Alternative definitions of epistasis: dependence and interaction

Michael J. Wade, Rasmus G. Winther, Aneil F. Agrawal and Charles J. Goodnight

Although epistasis is at the center of the Fisher-Wright debate, biologists not involved in the controversy are often unaware that there are actually two different formal definitions of epistasis. We compare concepts of genetic independence in the two theoretical traditions of evolutionary genetics, population genetics and quantitative genetics, and show how independence of gene action (represented by the multiplicative model of population genetics) can be different from the absence of gene interaction (represented by the linear additive model of quantitative genetics). The two formulations converge with weak selection but not with strong selection or, for multiple loci, when the aggregated interaction terms are not negligible. As a result of the different formulations of gene interaction, the presence or absence of linkage disequilibrium, *D*, does not necessarily indicate the presence or absence of fitness epistasis. Indeed, linkage disequilibrium is generated in 'additive' models in quantitative genetics whenever two (or more) loci experience simultaneous selection. As a research strategy, it is often practical, for theoretical or experimental reasons, to minimize gene interaction by assuming independence of gene action in regard to fitness, or by assuming linear additive effects of multiple loci on a phenotype. However, minimizing the role of epistasis in theoretical investigations hinders our understanding of the origins of diversity and the evolution of complex phenotypes.

Michael J. Wade*
Rasmus G. Winther
Dept of Biology and Dept of History and Philosophy of Science, Indiana University, Bloomington, IN 47405, USA.
*e-mail: mjwade@bio.indiana.edu

Aneil F. Agrawal
Dept of Biology, Indiana University, Bloomington, IN 47405, USA.

Charles J. Goodnight
Dept of Biology, University of Vermont, 1155 Marsh Life Science Building, Burlington, VT 05405-0086, USA.

When phenotypic traits are characterized at the molecular level, many have a complex GENETIC ARCHITECTURE[1,2] (see Glossary). By genetic architecture, we mean the genes as well as the INTERACTIONS among them (epistasis), and between genes and environments that affect trait expression. Epistasis is particularly important in several areas of current evolutionary research, including speciation, CANALIZATION, inbreeding depression, the evolution of sex and interdemic selection in Wright's shifting balance theory[3]. In speciation, epistasis plays a key role in reproductive isolation – genes that function well in conspecific genetic backgrounds function poorly when combined in interspecific hybrids. Diminished effects of one locus as a result of interactions with another are fundamental to concepts of developmental homeostasis and canalization[4]. With inbreeding and other systems of nonrandom mating, negative synergism between deleterious alleles can contribute to the decline and extinction of small, endangered populations[5] and so therefore is a concern for conservation biologists. Similarly, epistasis for fitness (FITNESS EPISTASIS) might be essential to the evolution of sex and recombination[6–8]. The shifting balance theory was proposed specifically as a mechanism for the evolution of 'co-adapted gene complexes'[9–11].

Definitions of epistasis

Although there are several ways of defining epistasis[12–15], most definitions in evolutionary genetics are concordant with the concept that epistasis occurs when the differences in the phenotypic value of an allele at one locus are dependent on differences in specific alleles at one or more other loci. It is illustrative to consider epistasis, or genetic context, in the light of the somewhat better understood phenomenon of genotype-by-environment interaction (G × E), or environmental context. Just as variation in the environment can result in changes in the magnitude or order of ALLELIC VALUES with G × E (Fig. 1a), with epistasis, changes in the magnitude or order of allelic values occur with variations in genetic background (Fig. 1b)[14–18]. Whenever effects at one locus depend in this way on the genetic background at other loci, genetically subdividing a population into DEMES creates an opportunity for different genetic backgrounds to arise via random genetic drift, selection, or mutation. It is then possible for the effect on viability fitness of allele A_1 to be greater than that of alternative allele A_2 in one deme, but lower than it in another deme (Fig. 1c). This kind of change in the order of allelic effects among demes means that migration between demes will mix positive and negative gene frequency changes and retard local adaptation. In this sense, epistasis represents a kind of genetic constraint on evolution.

