Published in final edited form as:
Theor Popul Biol. 2008 March ; 73(2): 307–316. doi:10.1016/j.tpb.2007.11.010.

On the Evolution of Epistasis III:

The Haploid Case with Mutation

Uri Liberman[1] and **Marcus Feldman**[2,*]

1*School of Mathematical Sciences, Tel Aviv University, Tel Aviv, Israel 69978, liberman@post.tau.ac.il*

2*Department of Biological Sciences, Stanford University, Stanford, CA 94305-5020, marc@charles.stanford.edu*

Abstract

Whether interaction between genes is better represented by synergistic or antagonistic epistasis has been a focus of experimental research in bacterial population genetics. Our previous research on evolution of modifiers of epistasis in diploid systems has indicated that the strength of positive or negative epistasis should increase provided linkage disequilibrium is maintained. Here we study a modifier of epistasis in fitness between two loci in a haploid system. Epistasis is modified in the neighborhood of a mutation selection balance. We show that when linkage in the three-locus system is tight, an increase in the frequency of a modifier allele that induces either more negative or more positive epistasis is possible. Epistasis here can be measured on either an additive or multiplicative scale.

Keywords

modifier theory; three-locus model; mutation-selection balance; viability-analogous equilibrium; antagonistic epistasis

0. Introduction

The evolution of epistasis is a topic largely absent from classical population genetic literature (Hansen et al., 2006). The subject has, however, received increased attention in the past decade as relevant data on haploid genetic systems have become available (Elena and Lenski 1997, 2001; Wilke and Adami 2001; You and Yin 2002; Sanjuan et al. 2004; Bonhoeffer et al. 2004; Burch and Chao 2004; Weinreich and Chao 2005; Shapiro et al. 2006). "Functional" (Hansen and Wagner 2001) or "physiological" (Cheverud and Routman 1995) epistasis refers to non-additive interaction among genes in production of phenotypes (see also Weinreich et al. 2005). This kind of epistasis is not a population property and is independent of allele frequencies. It can be expressed in terms of the parameters used to describe the phenotype of a multilocus individual.

Epistsis may occur at the genotypic level in determining fitness while not being detectable at the level of phenotype, as shown by Lunzer et al. (2005). In this paper we shall be concerned with the genotype-fitness relationship as well as the role of recombination. In Lunzer et al.

*Corresponding author

(2005), the locations of the variants were specified by amino-acids in a single enzyme; recombination was, therefore, not important in their chemostat competition experiments with *E. coli*.

In theoretical work on the evolution of recombination, the shape of the fitness as a function of the number of deleterious mutations has been shown to be critical (Feldman et al. 1980; Lenski et al. 1999; Otto and Lenormand 2002; Keightley and Otto 2006). With negative epistasis, for example, log fitness declines faster than linearly with increasing mutation number, and under such conditions recombination (or sex) is favored (Eshel and Feldman 1970; Feldman et al. 1980; Kondrashov 1988). Positive epistasis, on the other hand, will put sex at a disadvantage. Thus how the shape of this function evolves is also important. An approximate numerical treatment by Azevedo et al. (2006) concluded that when deleterious mutations can recombine there should be selection towards negative (synergistic) epistasis, but without recombination, positive epistasis should evolve. Understanding how epistasis evolves has broader implications, as epistatic interactions are "central to the evolution of genetic recombination" (Michalakis and Roze 2004). It should be noted, however, that in experiments with the RNA bacteriophage $\phi 6$, Froissart et al. (2004) were unable to detect any advantage to reassortment even in the presence of negative epistasis.

This manuscript is the third in a series of papers studying the evolution of epistasis. In our first two papers (Liberman and Feldman 2005,2006) we considered a diploid population, where the character under consideration is determined by two "major" loci with two possible alleles (*A* and *a*, *B* and *b*) at each locus and symmetric epistatic fitnesses. A third "modifier" locus is introduced, whose sole function is to determine the extent of epistasis in the selection at the major loci. By studying the equilibrium structure of such an epistatic-modifying system, we have shown that under well-specified conditions a modifying allele that increases epistasis succeeds. In other words, genetic interactions tend to become stronger.

In the present manuscript we assume that the population is haploid, rather than diploid. In this case viability selection alone cannot produce a stable polymorphism. Therefore we assume that at the two major loci there is a mutation-selection balance that produces a Hardy-Weinberg (HW) polymorphic equilibrium. As before, the modifier locus affects the extent of epistasis at the major loci. We analyze the external stability of the HW polymorphism towards the invasion of a new allele that modifies the existing epistasis. This model can be compared with our recent haploid analysis in which there are an arbitrary number of loci, each subject to the same rate of mutation to deleterious alleles (Desai et al. 2007). Fitness of each mutant was $(1 - s)$, but the haplotype with k mutations had fitness $e^{-sk(1+\varepsilon)}$, where ε describes the epistasis; $\varepsilon > 0$ entails negative or synergistic epistasis with $\varepsilon < 0$ representing positive or antagonistic epistasis. Antagonistic epistasis was shown to evolve under some simplifying assumptions on the relative rates of mutation and selection that we do not make in the present two-locus analysis.

The paper is organized as follows. Section 1 introduces the two-locus, two-allele haploid selection model. It is verified that with additive selection the mean fitness increases. In Section 2 we introduce independent unidirectional deleterious mutations at each locus (from $A \rightarrow a$, $B \rightarrow b$, each at rate μ) and describe the possible mutation-selection equilibria. In Section 3 a third modifier locus that controls epistasis is introduced and its dynamics explicated. In Section 4 the existence of a class of HW equilibria is revealed, in Section 5 we review the results, and in Sections 6, 7, and 8 we analyze the external stability of these HW polymorphic equilibria, thus determining the fate of a modifier allele that changes the extent of additive, or multiplicative, epistasis at the major loci. Detailed proofs of the main results are given in supplemental online material (SOM), sections A-E.

1. Two-locus haploid selection models

Consider a large population of haploid individuals and a character determined by two diallelic loci. Let *A*, *a* be the possible alleles at the first locus and *B*, *b* those at the second locus. Generations are non-overlapping, and the life cycle consists of random mating of gametes, recombination (in a diploid phase), Mendelian segregation, and viability selection in the haploid phase, in that order. We use the following notation:

Gamete:	AB	Ab	aB	ab
Frequency:	x_1	x_2	x_3	x_4
Viability:	w_1	w_2	w_3	w_4

(1.1)

Let r ($0 < r < 1$) be the recombination rate between the two loci. The frequencies x'_1,x'_2,x'_3,x'_4 of the four gametes in the next generation are given by

$$\begin{aligned} wx'_1 &= w_1(x_1 - rD) = w_1\widehat{x_1} \\ wx'_2 &= w_2(x_2 + rD) = w_2\widehat{x_2} \\ wx'_3 &= w_3(x_3 + rD) = w_3\widehat{x_3} \\ wx'_4 &= w_4(x_4 - rD) = w_4\widehat{x_4} \end{aligned} \tag{1.2}$$

where $D = x_1x_4 - x_2x_3$ is the *linkage disequilibrium* and w, the normalizing factor, is

$$w = \sum_{i=1}^{4} w_i\widehat{x_i} = \sum_{i=1}^{4} w_i x_i - rD(w_1 - w_2 - w_3 + w_4). \tag{1.3}$$

Unless mentioned to the contrary, it will be assumed that $w_i \neq w_j$ for $i \neq j$.

The expression

$$e_{\text{add}} = w_1 - w_2 - w_3 + w_4 \tag{1.4}$$

is the usual measure of *additive epistasis* associated with the fitness parameters w_1, w_2, w_3, w_4. Similarly

$$e_{\text{mult}} = w_1w_4 - w_2w_3 \tag{1.5}$$

is the measure of *multiplicative epistasis*.

Our interest centers on the evolution of epistasis. Therefore we would like to know, as a baseline, how a population evolves under non-epistatic selection, either additive or multiplicative, that is where $e_{\text{add}} = 0$ or $e_{\text{mult}} = 0$, respectively. We have the following two-part result.

Result 1

i. With additively non-epistatic selection, the mean fitness increases. That is

$$\sum_{i=1}^{4} w_i x'_i \geq \sum_{i=1}^{4} w_i x_i \tag{1.6}$$

with equality if and only if there is fixation in one of the four gametes. Moreover there is global convergence to fixation in the gamete with maximum viability.

ii. With multiplicatively non-epistatic selection, the only possible equilibria are the four fixations. Moreover the fixation in the gamete with maximum viability is the only locally stable equilibrium.

