage disequilibrium values. With moderate-to-loose linkage these polymorphisms are usually not maintained.

We can refine *Principle I* as follows:

Principle II. Consider the class of one locus multiallelic viability regimes that induce a stable polymorphism. The population system obtained by superposition on such viability regimes of one or several of the following mechanisms: (*i*) recombination, (*ii*) bisexuality, (*iii*) multidemic interactions (e.g., migration or population subdivision) usually evolves to a globally stable polymorphism.

For a selection structure in which increasing heterozygosity enhances the fitness, every stable equilibrium (with absolute linkage) constitutes an allelic polymorphism, and concomitantly with sufficiently tight linkage only *bona fide* polymorphisms occur (cf. *Result I*). In this light, large apparent allelic polymorphism as reported for xanthine dehydrogenase (12) and for esterase V (13) may correspond to clusters of tightly linked genes with a selection mechanism of the approximate structure **2.5**, in which $\gamma_0 < \gamma_1 < \ldots < \gamma_n$. Where the genomic complex consists of clusters of many tightly linked genes subject to small nonzero recombination (e.g., intragenic recombination), the expectation is that a few to several gametes occur with moderate frequency, while most gametes occur in trace frequency.

Another qualitative inference based on the study of the generalized symmetric selection regime **2.5** is as follows. Where the existence of appropriate heterozygosity at a few loci contributes significantly to fitness, then a "central frequency array" is stable irrespective of the extent of recombination. In particular, if a trait manifests a significant number of segregating types, each occurring with reasonable frequency and consistently observed in population samples at different locations, and where further cytological evidence or pedigree analysis attests to the property that the genes involved are mostly tightly linked, then an approximation to a selection regime of the form **2.5** may provide a basis for the observed variability.

The four-locus *HLA* complex in man covering about 1 map unit manifests approximately the frequency pattern just described. A mechanism consistent with the analytic theory proposes that a few loci may be decisive to fitness and carry numerous other loci (not necessarily closely linked) polymorphic. This inference conforms with the existence of the immune response gene documented in the parallel H2 system in mice. A similar proposal may apply for the variegated immunoglobulin variability involving clusters of tightly linked genes.

In contrast, consider a phenomenon in which a partial set of gamete types is mostly observed with trace amounts of other gamete possibilities. Suppose also that in different sampled populations the frequency data have the principal gametes often disparate across the population range. Where the underlying genes are recognized to be relatively tightly linked, then an explanation along the lines of *Result IV* states that probably *many* genes are needed to assure proper fitness (5).

Principle I states that if a stable polymorphism exists without recombination, then the recombination mechanism can displace it but not destroy it. A further proposition in this vein asserts that where a central type polymorphism is stable at *some* recombination level, then it persists stably when "more recombination" is in force (4, 8). This property does not apply for the near boundary equilibrium configuration in which such polymorphisms are generally not maintained under conditions of moderate to loose recombination rates.

It is established in refs. 6 and 8 that the redistribution of selection differentials from a monoecious population to a corresponding dioecious context, retaining the same recombination–segregation mechanisms, facilitates the establishment of a stable central H-W polymorphic equilibrium. Concomitantly, the contingencies of other polymorphic type equilibrium, in which only a partial set of haplotypes predominates, tend to be reduced with the advent of separate sexes as against a corresponding monoecious population.

What is the evolutionary advantage of cases of low recombination in the heterogametic sex with higher rates in the homogametic sex? The problem is considered enigmatic. A recent review of contrasts in rates of recombination between the sexes is given in ref. 14. Two aspects of increased chiasmata are widely recognized. The first is the need for chiasmata to ensure proper disjunction of the two homologous chromosomes at meiosis. The second is the evolutionary flexibility attendant to recombination events. It is of interest to consider evolutionary advantages in sex-dependent recombination rates in terms of the nature of polymorphism for dioecious as against monoecious selection recombination structures.

Multilocus theory (6) suggests that with loose linkage there is an increased tendency for the existence of polymorphism, entailing a gamete frequency configuration in near linkage equilibrium of all orders. With tight linkage, in contrast, the nature of polymorphism when extant involves a partial set of gametes manifesting some strong measures of genic associations. It appears that distinct sex differences in recombination frequencies can more easily accommodate the coexistence of both these types of polymorphisms, and which is established depends on initial conditions and environmental factors.

This work was supported in part by National Institutes of Health Grant GM 10452-15 and National Science Foundation Grant MCS76-80624-A01.

1. Karlin, S. & Liberman, U. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 6332–6336.
2. Lewontin, R. & Kojima, K. (1960) *Evolution* **14**, 458–472.
3. Feldman, M. W., Franklin, I. R. & Thomson, G. (1974) *Genetics* **76**, 135–162.
4. Karlin, S. & Liberman, U. (1979) *Genetics*, in press.
5. Karlin, S. (1977) in *Proceedings of the International Conference on Quantitative Genetics*, eds. Kempthorne, O. & Pollak, E. (Iowa State Univ. Press, Ames, IA), pp. 207–226.
6. Karlin, S. & Liberman, U. (1979) *Genetics*, in press.
7. Karlin, S. & Campbell, R. B. (1978) *Genet. Research*, in press.
8. Karlin, S. & Avni, H. (1979) *Theor. Popul. Biol.*, in press.
9. Karlin, S. & McGregor, J. (1972) *Theor. Popul. Biol.* **3**, 186–209.
10. Karlin, S. & Feldman, M. W. (1978) *Genetics* **90**, 813–825.
11. Franklin, I. R. & Lewontin, R. C. (1970) *Genetics* **65**, 701–734.
12. Singh, K., Lewontin, R. C. & Felton, A. (1976) *Genetics* **84**, 609–629.
13. Coyne, J. (1976) *Genetics* **84**, 593–607.
14. Callan, H. G. & Perry, P. E. (1977) *Philos. Trans. R. Soc. London* **277**, 227–233.