Population genetics and independence

The concept of allelic INDEPENDENCE is used in several different ways, leading to different definitions and different families of models[19]. In the most commonly used POPULATION GENETIC[20–22] model of epistasis, the MULTIPLICATIVE MODEL, the probablistic definition of independence of events is used. Thus, the joint effects on the viability of alleles at two different loci are considered to be independent (i.e. no epistasis) when the total viability of a genotype equals the product of the viability effects of the component loci. For example, if the viability fitness of genotype A_2A_2 is $(1-s)^2$ and that of genotype B_2B_2 is $(1-t)^2$, the viability of the $A_2A_2B_2B_2$ double homozygote would be $(1-t)^2(1-s)^2$ under this model (where $t, s>0$; t and s are the selection coefficients). Epistasis is modeled by introducing additional selection terms to the fitness of one or another of the multi-locus genotypes. Often, to reduce dimensionality and simplify mathematical analysis, the fitness of one genotype is assigned a value of one and other genotypes are assigned fitnesses relative to this standard, as one minus a selection coefficient (*cf.* Table I in Box 1). Empirically,

http://tree.trends.com 0169-5347/01/$ – see front matter PII: S0169-5347(01)02213-3

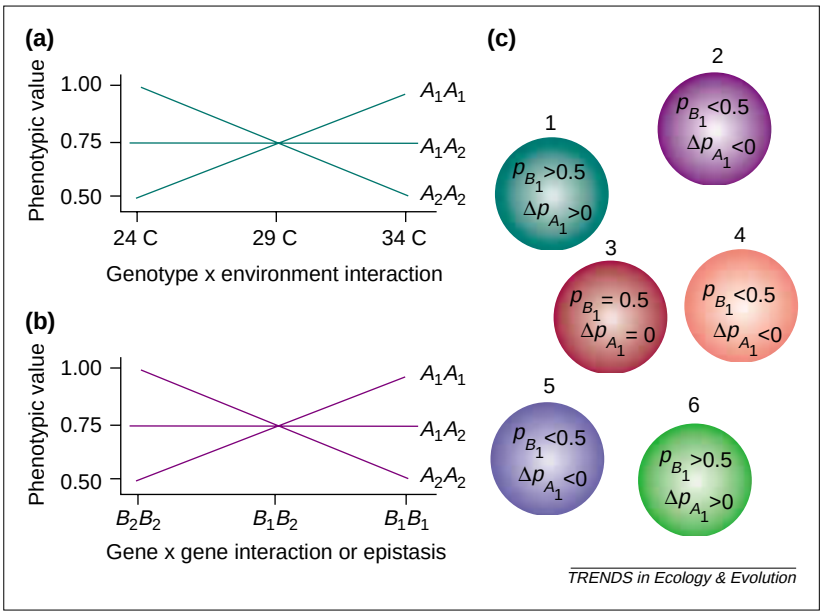

Fig. 1. A graphical view of interactions. (a) Genotype-by-environment interaction illustrating the change in the phenotypic values at the *A*-locus as a function of temperature. (b) Additive-by-additive epistasis illustrating the change in the phenotypic values at the *A*-locus as a function of genotype at the *B*-locus. (c) As a consequence of variations among demes in the frequency of the B_1 allele, p_{B1}, the allelic value of the A_1 allele changes from deme to deme. With natural selection favoring larger phenotypic values in every deme, this causes among-deme variations in the direction of selection experienced by the A_1 allele, Δp_{A1}. If the positive and negative values of Δp_{A1} are combined by migration, then total selection on the A_1 allele is reduced.

this procedure is problematic because: (1) it is rarely clear which genotype should be designated as the standard; (2) the average viability of the standard cannot usually be measured without error; (3) general environmental influences on viability negatively affect the statistical properties of viability ratios; and (4) the expected frequency of the standard genotype becomes vanishingly small as the number of loci increases. At present, multi-locus population genetic theory resembles a branch of applied mathematics whose findings inform evolutionary discussion, but whose many simplifying assumptions hinder empirical testing.

Quantitative genetics and interaction

In QUANTITATIVE GENETICS, there is only one family of models, based on linear regression[23]. In this regression model, alleles at loci, *A* and *B* (the independent variables), affect the phenotype (the dependent variable, *P*) in a manner described by Eqn 1:

$$P = b_1A + b_2B + b_{12}AB + e \qquad [1]$$

where b_1 and b_2 are the average effects of alleles at the *A* and *B* loci, respectively, on the phenotype *P*, whereas b_{12} describes the effect of interactions between alleles at the two loci (*e* reflects stochastic variations arising from the environment or from loci not being considered). The absence of epistasis means that the interaction coefficient b_{12} is zero; that is, the loci act additively (independently). This model is called the LINEAR ADDITIVE MODEL. For example, with fitness as the phenotype, if the deleterious effects on viability of allele A_2 are (–*s*), and B_2 are (–*t*), then the viability of the $A_2A_2B_2B_2$ double homozygote is $(1 - t - s)$ in the absence of epistasis (where mean viability is one; Box 1).