The proof is given in SOM, section A.

In view of Result 1, non-epistatic haploid selection does not allow polymorphic equilibrium. A possible source of polymorphism would be mutation, giving rise to a mutation-selection balance. Suppose, therefore, that in addition to haploid selection we have at each of the two loci independent unidirectional mutation $A \to a$ and $B \to b$. Let μ be the mutation rate, assumed to be the same at each locus. Then using the previous notation, the transformation of frequencies is given by

$$\begin{aligned} wx'_1 &= (1-\mu)^2 w_1 (x_1 - rD) \\ wx'_2 &= \mu(1-\mu) w_1 (x_1 - rD) + (1-\mu) w_2 (x_2 + rD) \\ wx'_3 &= \mu(1-\mu) w_1 (x_1 - rD) + (1-\mu) w_3 (x_3 + rD) \\ wx'_4 &= \mu^2 w_1 (x_1 - rD) + \mu w_2 (x_2 + rD) + \mu w_3 (x_3 + rD) + w_4 (x_4 - rD), \end{aligned} \tag{1.7}$$

where $w = \sum_{i=1}^{4} w_i \widehat{x_i} = \sum_{i=1}^{4} w_i x_i - rDe_{\text{add}}$ is the normalizing factor.

For the mutation-selection balance, suppose that $w_1 > w_2$, $w_3 > w_4$ so that *AB* is the gamete with the maximum fitness and *ab* the gamete with minimum fitness. These inequalitites will be assumed throughout the paper. Under these assumptions, the equilibrium $\mathbf{x}^* = (1, 0, 0, 0)$, corresponding to fixation in *AB*, is stable. Let $\mathbf{T}_\mu$ be the transformation of frequencies given in (1.7), then $\mathbf{T}_\mu \mathbf{x}^* = ((1-\mu)^2, \mu(1-\mu), \mu(1-\mu), \mu^2)$, and when $0 < \mu < 1$, $\mathbf{T}_\mu \mathbf{x}^*$ involves all four gametes with positive frequencies. Therefore from *small perturbation theory* (e.g., Karlin and McGregor 1971,1972) we know that for small μ there is a locally stable polymorphism near $\mathbf{x}^* = (1, 0, 0, 0)$. Thus in the haploid selection model with unidirectional mutation, a stable mutation-selection balance exists. Of course, the cases of non-epistatic selection are included in this result.

Although it is difficult to specify the equilibrium points of $\mathbf{T}_\mu$ explicitly, some classifications are possible. Direct computation produces the following result (see SOM, section B).

Result 2

The haploid selection-mutation model (1.7) has a polymorphic equilibrium with linkage equilibrium D = 0 if and only if selection is multiplicatively non-epistatic, in which case the mutation-selection equilibrium $\mathbf{x}^* = \left(x_1^*, x_2^*, x_3^*, x_4^*\right)$ is unique. It exists when

$$(1-\mu) w_1 > w_2, \quad (1-\mu) w_1 > w_3, \quad (1-\mu)^2 w_1 > w_4, \tag{1.8}$$

and it satisfies

$$\begin{aligned} x_2^* &= \frac{\mu w_1 x_1^*}{(1-\mu) w_1 - w_2}, \quad x_3^* = \frac{\mu w_1 x_1^*}{(1-\mu) w_1 - w_3}, \\ x_4^* &= \frac{\mu^2 w_1 x_1^*}{(1-\mu)^2 w_1 - w_4} \left[1 + \frac{w_2}{(1-\mu) w_1 - w_2} + \frac{w_3}{(1-\mu) w_1 - w_3} \right]. \end{aligned} \tag{1.9}$$

Remark 1—Result 2 entails that when there is multiplicative epistasis ($e_{\text{mult}} \neq 0$), equilibria with $D = 0$ must be boundary equilibria with $x_1 = 0$ and either $x_2 = 0$ or $x_3 = 0$. In fact three boundary equilibria with $D = 0$ are possible: (0, 0, 0, 1), $\left(0, 0, \tilde{x}_3, \tilde{x}_4\right)$, $\left(0, \tilde{\tilde{x}}_2, 0, \tilde{\tilde{x}}_4\right)$ where

$$\tilde{x}_3 = \frac{(1-\mu) w_3 - w_4}{w_3 - w_4}, \quad \tilde{x}_4 = \frac{\mu w_3}{w_3 - w_4}, \tag{1.10}$$

and

$$\tilde{\tilde{x}}_2 = \frac{(1-\mu) w_2 - w_4}{w_2 - w_4}, \quad \tilde{\tilde{x}}_4 = \frac{\mu w_2}{w_2 - w_4}. \tag{1.11}$$

As $w_1 > w_2$, $w_3 > w_4$, $\left(0,0,\tilde{x}_3,\tilde{x}_4\right)$ exists if $(1 - \mu)w_3 > w_4$ and $\left(0,\tilde{\tilde{x}}_2,0,\tilde{\tilde{x}}_4\right)$ exists if $(1 - \mu)w_2 > w_4$. Of course these equilibria may also exist when there is no multiplicative epistasis ($e_{\text{mult}} = 0$).

When a mutation-selection equilibrium exists for the recursion system (1.7), we have (see SOM, section B) the following result.

Result 3

At mutation-selection balance from (1.7), we must have

$$w > (1-\mu)\,w_2, \quad w > (1-\mu)\,w_3. \tag{1.12}$$

If $w_4 < w_2$, $w_3 < w_1$, then

$$w > w_4. \tag{1.13}$$

The sign of the linkage disequilibrium at a polymorphic equilibrium depends on the fitness parameters, as follows (see SOM, section B).

Result 4

(Felsenstein 1965, Feldman et al. 1980.) In the haploid mutation-selection model (1.7) at any interior equilibrium $\mathbf{x} = (x_1,x_2,x_3,x_4)$ with $x_i \neq 0$ for all i the linkage disequilibrium $D = x_1 x_4 - x_2 x_3$ satisfies

$$D\,(w_1 w_4 - w_2 w_3) > 0. \tag{1.14}$$

with $D = 0$ if and only if $w_1 w_4 = w_2 w_3$.

Thus the sign of D, at any polymorphic equilibrium, coincides with the sign of the multiplicative epistasis e_{mult}.

2. The epistasis modification model

Suppose that in addition to the two "major" loci there is a third *modifier locus*, with n possible alleles $M_1, M_2, \ldots, M_n$, whose sole function is to control epistasis among the haploid fitnesses at the two major loci. Let the frequencies of the four gametes AB, Ab, aB, ab at the major loci, in the present generation, be x_1, x_2, x_3, x_4, respectively. There is no mutation at the modifier locus. Then in this three-locus framework we can write

Gamete:	ABM_i	AbM_i	aBM_i	abM_i
Fitness:	w_1^i	w_2^i	w_3^i	w_4^i
Frequency:	$x_1 p_i$	$x_2 q_i$	$x_3 r_i$	$x_4 s_i$

(2.1)

where $0 \le p_i, q_i, r_i, s_i \le 1$ and $\sum_i p_i = \sum_i q_i = \sum_i r_i = \sum_i s_i = 1$. The population state in the present generation is therefore specified by the five vectors **x, p, q, r, s** where:

$$\begin{aligned}
\mathbf{x} &= (x_1,x_2,x_3,x_4)\\
\mathbf{p} &= (p_1,p_2,\ldots,p_n)\\
\mathbf{q} &= (q_1,q_2,\ldots,q_n)\\
\mathbf{r} &= (r_1,r_2,\ldots,r_n)\\
\mathbf{s} &= (s_1,s_2,\ldots,s_n).
\end{aligned} \tag{2.2}$$

The three-locus recombination distribution is determined by the four *crossover probabilities*

$$c_{00}, c_{01}, c_{10}, c_{11}, \tag{2.3}$$

where $c_{00} + c_{01} + c_{10} + c_{11} = 1$. Here c_{00} is the probability that there is no recombination between *A/a* and *B/b* nor between *B/b* and M_i/M_j. Similarly c_{01} is the probability that no recombination occurs between *A/a* and *B/b* but recombination occurs beween *B/b* and M_i/M_j. There is a corresponding interpretation for c_{10}, while c_{11} is the probability that recombination occurs both between *A/a* and *B/b* and between *B/b* and M_i/M_j. With this representation

$$r=c_{10}+c_{11}, \quad R=c_{01}+c_{11}, \tag{2.4}$$

where r is the recombination rate between the two major loci and R is the frequency of recombination between the modifier locus and the two major loci. As in the previous section we assume that μ is the rate of mutation $A \to a$, $B \to b$ at each locus and that recombination events are not affected by the modifier locus.