For two bi-allelic loci, there are four ways in which epistasis can be added to the regression[24,25]: (1 and 2) interactions between homozygotes at locus *A* and heterozygotes at locus *B* or vice versa (additive-by-dominance epistasis); (3) interactions between homozygotes at both loci (additive-by-additive epistasis; Fig. 1b); and (4) interactions between heterozygotes at both loci (dominance-by-dominance epistasis). The quantitative genetic approach lends itself to the experimental study of individual loci and their interactions when controlled breeding designs are coupled with molecular genetic markers. However, because the number of possible interactions increases much faster than do the numbers of loci, the application of the approach has its statistical limitations.

The two definitions are not equivalent

Terms or coefficients considered nonepistatic in the multiplicative model of population genetics can be epistatic in the linear additive model of quantitative genetics, and vice versa. This incongruence has implications both for how epistasis is measured in natural and laboratory populations, and for how inferences about its relevance to the evolutionary process are drawn. Simply put, observation is always theory laden[26–28]; with epistasis, there are two different theories, which are not concordant in the way they guide empirical observation. For example, although random genetic drift, mutation, migration or selection create associations between loci (linkage disequilibria), persistent associations are often considered to be evidence of selection[20]. However, whereas selection alone in the absence of epistasis, *sensu* the multiplicative model, does not generate LINKAGE DISEQUILIBRIUM, selection will generate it in the absence of epistasis, *sensu* the linear additive model[20]. Thus, depending on the model, the observation of linkage disequilibria might or might not signify epistasis for fitness, even in the absence of other evolutionary forces. To emphasize the difference between the approaches, we provide an illustration that the absence of epistasis in one model is not equivalent to the absence of epistasis in the other.

There is no doubt that interactions complicate theoretical analysis and that a variety of methods (such as transformation) or assumptions (such as independence), or both, have been made to eliminate the complicating features of epistasis from evolutionary theory. To paraphrase an earlier commentator[29], with respect to epistasis, realism has been traded for generality in both theories. Realism would require that epistasis be more explicitly incorporated in evolutionary theory.

Box 1. Converting models

Table I. The multiplicative model: no epistasis

	A_1A_1 **1**	A_1A_2 **$1-s$**	A_2A_2 **$(1-s)^2$**
B_1B_1 **1**	1	$1-s$	$(1-s)^2$
B_1B_2 **$1-t$**	$1-t$	$(1-t)(1-s)$	$(1-t)(1-s)^2$
B_2B_2 **$(1-t)^2$**	$(1-t)^2$	$(1-t)^2(1-s)$	$(1-t)^2(1-s)^2$

Table II. The linear additive model: epistasis

	A_1A_1 **$1+s$**	A_1A_2 **1**	A_2A_2 **$1-s$**
B_1B_1 **$1+t$**	$1+s+t$	$1+t$	$1+t-s+s^2$
B_1B_2 **1**	$1+s$	$1+ts$	$1-s+s^2+2ts-ts^2$
B_2B_2 **$1-t$**	$1-t+s+t^2$	$1-t+t^2+2ts-t^2s$	$1-t-s+t^2+4ts+s^2-2t^2s-2ts^2+t^2s$

Converting the multiplicative model to the linear additive model

The multiplicative model (Table I) with independence of gene action (i.e. no epistasis), when converted to the linear additive model (Table II), creates epistasis as indicated. The terms t and s are the selection coefficients. The first row and column of both tables contain the single-locus genotypes and the associated allelic effects appropriate to the respective model. Note that the epistatic terms in the linear additive model are all equal to or higher than second order (i.e. t^2, s^2, or ts). Hence, when selection is weak, the models converge.

Converting the linear additive model to the multiplicative model

The linear additive model (Table III) with absence of gene interaction (i.e. no epistasis), when converted to the multiplicative model (Table IV), creates epistasis as indicated. The first row and column of both tables contain the single-locus genotypes and the associated allelic effects appropriate to the respective model. Note that the epistatic terms in the multiplicative model are all equal to or higher than second order (i.e. t^2, s^2, or ts). Hence, when selection is weak, the models converge.

Table III. The linear additive model: no epistasis

	A_1A_1 **$1+s$**	A_1A_2 **1**	A_2A_2 **$1-s$**
B_1B_1 **$1+t$**	$1+t+s$	$1+t$	$1+t-s$
B_1B_2 **1**	$1+s$	1	$1-s$
B_2B_2 **$1-t$**	$1-t+s$	$1-t$	$1-t-s$

Table IV. The multiplicative model: epistasis

	A_1A_1 **1**	A_1A_2 **$1-s$**	A_2A_2 **$(1-s)^2$**
B_1B_1 **1**	1	$1-s$	$(1-s)^2-s^2$
B_1B_2 **$1-t$**	$1-t$	$(1-t)(1-s)-ts$	$(1-t)(1-s)^2-s^2-2ts+ts^2$
B_2B_2 **$(1-t)^2$**	$(1-t)^2-t^2$	$(1-t)^2(1-s)-t^2-2ts+t^2s$	$(1-t)^2(1-s)^2-t^2-4ts-s^2+2t^2s+2ts^2-t^2s^2$

The relationship between the models in the absence of epistasis

Independence, *sensu* the multiplicative model, is not equivalent to the absence of gene interaction, *sensu* the linear additive model. We illustrate this with two bi-allelic loci, A and B. The A_2 allele reduces fitness by a factor s relative to the A_1 allele. The B_2 allele reduces fitness by a factor t relative to the B_1 allele. Under the definition of independence from the multiplicative model, the absence of epistasis produces the effects shown in Table I in Box 1. However, when converted to the linear additive model, epistasis is clearly present (Table II in Box 1). Reciprocally, the absence of gene interaction in the linear additive model produces the effects shown in Table III in Box 1. When we translate this into the terms of the multiplicative model, again, epistatic terms are manifest (Table IV in Box 1). Note, however, that all of the epistatic terms in Box 1 are second or higher order functions of s and t. Terms of this magnitude are often (but not always) assumed to be zero in analyses of population genetic models. In a two locus bi-allelic model, these terms are not only small with weak selection (i.e. 'nearly neutral' alleles) but also few in number. When more than two loci are considered, the number of higher order terms can grow so large that even extremely small pairwise effects might be large when summed together[18].

We now turn to two phenomena, linkage disequilibrium and episodic selection, in which differences between the models generate epistasis. In the light of this, clearer definitions of, and empirical measurement criteria for, epistasis are required.

The generation of linkage disequilibrium with and without epistasis

Although several evolutionary forces, such as migration, random genetic drift, mutation and selection, can cause linkage disequilibria, its persistence is often regarded as the signature of multi-locus epistatic selection[20,22,30,31]. For single large panmictic populations without mutation, gametic phase disequilibria between recombining loci eventually disappear without epistatic selection. One of the mathematical advantages of the multiplicative model is that selection does not

Box 2. Why linkage disequilibrium is not always a signature of epistatic selection

Consider the haploid version of the linear additive model (Box 1), in which the fitness of an individual of genotype $A_1B_1 = 1$, $A_1B_2 = (1 - t)$, $A_2B_1 = (1 - s)$, and $A_2B_2 = (1 - t - s)$. Let the frequency of the A_1 allele be p_A and that of the B_1 allele be p_B. If we assume that a population is initially in linkage equilibrium (i.e. $D = 0$), then the frequencies of the four haploid genotypes before selection conform to the Hardy-Weinberg expectations. The frequency of each genotype after selection equals the product of its frequency before selection times its relative fitness, which equals its genotypic fitness divided by the mean population fitness, W, which is given by Eqn I:

$$W = 1 - (1 - p_B)(t) - (1 - p_A)(s) \quad [I]$$

The linkage disequilibrium generated by a single generation of selection, is given by Eqn II:

$$D' = -(1 / W)^2(p_B)(p_A)(1 - p_B)(1 - p_A)(ts) \quad [II]$$

Clearly, the linkage disequilibrium generated in the linear additive model is proportional to the product ts. Here, linkage disequilibrium is generated despite the absence of epistasis for fitness.