We observe the population at the gametic stage and in the course of one generation the population undergoes random mating, recombination, Mendelian segregation, haploid selection, and mutation, in that order. For ease of notation we represent the transformation of frequencies before mutation. Starting with the population state **x, p, q, r, s** we have:

Gamete:	ABM_i	AbM_i	aBM_i	abM_i
"New" frequency after selection and before mutation:	$\tilde{x}_1\tilde{p}_i$	$\tilde{x}_2\tilde{q}_i$	$\tilde{x}_3\tilde{r}_i$	$\tilde{x}_4\tilde{s}_i$
"New" frequency after mutation:	$x'_1p'_i$	$x'_2q'_i$	$x'_3r'_i$	$x'_4s'_i$.

(2.5)

For the "new" frequencies before mutation we can use the transformation of frequencies described in Liberman and Feldman (1986). Following that transformation and using the vector notation

$$\mathbf{w}_k=\left(w_k^1,w_k^2,\ldots,w_k^n\right) \quad \text{for} \quad k=1,2,3,4 \tag{2.6}$$

we have:

$$\begin{aligned}
w\tilde{x}_1\,\tilde{\mathbf{p}} = {} & x_1^2\mathbf{p}\circ\mathbf{w}_1+x_1x_2\left[(c_{00}+c_{10})\,\mathbf{p}\circ\mathbf{w}_1+(c_{01}+c_{11})\,\mathbf{q}\circ\mathbf{w}_1\right]\\
& +x_1x_3\left[(c_{00}+c_{11})\,\mathbf{p}\circ\mathbf{w}_1+(c_{01}+c_{10})\,\mathbf{r}\circ\mathbf{w}_1\right]\\
& +x_1x_4\left[c_{00}\mathbf{p}\circ\mathbf{w}_1+c_{01}\mathbf{s}\circ\mathbf{w}_1\right]\\
& +x_2x_3\left[c_{11}\mathbf{q}\circ\mathbf{w}_1+c_{10}\mathbf{r}\circ\mathbf{w}_1\right]\\
w\tilde{x}_2\,\tilde{\mathbf{q}} = {} & x_2^2\mathbf{q}\circ\mathbf{w}_2+x_1x_2\left[(c_{01}+c_{11})\,\mathbf{p}\circ\mathbf{w}_2+(c_{00}+c_{10})\,\mathbf{q}\circ\mathbf{w}_2\right]\\
& +x_2x_4\left[(c_{00}+c_{11})\,\mathbf{q}\circ\mathbf{w}_2+(c_{01}+c_{10})\,\mathbf{s}\circ\mathbf{w}_2\right]\\
& +x_1x_4\left[c_{11}\mathbf{p}\circ\mathbf{w}_2+c_{10}\mathbf{s}\circ\mathbf{w}_2\right]\\
& +x_2x_3\left[c_{00}\mathbf{q}\circ\mathbf{w}_2+c_{01}\mathbf{r}\circ\mathbf{w}_2\right]\\
w\tilde{x}_3\,\tilde{\mathbf{r}} = {} & x_3^2\mathbf{r}\circ\mathbf{w}_3+x_1x_3\left[(c_{01}+c_{10})\,\mathbf{p}\circ\mathbf{w}_3+(c_{00}+c_{11})\,\mathbf{r}\circ\mathbf{w}_3\right]\\
& +x_3x_4\left[(c_{00}+c_{10})\,\mathbf{r}\circ\mathbf{w}_3+(c_{01}+c_{11})\,\mathbf{s}\circ\mathbf{w}_3\right]\\
& +x_1x_4\left[c_{10}\mathbf{p}\circ\mathbf{w}_3+c_{11}\mathbf{s}\circ\mathbf{w}_3\right]\\
& +x_2x_3\left[c_{00}\mathbf{r}\circ\mathbf{w}_3+c_{01}\mathbf{q}\circ\mathbf{w}_3\right]\\
w\tilde{x}_4\,\tilde{\mathbf{s}} = {} & x_4^2\mathbf{s}\circ\mathbf{w}_4+x_2x_4\left[(c_{00}+c_{11})\,\mathbf{s}\circ\mathbf{w}_4+(c_{01}+c_{10})\,\mathbf{q}\circ\mathbf{w}_4\right]\\
& +x_3x_4\left[(c_{00}+c_{10})\,\mathbf{s}\circ\mathbf{w}_4+(c_{01}+c_{11})\,\mathbf{r}\circ\mathbf{w}_4\right]\\
& +x_1x_4\left[c_{01}\mathbf{p}\circ\mathbf{w}_4+c_{00}\mathbf{s}\circ\mathbf{w}_4\right]\\
& +x_2x_3\left[c_{10}\mathbf{q}\circ\mathbf{w}_4+c_{11}\mathbf{r}\circ\mathbf{w}_4\right].
\end{aligned} \tag{2.7}$$

Here $\circ$ is the Schur product of vectors, namely, if $\mathbf{a} = (a_1, a_2, ..., a_n)$ and $\mathbf{b} = (b_1, b_2, ..., b_n)$ then

$$\mathbf{a}\circ\mathbf{b}=(a_1b_1,a_2b_2,\ldots,a_nb_n), \tag{2.8}$$

and w is a normalizing factor such that

$$\tilde{x}_1\sum_{i=1}^{n}\tilde{p}_i+\tilde{x}_2\sum_{i=1}^{n}\tilde{q}_i+\tilde{x}_3\sum_{i=1}^{n}\tilde{r}_i+\tilde{x}_4\sum_{i=1}^{n}\tilde{s}_i=1. \tag{2.9}$$

The new frequencies after mutation are

$$\begin{aligned} wx'_1\mathbf{p}' &= (1-\mu)^2 w\tilde{x}_1\,\tilde{\mathbf{p}} \\ wx'_2\mathbf{q}' &= \mu(1-\mu)\,w\tilde{x}_1\,\tilde{\mathbf{p}} + (1-\mu)\,w\tilde{x}_2\,\tilde{\mathbf{q}} \\ wx'_3\mathbf{r}' &= \mu(1-\mu)\,w\tilde{x}_1\,\tilde{\mathbf{p}} + (1-\mu)\,w\tilde{x}_3\,\tilde{\mathbf{r}} \\ wx'_4\mathbf{s}' &= \mu^2 w\tilde{x}_1\,\tilde{\mathbf{p}} + \mu w\tilde{x}_2\,\tilde{\mathbf{q}} + \mu w\tilde{x}_3\,\tilde{\mathbf{r}} + w\tilde{x}_4\,\tilde{\mathbf{s}}\,, \end{aligned} \tag{2.10}$$

where w is the same normalizing factor as in (2.7).

3. Existence of Hardy-Weinberg equilibria

A special class of equilibria of the system (2.10) are the Hardy-Weinberg (HW) equilibria, where the modifier locus is in linkage equilibrium with the major loci. The three-locus gametic frequencies at these equilibria are given by a Kronecker product vector $\mathbf{x} \otimes \mathbf{y}$, where

$$\begin{aligned} &\mathbf{x} = (x_1, x_2, x_3, x_4)\,, \quad \mathbf{y} = (y_1, y_2, \ldots, y_n)\,, \\ \text{and} \quad &x \otimes y = (x_1y_1, x_1y_2 \ldots x_1y_n \ldots x_4y_1 \ldots x_4y_n)\,. \end{aligned} \tag{3.1}$$

Thus, the frequencies of the $4n$ three-locus gametes are

Gamete:	ABM_i	AbM_i	aBM_i	abM_i
Frequency:	x_1y_i	x_2y_i	x_3y_i	x_4y_i.