Table I. The generation of linkage disequilibrium by epistatic selection

Epistasis?	Multiplicative model	Linear additive model
No	$D' = 0$	$D' \neq 0$
Yes	$D' \neq 0$	$D' \neq 0$ or $D' = 0$

By contrast, with the haploid version of the multiplicative model of fitness (Table I in Box 1), the fitness of the A_2B_2 genotype is $(1 - t)(1 - s)$. If we assume that a population is initially in linkage equilibrium (i.e. $D = 0$), then it will remain so despite selection. This happens because the initial value of D appears as a factor of D', given in Eqn III:

$$D' = (1 / W)^2(1 - s)(1 - t)(D) = 0 \quad [III]$$

where the average fitness in the population, W, equals that above with the added term of $(+ts)(1 - p_A)(1 - p_B)$. Here, linkage disequilibrium is not generated in the absence of epistasis in the multiplicative model, and despite the presence of epistasis in the linear additive model. Table I summarizes the differences between the two models with respect to whether linkage disequilibrium is a signature of epistatic selection.

generate linkage disequilibria. That is, if D is zero before selection, it is also zero after a single generation of selection ($D' = 0$). This conclusion is clear when linkage disequilibrium after selection, D', is expressed as a multiplicative function of the linkage disequilibrium before selection (Box 2). When this multiplicative model is analyzed from the perspective of the linear additive model, there is epistasis (Box 1). This means that there is at least one form of epistatic selection under the linear additive model that does *not* generate linkage disequilibrium. We can also show the opposite.

With the linear additive model and no gene interaction (Box 2), if D is zero before selection, it is not zero after selection ($D' \neq 0$). In this case, we cannot express D' as a multiplicative function of D. Instead, selection without gene interaction for fitness produces an amount of linkage disequilibrium, D', proportional to the product, ts. Consistent with our earlier remarks, when s and t are small, the linkage disequilibrium created by selection, D', will also be small.

This linkage disequilibrium, created by selection in the absence of gene interaction, *sensu* the linear additive model, contributes to the phenomenon known as INTERFERENCE[32,33], whereby the change in gene frequency at one locus is slowed by simultaneous selection at another locus. Note that, because the term ts is always positive (Box 2), selection always generates a negative genetic correlation between advantageous alleles at different loci or between disadvantageous alleles (i.e. $D<0$). Thus, alleles at different loci simultaneously selected in the same direction always interfere with one another to reduce the rate of single gene evolution. Conversely, when an allele is positively selected at one locus and an allele at another locus is negatively selected, then the term ts is always negative (Box 2) and selection always generates a positive genetic correlation between them. This is also interference. The generation of linkage disequilibrium by selection in the absence of gene interaction for fitness is counterintuitive for many who associate the existence of nonzero D with epistatic selection. However, under the linear additive model, the fitness of the least-fit genotype, $A_2A_2B_2B_2$, for example, is lower than it would be under multiplicative independence (compare entries in Tables III and IV of Box 1). When genotypic fitnesses are not multiplicatively independent, linkage disequilibrium is generated by selection and there is interference.

In summary, considering both cases of no epistasis, it is clear that there is no necessary connection between epistatic selection and the existence of nonzero values of D, as is often expected (Table I in Box 2).

Glossary

Additivity: characterizes the mapping of phenotype onto genotype when the population mean genotypic value equals the sum of the genotypic values for each locus considered separately.
Allelic value: the regression of genotypic value on allele number. In the absence of dominance (allelic interactions within a locus) and epistasis, genotypic value is the sum of allelic values.
Canalization: the internal process of regulation during development that constrains phenotypic variation.
Deme: a group of individuals among which mating is random.
Genetic architecture: the characterization of a phenotype in terms of the direct effects of genes and environment as well as the genetic and environmental interactions affecting trait expression.
Fitness epistasis: occurs when the fitness value of an allele at one locus depends upon alleles at one or more other loci.
Independence: refers to a pair of events, *x* and *y*, such that the probability of *x* conditional on *y* is the same as the probability of *x*.
Interaction: refers to cases when the phenotype cannot be predicted from knowledge of the average effects of either genes or environment alone.
Interference: the reduction in the rate of evolution of one allele that occurs whenever there is simultaneous selection at another locus.
Linear additive model: the family of regression models in quantitative genetics. Under this model, the phenotype of a multi-locus genotype does not include any terms for interactions among loci in the absence of epistasis.
Linkage disequilibrium (*D*): the statistical association, or covariance, of alleles at two loci. For alleles A_1 and B_1 occurring at frequencies p_{A1} and p_{B1} at loci *A* and *B*, respectively, linkage disequilibrium equals $(p_{A1B1} - p_{A1}p_{B1})$, where p_{A1B1} is the frequency of $A_1 - B_1$ haplotypes.
Multiplicative model: one family of fitness models in population genetics. Under this model, the fitness of a multi-locus genotype equals the product of the fitnesses of the single-locus genotypes.
Population genetics: the study of specific genes in populations and the micro-evolutionary forces that affect their frequencies. Experimental population genetics tends to be concerned with change in the frequency of specific genes or molecular markers.
Quantitative genetics: the study of polygenic traits and the statistical genetic properties of such traits when subjected to artificial selection and micro-evolutionary forces. Although founded on explicit genetic models, experimental quantitative genetics, in practice, tends to be concerned with measurable properties of individuals and populations.