(3.2)

Comparing (3.2) with (2.1), at the HW equilibria we have $\mathbf{p} = \mathbf{q} = \mathbf{r} = \mathbf{s} = \mathbf{y}$. Using the transformation equations (2.7) and (2.10), at equilibrium we have, e.g.

$$\begin{aligned} wx_1\mathbf{y} &= (1-\mu)^2\mathbf{y} \circ \mathbf{w}_1 \left[x_1^2 + x_1x_2 + x_1x_3 + x_1x_4\,(c_{00}+c_{01}) + x_2x_3\,(c_{10}+c_{11}) \right] \\ &= (1-\mu)^2\mathbf{y} \circ \mathbf{w}_1\,[x_1 - rD]\,. \end{aligned} \tag{3.3}$$

Thus, using (2.4), the four equilibrium equations become

$$\begin{aligned} wx_1\mathbf{y} &= (1-\mu)^2\mathbf{y} \circ \mathbf{w}_1\,[x_1 - rD] \\ wx_2\mathbf{y} &= \mu(1-\mu)\,\mathbf{y} \circ \mathbf{w}_1\,[x_1 - rD] + (1-\mu)\,\mathbf{y} \circ \mathbf{w}_2\,[x_2 + rD] \\ wx_3\mathbf{y} &= \mu(1-\mu)\,\mathbf{y} \circ \mathbf{w}_1\,[x_1 - rD] + (1-\mu)\,\mathbf{y} \circ \mathbf{w}_3\,[x_3 + rD] \\ wx_4\mathbf{y} &= \mu^2\mathbf{y} \circ \mathbf{w}_1\,[x_1 - rD] + \mu\mathbf{y} \circ \mathbf{w}_2\,[x_2 + rD] + \mu\mathbf{y} \circ \mathbf{w}_3\,[x_3 + rD] + \mathbf{y} \circ \mathbf{w}_4\,[x_4 - rD] \end{aligned} \tag{3.4}$$

Summing all components for each of the equations in (3.4) using $\sum_{i=1}^{n} y_i = 1$, we find

$$\begin{aligned} wx_1 &= (1-\mu)^2\,(\mathbf{y},\mathbf{w}_1)\,[x_1 - rD] \\ wx_2 &= \mu(1-\mu)\,(\mathbf{y},\mathbf{w}_1)\,[x_1 - rD] + (1-\mu)\,(\mathbf{y},\mathbf{w}_2)\,[x_2 + rD] \\ wx_3 &= \mu(1-\mu)\,(\mathbf{y},\mathbf{w}_1)\,[x_1 - rD] + (1-\mu)\,(\mathbf{y},\mathbf{w}_3)\,[x_3 + rD] \\ wx_4 &= \mu^2\,(\mathbf{y},\mathbf{w}_1)\,[x_1 - rD] + \mu\,(\mathbf{y},\mathbf{w}_2)\,[x_2 + rD] + \mu\,(\mathbf{y},\mathbf{w}_3)\,[x_3 + rD] + (\mathbf{y},\mathbf{w}_4)\,[x_4 - rD]\,, \end{aligned} \tag{3.5}$$

where $(\mathbf{a}, \mathbf{b})$ is the scalar product of the two vectors $\mathbf{a} = (a_1, \ldots, a_n)$ and $\mathbf{b} = (b_1, \ldots, b_n)$, namely

$$(\mathbf{a},\mathbf{b}) = \sum_{i=1}^{n} a_i b_i. \tag{3.6}$$

Observe that equations (3.5) are equivalent to the equilibrium equations at the major loci resulting from (1.7) with associated fitness parameters

Gamete:	AB	Ab	aB	ab
Fitness:	$(\mathbf{y},\mathbf{w}_1)$	$(\mathbf{y},\mathbf{w}_2)$	$(\mathbf{y},\mathbf{w}_3)$	$(\mathbf{y},\mathbf{w}_4)$

(3.7)

where

$$(\mathbf{y},\mathbf{w}_k) = \sum_{i=1}^{n} w_k^i y_i$$

is the average fitness of the k-th gamete with weights given by $\mathbf{y} = (y_1, y_2 \ldots y_n)$.

Now substituting (3.5) into (3.4) gives for the first equation

$$(1-\mu)^2\,(\mathbf{y},\mathbf{w}_1)\,[x_1 - rD]\,\mathbf{y} = (1-\mu)^2\mathbf{y}\circ\mathbf{w}_1\,[x_1 - rD], \tag{3.8}$$

and since $0 < \mu < 1$ we have

$$[x_1 - rD]\,[(\mathbf{y},\mathbf{w}_1)\,\mathbf{y} - \mathbf{y}\circ\mathbf{w}_1] = 0. \tag{3.9}$$

Similarly, for the second equation we have

$$\mu(1-\mu)\,[x_1 - rD]\,[(\mathbf{y},\mathbf{w}_1)\,\mathbf{y} - \mathbf{y}\circ\mathbf{w}_1] + (1-\mu)\,[x_2 + rD]\,[(\mathbf{y},\mathbf{w}_2)\,\mathbf{y} - \mathbf{y}\circ\mathbf{w}_2] = 0. \tag{3.10}$$

Using (3.9) and $0 < \mu < 1$, (3.10) gives

$$[x_2 + rD]\,[(\mathbf{y},\mathbf{w}_2)\,\mathbf{y} - \mathbf{y}\circ\mathbf{w}_2] = 0. \tag{3.11}$$

In general we derive the following four equations at equilibrium:

$$\begin{aligned}
[x_1 - rD]\,[(\mathbf{y},\mathbf{w}_1)\,\mathbf{y} - \mathbf{y}\circ\mathbf{w}_1] &= 0\\
[x_2 + rD]\,[(\mathbf{y},\mathbf{w}_2)\,\mathbf{y} - \mathbf{y}\circ\mathbf{w}_2] &= 0\\
[x_3 + rD]\,[(\mathbf{y},\mathbf{w}_3)\,\mathbf{y} - \mathbf{y}\circ\mathbf{w}_3] &= 0\\
[x_4 - rD]\,[(\mathbf{y},\mathbf{w}_4)\,\mathbf{y} - \mathbf{y}\circ\mathbf{w}_4] &= 0.
\end{aligned} \tag{3.12}$$

We assume that at the major loci we have an interior equilibrium, which entails that $x_k \mp rD \neq 0$ (± depending on k) for all k. We conclude that a HW equilibrium exists if and only if for each $k = 1, 2, 3, 4$

$$(\mathbf{y},\mathbf{w}_k)\,\mathbf{y} = \mathbf{y}\circ\mathbf{w}_k. \tag{3.13}$$

Observe that the equilibrium equations (3.13) are exactly the one-locus multiallellic equilibrium equations with n alleles M_1, M_2, ..., M_n and haploid fitness parameters $w_k^1, w_k^2, \ldots, w_k^n$. We have called these HW equilibria, when they exist, *viability analogous HW* (VAHW) equilibria (Liberman and Feldman, 1986). We summarize these findings in the following:

Result 5

If a frequency vector $\mathbf{y}$ exists such that $(\mathbf{y},\mathbf{w}_k)\mathbf{y} = \mathbf{y}\circ\mathbf{w}_k$ for k = 1, 2, 3, 4, and $\mathbf{x} = (x_1,x_2,x_3,x_4)$, the frequencies vector of AB, Ab, ab, and ab, satisfies the two-locus mutation-selection equilibrium equations (3.5) with fitness parameters $(\mathbf{y}, \mathbf{w}_k)$ for k = 1, 2, 3, 4, then $\mathbf{x} \otimes \mathbf{y}$ is a VAHW equilibrium of the three-locus system (2.10).

Remark 5—The VA equations (3.13) are of the form

$$(\mathbf{w},\mathbf{z})\,\mathbf{z} = \mathbf{w}\circ\mathbf{z} \tag{3.14}$$

where $\mathbf{w} = (w_1, w_2,..., w_n)$ and $\mathbf{z} = (z_1, z_2,..., z_n)$. Observe that if $\bar{w} = (\mathbf{w},\mathbf{z})$, then

$$\bar{w}\,z_i = w_i z_i \quad i = 1,2,\ldots,n. \tag{3.15}$$

Hence, unless $w_1 = w_2 = ... = w_n$, at least one of z_i-s must be zero. In fact, if $I \subset \{1, 2,..., n\}$ is the set such that $w_i = w_j$ for $i, j \in I$, then any solution $\mathbf{z}^I$ of (3.14) corresponding to I has

$$z_i^I = 0 \quad \text{for} \quad i \notin I, \quad z_i^I \geq 0 \quad \text{for} \quad i \in I \quad \text{with} \sum_{i\in I} z_i^I = 1. \tag{3.16}$$

Therefore, unless there are special relations among the vectors $\mathbf{w}_1$, $\mathbf{w}_2$, $\mathbf{w}_3$, $\mathbf{w}_4$, the only solutions of the four equations (3.13) are the n fixations in any one of the n alleles M_1, M_2,..., M_n at the modifier locus. On the other hand, if, e.g., $\mathbf{w}_k = w_k\mathbf{e}$, where $\mathbf{e} = (1, 1,..., 1)$ for all k

= 1, 2, 3, 4, such that the fitnesses of the four gametes *AB*, *Ab*, *aB*, *ab* at the major loci are the same for each allele at the modifier locus, then any vector **y** with non-negative components and $\sum_{i=1}^{n} y_i = 1$ is a solution of the four equations (3.13).

4. External stability of a VAHW equilibrium

Let $\mathbf{x} \otimes \mathbf{y}$ be a VAHW equilibrium where $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ satisfies the equilibrium equations (3.14) and **x** solves the equations (3.5) with fitness parameters

$$w_k = (\mathbf{y}, \mathbf{w}_k) \quad k = 1,2,3,4. \tag{4.1}$$

Following Remark 5, only one of $y_1, y_2, \ldots, y_n$ is non-zero. Suppose now that a new allele M_i, where i is not one of the indices 1, 2,..., n, is introduced at the modifier locus, with new fitness parameters

$$w_1^i, w_2^i, w_3^i, w_4^i. \tag{4.2}$$

Mutation does not occur at the modifier locus after the initial appearance of M_i. The linear approximation for p_i, q_i, r_i, s_i, the frequencies of the new rare allele M_i among the four gametes *AB*, *Ab*, *aB*, *ab*, respectively, is derived from (2.7) and (2.10) and reduces to the following equations:

$$
\begin{aligned}
w x_1 p_i' = (1-\mu)^2 w_1^i \quad & \{x_1^2 p_i + x_1 x_2 [(c_{00}+c_{10}) p_i + (c_{01}+c_{11}) q_i] \\
& + x_1 x_3 [(c_{00}+c_{11}) p_i + (c_{01}+c_{10}) r_i] \\
& + x_1 x_4 [c_{00} p_i + c_{01} s_i] \\
& + x_2 x_3 [c_{11} q_i + c_{10} r_i]\} \\
w x_2 q_i' = \mu (1-\mu) w_1^i \quad & \{x_1^2 p_i + x_1 x_2 [(c_{00}+c_{10}) p_i + (c_{01}+c_{11}) q_i] \\
& + x_1 x_3 [(c_{00}+c_{11}) p_i + (c_{01}+c_{10}) r_i] \\
& + x_1 x_4 [c_{00} p_i + c_{01} s_i] \\
& + x_2 x_3 [c_{11} q_i + c_{10} r_i]\} \\
+ (1-\mu) w_2^i \quad & \{x_2^2 q_i + x_1 x_2 [(c_{01}+c_{11}) p_i + (c_{00}+c_{10}) q_i] \\
& + x_2 x_4 [(c_{00}+c_{11}) q_i + (c_{01}+c_{10}) s_i] \\
& + x_1 x_4 [c_{11} p_i + c_{10} s_i] \\
& + x_2 x_3 [c_{00} q_i + c_{01} r_i] \\
w x_3 r_i' = \mu (1-\mu) w_1^i \quad & \{x_1^2 p_i + x_1 x_2 [(c_{00}+c_{10}) p_i + (c_{01}+c_{11}) q_i] \\
& + x_1 x_3 [(c_{00}+c_{11}) p_i + (c_{01}+c_{10}) r_i] \\
& + x_1 x_4 [c_{00} p_i + c_{01} s_i] \\
& + x_2 x_3 [c_{11} q_i + c_{10} r_i]\} \\
+ (1-\mu) w_3^i \quad & \{x_3^2 r_i + x_1 x_3 [(c_{01}+c_{10}) p_i + (c_{00}+c_{11}) r_i] \\
& + x_3 x_4 [(c_{00}+c_{10}) r_i + (c_{01}+c_{11}) s_i] \\
& + x_1 x_4 [c_{10} p_i + c_{11} s_i] \\
& + x_2 x_3 [c_{00} r_i + c_{01} q_i] \\
w x_4 s_i' = \mu^2 w_1^i \quad & \{x_1^2 p_i + x_1 x_2 [(c_{00}+c_{10}) p_i + (c_{01}+c_{11}) q_i] \\
& + x_1 x_3 [(c_{00}+c_{11}) p_i + (c_{01}+c_{10}) r_i] \\
& + x_1 x_4 [c_{00} p_i + c_{01} s_i] \\
& + x_2 x_3 [c_{11} q_i + c_{10} r_i]\} \\
+ \mu^2 w_2^i \quad & \{x_2^2 q_i + x_1 x_2 [(c_{01}+c_{11}) p_i + (c_{00}+c_{10}) q_i] \\
& + x_2 x_4 [(c_{00}+c_{11}) q_i + (c_{01}+c_{10}) s_i] \\
& + x_1 x_4 [c_{11} p_i + c_{10} s_i] \\
& + x_2 x_3 [c_{00} q_i + c_{01} r_i]\} \\
+ \mu w_3^i \quad & \{x_3^2 r_i + x_1 x_3 [(c_{01}+c_{10}) p_i + (c_{00}+c_{11}) r_i] \\
& + x_3 x_4 [(c_{00}+c_{10}) r_i + (c_{01}+c_{11}) s_i] \\
& + x_1 x_4 [c_{10} p_i + c_{11} s_i] \\
& + x_2 x_3 [c_{00} r_i + c_{01} q_i]\} \\
+ w_4^i \quad & \{x_4^2 s_i + x_2 x_4 [(c_{00}+c_{11}) s_i + (c_{01}+c_{10}) q_i] \\
& + x_3 x_4 [(c_{00}+c_{10}) s_i + (c_{01}+c_{11}) r_i] \\
& + x_1 x_4 [c_{01} p_i + c_{00} s_i] \\
& + x_2 x_3 [c_{10} q_i + c_{11} r_i]\}. \qquad (4.3)
\end{aligned}
$$

Let **L** be the matrix corresponding to the linear approximation (4.3) and write

$$
\begin{bmatrix} p_i' \\ q_i' \\ r_i' \\ s_i' \end{bmatrix} = \mathbf{L} \begin{bmatrix} p_i \\ q_i \\ r_i \\ s_i \end{bmatrix}. \qquad (4.4)
$$

Then the external stability of $\mathbf{x} \otimes \mathbf{y}$ is determined by the largest root in absolute value of the fourth degree polynomial $Q(z)$, where

$$
Q(z) = \det[\mathbf{L} - z\mathbf{I}]. \qquad (4.5)
$$

As the leading coefficient (of z^4) in $Q(z)$ is 1, the value of $Q(z)$ as $z \to \infty$ is positive. Here **I** is the 4×4 identity matrix. We shall be especially interested in (4.5) at $z = 1$, namely $Q(1)$, since if $Q(1) < 0$, then $Q(z)$ has at least one positive root greater than 1.

We are interested in the influence of the sign and amount of epistasis introduced by the new allele M_i on the external stability of the VAHW equilibrium $\mathbf{x} \otimes \mathbf{y}$. We will assume that M_i determines the associated fitness parameters in a linear way, namely

$$w_k^i = w_k + w_k \delta_k \varepsilon_i \quad k=1,2,3,4, \tag{4.6}$$

where the δ_k-s can be either positive or negative, and $\varepsilon_i > 0$. Under this assumption $Q(1)$ is given by the following.

Result 6

Under the linear selection modification assumption (4.6) for small ε_i, and up to non-linear terms in ε_i,

$$Q(1) = \frac{(1-\mu)^2 w_1}{w^4 x_1} \varepsilon_i \det(\mathbf{N}) \tag{4.7}$$

where $\mathbf{N}$ is the 4×4 matrix given by

$$\begin{aligned}
n_{11} &= -x_1 c_{01} \\
n_{12} &= (x_1+x_3)\, c_{11} \\
n_{13} &= (x_1+x_2)\, c_{10} \\
n_{14} &= \delta_1 (x_1 - rD) \\
n_{21} &= -(1-\mu)\, w_2 (x_1+x_2)(c_{01}+c_{10}) \\
n_{22} &= (1-\mu)\, w_2 (c_{00}+c_{11}) - w \\
n_{23} &= 0 \\
n_{24} &= (1-\mu)\, w_2 [\delta_1 (x_1 - rD) + \delta_2 (x_2+rD)] \\
n_{31} &= -(1-\mu)\, w_3 (x_1+x_3)(c_{01}+c_{11}) \\
n_{32} &= 0 \\
n_{33} &= (1-\mu)\, w_3 (c_{00}+c_{10}) - w \\
n_{34} &= (1-\mu)\, w_3 [\delta_1 (x_1 - rD) + \delta_3 (x_3+rD)] \\
n_{41} &= -\mu w_2 (x_1+x_2)(c_{01}+c_{10}) - \mu w_3 (x_1+x_3)(c_{01}+c_{11}) - w_4 [x_4 + x_2 (c_{00}+c_{11}) + x_3 (c_{00}+c_{10}) + x_1 c_{00}] + w \\
n_{42} &= \mu w_2 (c_{00}+c_{11}) - w_4 [(x_2+x_4)(c_{00}+c_{11}) + (x_1+x_3) c_{00}] + w \\
n_{43} &= \mu w_3 (c_{00}+c_{10}) - w_4 [(x_3+x_4)(c_{00}+c_{10}) + (x_1+x_2) c_{00}] + w \\
n_{44} &= \mu w_2 [\delta_1 (x_1 - rD) + \delta_2 (x_2+rD)] + \mu w_3 [\delta_1 (x_1 - rD) + \delta_3 (x_3+rD)] + w_4 \delta_4 (x_4 - rD)
\end{aligned} \tag{4.8}$$

The proof is given in SOM, section C.