Additive traits and multiplicative fitnesses: a recipe for epistatic selection

In studies of phenotypic selection in nature, selection is partitioned into sequential episodes, such as germination or birth, early viability, late viability and fecundity[20,34–36]. In these studies, total lifetime fitness is equal to the product of the fitnesses at each stage in the life history. The logarithmic transformation of lifetime fitness defined as a product of components then becomes the sum of the logs of individual components. Logarithmic transformation is also recommended as a statistical means of removing epistasis (as could be the case with lifetime fitness) when '…the effects of different loci [are] combined by multiplication rather than by addition'[23]. This is an implicit reference to the multiplicative model as the basis of the underlying biology: we could hypothesize that the *A* locus affects early viability whereas the *B* locus affects later viability, for example. However, removing epistasis by switching between absolute and logarithmic scales can itself be problematic[23] because epistasis cannot be removed by such a transformation whenever a difference in rank order of fitnesses exists among genotypes at one locus when measured across different genetic backgrounds. (This is always the case with the kind of additive-by-additive epistasis shown in Fig. 1b.) Furthermore, with linear additive traits contributing to multiplicative components of total fitness, epistasis cannot be removed by switching between scales as we show here.

If allelic effects are additive for any of the fitness components (as is typically assumed of fecundity and other size-correlated fitness attributes), total fitness will exhibit epistasis under both models, and this epistasis cannot be eliminated by logarithmic transformation. For example, let two additively acting loci, *A* and *B* (e.g. Table III in Box 1), affect egg-to-adult viability fitness, W_V, and two other similarly additively acting loci, *C* and *D* with effects *u* and *v*, affect fecundity fitness, W_F. If viability and fecundity affect fitness independently, the total lifetime fitness, W_T, of a genotype equals the product, $(W_V)(W_F)$. An individual of genotype, $A_2A_2B_2B_2\,C_2C_2D_2D_2$, would have W_V of $(1 - 2t - 2s)$, W_F of $(1 - 2u - 2v)$ and lifetime fitness of $(1 - 2t - 2s)(1 - 2u - 2v)$. Taking the logarithm of lifetime fitness would serve to make the fitness components 'additive' on a log scale (Eqn 2):

$$\mathrm{Log}[W_T] = \mathrm{Log}[(W_V)(W_F)] = \mathrm{Log}[W_V] + \mathrm{Log}[W_F] \quad [2]$$

However, the ADDITIVITY of the underlying loci of each fitness component so transformed is not preserved, because (Eqn 3):

$$\mathrm{Log}[W_V] = \mathrm{Log}[(1 - 2t - 2s)] \sim \mathrm{Log}[(1 - 2t)] + \mathrm{Log}[(1 - 2s)] \quad [3]$$

to order s^2, t^2 and st, but only when *s* and *t* are both small (<0.01). When *s* and *t* are equal to or greater than 0.05, then Eqn 2 is accurate only to order *s* or *t*. (A similar argument applies to W_F.) Genetic inviability or infertility for some genotypes at any stage in the life history not only represents strong selection but also presents the problem of taking the logarithm of zero[37]. Thus, with additive traits and multiplicative fitness components, that is, the case deemed most representative of natural populations, there will always be epistasis for fitness with strong selection.

Conclusion: independence and reductionism

The definitions of independence and interaction are not congruent. The definition employed determines how we observe epistasis and also the evolutionary consequences attributed to it (e.g. the generation of linkage disequilibrium). There is no doubt that theoretical analyses in both population genetics and quantitative genetics are facilitated by diminishing or removing epistasis. However, attempts to remove it under one model produce it for the other (Box 1) and the models only converge when all higher order selection terms are relatively small in aggregate. What difference does it make for our perception of the evolutionary process if epistasis is diminished or absent in much of the mathematical theory? How do the simplifying assumptions, which stem in large part from mathematical convenience and might not be representative of living systems, contribute to

shaping what has been called the 'simple reductionistic'[38] view of allelic effects and the debate over the importance of epistasis in evolution? The simplifying assumptions can be summarized as two heuristics that are common in evolutionary discussion: the *ceteris paribus* ('other things being equal') clause and reductionism, which conflates the structural and functional identification of loci.

Scientific models are known to be sensitive to the *ceteris paribus* clause, which refers to background factual conditions and conceptual assumptions. For example, in physics, it has been shown that the apparently context-independent laws of physics are really '*ceteris paribus* laws' which are not only sensitive to context but also are false (i.e. they do not correspond to nature) whenever the clause is omitted (i.e. when all pertinent factual conditions are not specified)[39]. We believe that an analogous situation exists in evolutionary biology regarding assumptions about genetic backgrounds and gene interactions. Here, the *ceteris paribus* clause includes the assumption of an absence of epistasis because of the related assumption of the averaging out of genetic context in sufficiently large natural populations[40,41]. Presupposing the absence of epistasis justifies both the simplifying assumptions mitigating interaction in models, and the employment of transformation to statistically remove epistasis in observations. Because the clause is often implicit or hidden, the absence of epistasis in data analyses or theoretical models can be falsely interpreted as its true absence in nature, rather than as a consequence of employing the research heuristic of *ceteris paribus* as an unstated background assumption regarding the absence of epistasis.

Reductionism is characteristic of any genetic analysis, in which properties of the genetic system as a whole are derived solely from the independent properties of the loci from which the system is constructed. Relationships between gene loci are often ignored because loci and their effects are assumed to be not only structurally decomposable (genes can be linearly mapped) but also functionally additive (no interaction effects exist). However, as we have demonstrated, the meaning of functional additivity (independence or absence of interaction) depends on the type of model employed. With epistasis, gene function cannot be decomposed as easily as gene structure[27,41–45]. In some discussions, an individual organism is considered to be embedded in a biological hierarchy composed of nested levels of interacting components[46–49], ranging from chromosomes, to cells, to populations, and even to ecological communities. In this formulation, interaction can occur at any level of organization above the locus. This means that the functional properties of structurally decomposable single genes, in general, cannot be understood as invariant, monadic properties of gene loci and epistasis must be considered in evolutionary analysis.

Epistasis is not consistently defined in the different theories, making it difficult to empirically assess its role in evolutionary genetics and adaptive evolution. Some of the most favored single-gene examples of this field, such as sickle cell anemia, reveal significant epistasis when examined more thoroughly[17]. We believe that a comprehensive reconsideration of the importance of epistasis in the evolutionary process is warranted[3].

Acknowledgements
We thank Jeff Demuth, Elihu Gerson, James Griesemer, Norman Johnson, Doug Karpa-Wilson, Sandhya Kilaru, Jake Moorad and Meaghan Saur for useful comments and discussion.