5. The evolution of an epistasis modifier—Results

Does a new modifier allele M_i determining new fitness parameters $w_1^i, w_2^i, w_3^i, w_4^i$ that induce "more" positive or "more" negative epistasis invade a population that is at a VAHW equilibrium? Using the additive and the multiplicative epistasis measures defined in (1.4) and (1.5), the new epistasis measures with the M_i allele present at the modifier locus are, respectively,

$$\begin{aligned}
e^i_{\text{add}} &= w_1^i - w_2^i - w_3^i + w_4^i \\
e^i_{\text{mult}} &= w_1^i w_4^i - w_2^i w_3^i.
\end{aligned} \tag{5.1}$$

Since the modifier allele M_i determines the fitness parameters linearly as in (4.6), the new measure of additive epistasis is

$$e^i_{\text{add}} = (w_1 - w_2 - w_3 + w_4) + (w_1\delta_1 - w_2\delta_2 - w_3\delta_3 + w_4\delta_4)\, \varepsilon_i. \tag{5.2}$$

As $(w_1 - w_2 - w_3 + w_4)$ is the "present" measure of additive epistasis, the change in additive epistasis is $(w_1\delta_1 - w_2\delta_2 - w_3\delta_3 + w_4\delta_4)\varepsilon_i$. If $(w_1 - w_2 - w_3 + w_4) > 0$, for example, then as we

assume that $\varepsilon_i > 0$, there is "more" or "less" additive epistasis depending on whether ($w_1\delta_1$ - $w_2\delta_2$ - $w_3\delta_3$ + $w_4\delta_4$) is positive or negative, respectively.

Similarly, with multliplicative epistasis the new measure is

$$e^i_{\text{mult}} = (w_1w_4 - w_2w_3) + (w_1w_4\delta_1 - w_2w_3\delta_2 - w_2w_3\delta_3 + w_1w_4\delta_4)\,\varepsilon_i + (w_1w_4\delta_1\delta_4 - w_2w_3\delta_2\delta_3)\,\varepsilon_i^2. \tag{5.3}$$

In our analysis we assume ε_i to be positive and "small" and ignore non-linear terms in ε_i. If w_1w_4 - $w_2w_3 > 0$, for example, then up to non-linear terms in ε_i, the change in multiplicative epistasis is determined by the sign of the term ($w_1w_4\delta_1$ - $w_2w_3\delta_2$ - $w_2w_3\delta_3$ + $w_1w_4\delta_4$), and "more" multiplicative epistasis corresponds to this term being positive while there is "less" multiplicative epistasis when this term is negative.

The fate of the new modifier alleles is determined by eigenvalues of the linear approximation matrix **L** of (4.3) or the sign of $Q(1)$ of (4.7). Unfortunately the general analysis of the eigenvalues of **L** and the sign of $Q(1)$ is too complicated. However, we can provide some results that shed light on the general trend of evolution. Specifically, we will discuss the following cases:

- **i.** $R \approx 0$, where the modifier locus and the major loci are tightly linked;
- **ii.** $r \approx 0$, where the two major loci are tightly linked;
- **iii.** Modifier changing only one fitness parameter.

6. The case $R \approx 0$

When $R = 0$ we have

$$c_{00} = 1 - r, \quad c_{01} = 0, \quad c_{10} = r, \quad c_{11} = 0. \tag{6.1}$$

As a result, the matrix **N** of (4.8) that determines the external stability of a VAHW equilibrium $\mathbf{x} \otimes \mathbf{y}$, has the special form

$$\mathbf{N} = \begin{bmatrix} 0 & 0 & n_{13} & n_{14} \\ n_{21} & n_{22} & 0 & n_{24} \\ 0 & 0 & n_{33} & n_{34} \\ n_{41} & n_{42} & n_{43} & n_{44} \end{bmatrix}. \tag{6.2}$$

It is easily seen that in this case $\det(\mathbf{N}) = -N_1 \cdot N_2$ where

$$N_1 = \det\begin{bmatrix} n_{13} & n_{14} \\ n_{33} & n_{34} \end{bmatrix}, \quad N_2 = \det\begin{bmatrix} n_{21} & n_{22} \\ n_{41} & n_{42} \end{bmatrix}. \tag{6.3}$$

Let $p = x_1 + x_2$, then from (4.8) we have

$$\begin{aligned} n_{13} &= rp & n_{14} &= \delta_1 (x_1 - rD) \\ n_{33} &= (1-\mu)(1-r)w_3 - w & n_{34} &= (1-\mu)w_3[\delta_1 (x_1 - rD) + \delta_3 (x_3 + rD)], \end{aligned} \tag{6.4}$$

and

$$\begin{aligned} n_{21} &= -(1-\mu)w_2 rp & n_{22} &= (1-\mu)w_2(1-r) - w \\ n_{41} &= -\mu w_2 rp - w_4(1-rp) + w & n_{42} &= \mu w_2(1-r) - w_4(1-r) + w. \end{aligned} \tag{6.5}$$

Direct computation of N_2, using the normalizing factor w of (1.3) reveals that

$$N_2 = (w - w_4)[w - (1-\mu)w_2](1-r) + rw[(w_1 - w_2)(x_1 - rD) + (w_3 - w_4)(x_3 + rD)]. \tag{6.6}$$

As we assume $w_1 > w_2$, $w_3 > w_4$, at equilibrium $w > w_4$, $w > (1 - \mu)w_2$ and we conclude that $N_2 > 0$ for all values of r and μ.

Following Result 6, a new modifier allele will invade the population if $N_1 > 0$. Now

$$N_1=\delta_1 (x_1 - rD) \quad \{w - (1-\mu) w_3 [1 - r(1+p)]\} + \delta_3 (x_3+rD)(1-\mu) w_3 rp. \tag{6.7}$$

Thus, when $r = 0$ N_1 is positive if and only if $\delta_1 > 0$, whereas when $r > 0$, N_1 is positive if

$$\delta_3 > -\delta_1 \frac{x_1 - rD}{x_3+rD} \cdot \frac{w-(1-\mu)[1-r(1+p)]w_3}{(1-\mu) rpw_3}. \tag{6.8}$$

Thus when $r = 0$ and $\delta_1 > 0$, with all the other δ_k values small in absolute value, then a new allele associated with more positive additive or multiplicative epistasis invades when $\varepsilon_i > 0$ and small. Moreover, when $r > 0$ condition (6.8) is valid when $\delta_1 > 0$ and $\delta_3 < 0$ but small in absolute value, but also when $\delta_1 < 0$ and $\delta_3 > 0$ and large enough. In these two cases when $|\delta_2|$ and $|\delta_4|$ are small, the modifier allele can invade if it introduces more positive as well as "more" negative additive or multiplicative epistasis. Thus, using small perturbation theory, we have the following result.

Result 7

Suppose the modifier locus is tightly linked to the two major loci (R ≈ 0) and let $\mathbf{x} \otimes \mathbf{y}$ be a VAHW equilibrium. Then it is possible that anew modifier allele that induces more positive or more negative additive or multiplicative epistasis invades the population for all values of r and μ for which $\mathbf{x} \otimes \mathbf{y}$ exists and is stable.

7. The case $r \approx 0$

If the two major loci are absolutely linked ($r = 0$), then the equilibrium equations at the major loci reduce to

$$\begin{aligned} wx_1 &= (1-\mu) w_1 x_1 \\ wx_2 &= \mu(1-\mu) w_1 x_1 + (1-\mu) w_2 x_2 \\ wx_3 &= \mu(1-\mu) w_1 x_1 + (1-\mu) w_3 x_3 \\ wx_4 &= \mu^2 w_1 x_1 + \mu w_2 x_2 + \mu w_3 x_3 + w_4 x_4. \end{aligned} \tag{7.1}$$

Solving the equilibrium equations yields the four types of equilibria represented in Table 1. It should be noted that the "corner" and the two "edge" equilibria have linkage equilibrium ($D = 0$), whereas the "face" equilibrium has negative linkage equilibrium ($D < 0$). Following Result 4, at the polymorphic equilibrium the sign of D coincides with the sign of the multiplicative epistasis $e_{\text{mult}} = w_1 w_4 - w_2 w_3$, with $D = 0$ when $e_{\text{mult}} = 0$ and selection is multiplicative.

In Table 2 we list sufficient conditions for a new modifier allele to invade, i.e., the VAHW equilibrium is externally unstable. The complete details of the analysis are given in SOM, section D. As all four types of equilibria are independent of R, we chose sufficient conditions that are also independent of R. Other sufficient conditions that depend on R are also available and can be obtained using SOM, section D.

Following Table 2, and the additive and multiplicative measures of epistasis associated with a new modifying allele described in (5.2) and (5.3), it is easily seen that a new modifier that changes epistasis to be "more positive" or "more negative," both in the additive and multiplicative sense, can invade. For example, if the present VAHW equilibrium $\mathbf{x} \otimes \mathbf{y}$ has $\mathbf{x} = \mathbf{x}^*$, the polymorphic equilibrium of Table 1, and originally there is positive additive epistasis $(w_1 - w_2 - w_3 + w_4) \geq 0$, or positive multiplicative epistasis $w_1 w_4 - w_2 w_3 \geq 0$, then from Table 2 if δ_1, $\delta_4 > 0$ and δ_2, $\delta_3 > 0$ but small, a modifier allele that increases epistasis (see (5.2) and (5.3)) will invade. Similarly, if originally we have negative epistasis, $(w_1 - w_2 - w_3 + w_4) \leq 0$ or $w_1 w_4 - w_2 w_3 \leq 0$, and δ_2, $\delta_3 > 0$ with δ_1, $\delta_4 > 0$ but small, then a modifier that induces "more negative" epistasis invades.

Using small perturbation theory the same phenomenon occurs also when $r \approx 0$, and we have the following result:

Result 8

If the two major loci are tightly linked ($r \approx 0$), then a new modifier allele associated with more positive or more negative additive or multiplicative epistasis can invade a population which is at a VAHW equilibrium.

8. Modification of one fitness parameter

We now discuss the four special cases where the new modifier allele changes only one of the fitness parameters. Here the only constraints on recombination are $0 \le R \le \frac{1}{2}$ and $0 \le r \le \frac{1}{2}$.

Case I

$\delta_1 = \delta_2 = \delta_3 = 0$, $w_4\delta_4 = 1$

In this case the modifier allele M_i does not alter the fitness parameters of the gametes *AB*, *Ab*, *aB* but changes the fitness of *ab* from w_4 to $w_4 + \varepsilon_i$. Thus in the presence of M_i the additive epistasis measure is

$$e^i_{\text{add}} = (w_1 - w_2 - w_3 + w_4) + \varepsilon_i, \tag{8.1}$$

while the multiplicative epistasis measure is

$$e^i_{\text{mult}} = (w_1 w_4 - w_2 w_3) + w_1 \varepsilon_i. \tag{8.2}$$

As we assume $\varepsilon_i > 0$, in both of these cases M_i introduces "more" positive epistasis. In this case we have the following result (see SOM, section E).

Result 9—If selection modification is of the form in Case I, then the VAHW equilibrium $\mathbf{x}\otimes\mathbf{y}$ is externally unstable towards the introduction of "more" positive (additive or multiplicative) epistasis. In particular the population evolves away from no (additive or multiplicative) epistasis towards positive epistasis.

Remark 6: Although Result 7 holds for small values of ε_i, Result 9 for Case I is true for any ε_i. This is easily seen from the expansion of $Q(1)$ given in SOM, section C under the assumptions of Case I.

Case II

$\delta_1 = \delta_2 = \delta_4 = 0$, $\delta_3 w_3 = 1$

In this case the modifier allele M_i changes only the fitness of *aB* from w_3 to $w_3 + \varepsilon_i$. Thus the total additive epistasis becomes

$$e^i_{\text{add}} = (w_1 - w_2 - w_3 + w_4) - \varepsilon_i, \tag{8.3}$$

and the multiplicative epistasis is

$$e^i_{\text{mult}} = (w_1 w_4 - w_2 w_3) - w_2 \varepsilon_i. \tag{8.4}$$

Hence, in both cases when $\varepsilon_i > 0$, M_i reduces the amount of positive epistasis or produces more negative epistasis. In this case we have (see SOM, section E)

Result 10—If the selection modification follows Case II, then any VAHW equilibrium $\mathbf{x} \otimes \mathbf{y}$ is externally unstable towards the introduction of more negative, but small, epistasis provided the mutation rate μ is sufficiently small.

Case III

$\delta_1 = \delta_3 = \delta_4 = 0, \delta_2 w_2 = 1$

In our model there is complete symmetry between the second and the third gametes *Ab* and *aB* respectively. Therefore the results of Case II apply also to this case; that is, Result 10 also holds under the assumptions of Case III.

Case IV

$\delta_2 = \delta_3 = \delta_4 = 0, \delta_1 w_1 = 1$

Here the new measures of epistasis are

$$e^i_{\text{add}} = (w_1 - w_2 - w_3 + w_4) + \varepsilon_i, \quad e^i_{\text{mult}} = (w_1 w_4 - w_2 w_3) + w_4 \varepsilon_i, \tag{8.5}$$

as the modifier allele M_i changes only the fitness of *AB* from w_1 to $w_1 + \varepsilon_i$. Hence epistasis becomes more positive when $\varepsilon_i > 0$. In this case we have a similar result to Result 9, namely:

Result 11—If selection modification follows Case IV, then the VAHW equilibrium $\mathbf{x} \otimes \mathbf{y}$ is externally unstable towards the introduction of "more" additive or multiplicative epistasis, provided mutation rates are sufficiently small.

The proof is given in SOM, section E.

9. Conclusions

A general analysis of the general problem with arbitrary recombination rates and general fitnesses and fitness perturbations appears to be intractable. In the case of no recombination our results are completely in line with those of Desai et al. (2007) in that a modifier increasing the level of antagonistic epistasis will be favored. When both *R* and *r* are very close to zero we again have evolution of more positive (multiplicative or additive) epistasis. However, when either *r* or *R* or both are larger, it is possible that more antagonistic or more synergistic epistasis evolves. This possibility was not seen in the multilocus analysis of Desai et al. The apparent discrepancy occurs because the fitness models are not exactly the same; by changing w_4, for example, in the present paper we have affected *s* and *ε* in their terms, not just the *ε* parameter.

Our results appear to differ from those of Azevedo et al. (2006), although these authors did not use a modifier gene to alter epistasis. They found that with recombination there should generally be selection towards (negative) synergistic epistasis. At best their conclusion is incomplete; as revealed by our modifier model the amount of recombination among the interacting genes can be important. Our analyses address evolution of epistasis between a single pair of diallelic loci, while the simulations by Azevedo et al. tracked ten diallelic loci. In principle, there could be 2^{10} - 10 - 1 (i.e., 1,013) epistatic interactions with ten diallelic loci, and it may be that the discrepancy between our results and theirs is due to their summary of these interactions by a single scalar.

In their work with deleterious mutations in *E. coli*, Elena and Lenski (1997) found no signal of epistasis in one set of experiments, but both positive and negative epistasis in another set of experiments. A recent reanalysis of these data (Beerenwinkel et al. 2007) found that among 27 double mutants there were more positive epistatic interactions between more deleterious mutations. The overall fitness landscape in therefore more complicated than the widely used

form $e^{-sk(1+\varepsilon)}$ (e.g., Desai et al. 2007). However, neither the data of Elena and Lenski nor the analysis by Beerenwinkel et al. really address how the shape of the fitness landscape, and in particular the signs of pairwise epistasis measures, might evolve. Our analysis suggests that across a genome with multiple mutations, since either positive or negative pairwise epistasis might evolve, a complex fitness landscape should be expected.

All of our results assume the population is initially in the neighborhood of a mutation selection balance. We cannot say what the dynamics of epistasis might be if it changes during a transient phase of the evolution. These haploid results would appear to confirm and extend our earlier findings for diploid populations, namely that an increased level of interaction between genes is likely to be favored by evolution. In the diploid case (Liberman and Feldman 2006) we saw that this was not necessarily a consequence of an increase in the mean fitness with increasing epistasis. It remains to be seen whether the same is true in the haploid model. It would also be interesting if, as was seen in Liberman and Feldman (2006), the way in which the equilibrium mean fitness changed with epistasis depended on the recombination rates.

REFERENCES

Azevedo RBR, Lohaus R, Srinivasan S, Dang KK, Burch CL. Sexual reproduction selects for robustness and negative epistasis in artificial gene networks. Nature 2006;440:87–90. [PubMed: 16511495]

Beerenwinkel N, Pachter L, Sturmfels B, Elena SF, Lenski RE. Analysis of epistatic interactions and fitness landscapes using a new geometric approach. BMC Evol. Biol 2007;7:60. [PubMed: 17433106]

Bonhoeffer S, Chappey C, Parkin NT, Whitcomb JM, Petropoulos CJ. Evidence for positive epistasis in HIV-1. Science 2004;306:1547–1550. [PubMed: 15567861]

Burch CL, Chao L. Epistasis and its relationship to canalization in the RNA virus *φ6*. Genetics 2004;167:559–567. [PubMed: 15238511]

Cheverud JM, Routman EJ. Epistasis and its contribution to genetic variance components. Genetics 1995;139:1455–1461. [PubMed: 7768453]

Desai MM, Weissman D, Feldman MW. Evolution can favor antagonistic epistasis. Genetics. In press

Elena SF, Lenski RE. Test of synergistic interactions among deleterious mutations in bacteria. Nature 1997;390:395–398. [PubMed: 9389477]

Elena SF, Lenski RE. Epistasis between new mutations and genetic background and a test of genetic canalization. Evolution 2001;55:1746–1752. [PubMed: 11681730]

Feldman MW, Christiansen FB, Brooks LD. Evolution of recombination in a constant environment. Proc. Natl. Acad. Sci. USA 1980;77:4838–4841. [PubMed: 16592864]

Felsenstein J. The effect of linkage on directional selection. Genetics 1965;52:349–363. [PubMed: 5861564]

Froissart R, Wilke CO, Montville R, Remold SK, Chao L, Turner PE. Coinfection weakens selection against epistatic mutations in RNA viruses. Genetics 2004;168:9–19. [PubMed: 15454523]

Hansen TF, Wagner GP. Modeling genetic architecture: a multilinear model of gene interaction. Theor. Pop. Biol 2001;59:61–86. [PubMed: 11243929]

Hansen TF, lvarez-Castro JM, Carter AJR, Hermisson J, Wagner GP. Evolution of genetic architecture under directional selection. Evolution 2006;60:1523–1536. [PubMed: 17017054]

Karlin S, McGregor J. On mutation selection balance for two-locus haploid and diploid populations. Theor. Popul. Biol 1971;2:60–70. [PubMed: 5170714]

Karlin S, McGregor J. Application of method of small parameters to multi-niche population genetic models. Theor. Pop. Biol 1972;3:186–209. [PubMed: 4667083]

Keightley PD, Otto SP. Interference among deleterious mutations favors sex and recombination in finite populations. Nature 2006;443:89–92. [PubMed: 16957730]

Lenski RE, Ofria C, Collier TC, Adami C. Genome complexity, robustness, and genetic interactions in digital organisms. Nature 1999;400:661–664. [PubMed: 10458160]

Liberman U, Feldman MW. Modifiers of mutation rate: a general reduction principle. Theor. Pop. Biol 1986;30:125–142. [PubMed: 3750215]

Liberman U, Feldman MW. On the evolution of epistasis I: diploids under selection. Theor. Popul. Biol 2005;67:141–160. [PubMed: 15808333]

Liberman U, Feldman MW. Evolutionary theory for modifiers of epistasis: a general symmetric model. Proc. Natl. Acad. Sci. USA 2006;103:19402–19406. [PubMed: 17167052]

Lunzer M, Miller SP, Felsheim R, Dean AM. The biochemical architecxture of an ancient adaptive landscape. Science 2005;310:499–501. [PubMed: 16239478]

Michalakis Y, Roze D. Epistasis in RNA viruses. Science 2004;306:1492–1493. [PubMed: 15567846]

Otto SP, Lenormand T. Resolving the paradox of sex and recombination. Nat. Rev. Genet 2002;3:252–261. [PubMed: 11967550]

Sanjuan R, Moya A, Elena SF. The distribution of fitness effects caused by single-nucleotide substitutions in an RNA virus. Proc. Natl. Acad. Sci. USA 2004;101:8396–8401. [PubMed: 15159545]

Shapiro B, Rambaut A, Pybus OG, Holmes EC. A phylogenetic method for detecting positive epistasis in gene sequences and its application to RNA virus evolution. Mol. Biol. Evol 2006;23:1724–1730. [PubMed: 16774976]

Weinreich DM, Chao L. Rapid evolutionary escape by large populations from local fitness peaks is likely in nature. Evolution 2005;59:1175–1182. [PubMed: 16050095]

Weinreich DM, Watson RA, Chao L. Sign epistasis and genetic constraint on evolutionary trajectories. Evolution 2005;59:1165–1174. [PubMed: 16050094]

Wilke CO, Adami C. Interaction between directional epistasis and averagemutational effects. Proc. R. Soc. B 2001;268:1469–1474.

You L, Yin J. Dependence of epistasis on environment and mutation severity as revealed by *in Silico* mutagenesis of phage T7. Genetics 2002;160:1273–1281. [PubMed: 11973286]

Table 1

Possible equilibria when $r = 0$

Type	*Frequencies*	*Existence Conditions*
Corner	$\bar{\mathbf{x}} = (0, 0, 0, 1)$	none
Edge	$\tilde{\mathbf{x}} = \left(0,\ 0,\ \tilde{x}3,\ \tilde{x}4\right)$, $\tilde{x}3 = \frac{(1-\mu)w_3 - w_4}{w_3 - w_4}$, $\tilde{x}4 = \frac{\mu w_3}{w_3 - w_4}$	$(1 - \mu)w_3 > w4$ $(1 - \mu)w_2 > w_4$
	$\approx{\mathbf{x}} = \left(0,\ \approx{x}2,\ 0,\ \approx{x}4\right)$, $\approx{x}2 = \frac{(1-\mu)w_2 - w_4}{w_2 - w_4}$, $\approx{x}4 = \frac{\mu w_2}{w_2 - w_4}$	
Face	$\hat{\mathbf{x}} = (0,\ a,\ a,\ 1 - 2a)$, $a = \frac{(1-\mu)w_2 - w_4}{2\left(w_2 - w_4\right)}$, $1 - a = \frac{\mu w_2}{w_2 - w_4}$	$w_2 = w_3$, $(1 - \mu)w_2 > w_4$
Polymorphism*	$\mathbf{x}^* = \left(x_1^*,\ x_2^*,\ x_3^*,\ x_4^*\right)$	$(1 - \mu)w_1 > w_2$, $(1 - \mu)w_1 > w_3$ $(1 - \mu)^2 w_1 > w_4$

*With the polymorphism $\mathbf{x}^*$ we have

$$x_2^* = \frac{\mu w_1 x_1^*}{(1-\mu)w_1 - w_2}, \quad x_3^* = \frac{\mu w_1 x_1^*}{(1-\mu)w_1 - w_3}, \quad x_4^* = \frac{\mu^2 w_1 x_1^*}{(1-\mu)^2 w_1 - w_4}\left[1 + \frac{w_2}{(1-\mu)w_1 - w_2} + \frac{w_3}{(1-\mu)w_1 - w_3}\right].$$

Table 2

Sufficient conditions for external instability

Equilibrium	Sufficient condition for external instability
Corner $\bar{\mathbf{x}} = (0, 0, 0, 1)$	$\delta_4 > 0$
Edge $\tilde{\mathbf{x}} = (0, 0, \tilde{x}_3, \tilde{x}_4)$	$\delta_3 > 0$, $w_3\delta_3 x_3 + w_4\delta_4 x_4 > 0$
Face $\hat{\mathbf{x}} = (0, \alpha, \alpha, 1 - 2\alpha)$	$\delta_2 = \delta_3 > 0$, $w_2\delta_2 x_2 + w_3\delta_3 x_3 + w_4\delta_4 x_4 > 0$
Polymorphism[†] $\mathbf{x}^* = (x_1^*, x_2^*, x_3^*, x_4^*)$	$\delta_1 > 0, \delta_2 > 0, \delta_3 > 0, \delta_4 > 0$ or $\delta_1 > 0, \sum_{k=1}^{4} w_k \delta_k x_k > 0, \quad \sum_{k=1}^{4} \delta_k x_k < 0$

[†]The conditions are given for the special case where $w_2 = w_3$ in which case $x_2^* = x_3^*$.