References

1 Clark, A.G. and Wang, L. (1998) Epistasis in measured genotypes: *Drosophila P* element insertions. *Genetics* 147, 157–163
2 Mackay, T. and Fry, J. (1996) Polygenic mutation in *Drosophila melanogaster*: genetic interactions between selection lines and candidate quantitative trait loci. *Genetics* 144, 671–688
3 Wolf, J.B. *et al*., eds (2000) *Epistasis and the Evolutionary Process*, Oxford University Press
4 Gibson, G. and Wagner, G.P. (2000) Canalization in evolutionary genetics: a stabilizing theory? *Bioessays* 22, 372–380
5 Lynch, M. *et al*. (1993) The mutational meltdown in asexual populations. *Heredity* 84, 339–344
6 Kimura, M. and Maruyama, T. (1966) The mutational load with epistatic gene interactions in fitness. *Genetics* 54, 1337–1351
7 Kondrashov, A.S. (1988) Deleterious mutations and the evolution of sexual reproduction. *Nature* 336, 435–440
8 Charlesworth, B. (1990) Mutation–selection balance and the evolutionary advantage of sex and recombination. *Genet. Res.* 55, 199–221
9 Wright, S. (1931) Evolution in mendelian populations. *Genetics* 16, 97–159
10 Wade, M.J. (1996) Adaptation in subdivided populations: kin selection and interdemic selection. In *Adaptation* (Rose, M.R. and Lauder, G.V., eds), pp. 381–405, Academic Press
11 Wade, M.J. and Goodnight, C.J. (1991) Wright's Shifting Balance Theory: an experimental study. *Science* 253, 1015–1018
12 Wade, M.J. (1992) Epistasis. In *Keywords in Evolutionary Biology* (Keller, E.F. and Lloyd, E.A., eds), pp. 87–91, Harvard University Press
13 Phillips, P.C. (1998) The language of gene interaction. *Genetics* 149, 1167–1171
14 Phillips, P.C. *et al*. (2000) Beyond the average: the evolutionary importance of gene interactions and variability of epistatic effects. In *Epistasis and the Evolutionary Process* (Wolf, J.B. *et al*., eds), pp. 20–38, Oxford University Press
15 Brodie, E.D., III (2000) Why evolutionary genetics does not always add up. In *Epistasis and the Evolutionary Process* (Wolf, J.B. *et al*., eds), pp. 3–19, Oxford University Press
16 Wagner, G.P. *et al*. (1998) Genetic measurement theory of epistatic effects. *Genetica* 102/103, 569–580
17 Templeton, A. (2000) Epistasis and complex traits. In *Epistasis and the Evolutionary Process* (Wolf, J.B. *et al*., eds), pp. 41–57, Oxford University Press
18 Wade, M.J. (2000) Epistasis as a genetic constraint within populations and an accelerant of adaptive divergence among them. In *Epistasis and the Evolutionary Process* (Wolf, J.B. *et al*., eds), pp. 213–231, Oxford University Press
19 Christiansen, F.B. (2000) *Population Genetics of Multiple Loci*, John Wiley & Sons
20 Hedrick, P.W. (2000) *Genetics of Populations* (2nd edn), Jones & Bartlett Publishers
21 Hartl, D.L. and Clark, A.G. (1989) *Principles of Population Genetics* (2nd edn), Sinauer Associates Publishers
22 Otto, S.P. and Feldman, M.W. (1997) Deleterious mutations, variable epistatic interactions, and the evolution of recombination. *Theor. Popul. Biol.* 51, 134–147
23 Falconer, D.S. and Mackay, T.F.C. (1996) *Introduction to Quantitative Genetics* (4th edn), Longman Group
24 Cockerham, C.C. (1954) An extension of the concept of partitioning hereditary variance for analysis of covariances among relatives when epistasis is present. *Genetics* 39, 859–882
25 Kempthorne, O. (1955) The theoretical values of correlations between relatives in random mating populations. *Genetics* 40, 153–167
26 Kuhn, T. (1962) *The Structure of Scientific Revolutions*, University of Chicago Press

27 Wimsatt, W.C. (1974) Complexity and organization. *Philos. Sci. Assoc.* 1, 67–86
28 Lloyd, E.A. (1988) *The Structure and Confirmation of Evolutionary Theory*, Princeton University Press
29 Levins, R. (1968) *Evolution in Changing Environments*, Princeton University Press
30 Felsenstein, J. (1965) The effect of linkage on directional selection. *Genetics* 52, 349–363
31 Lewontin, R.C. (1974) *The Genetic Basis of Evolutionary Change*, Columbia University Press
32 Hill, W.G. and Robertson, A. (1966) The effect of linkage on the limits to artificial selection. *Genet. Res.* 8, 269–294
33 Barton, N.H. (1995) Linkage and the limits to natural selection. *Genetics* 140, 821–841
34 Arnold, S.J. and Wade, M.J. (1984) On the measurement of selection in natural and laboratory populations: theory. *Evolution* 38, 709–719
35 Wade, M.J. and Kalisz, S. (1990) The causes of natural selection. *Evolution* 44, 1947–1955
36 Wolf, J.B. and Wade, M.J. On the assignment of fitness to parents and offspring: whose fitness is it and when does it matter? *J. Evol. Biol.* (in press)
37 Willis, J.H. (1993) Effects of different levels of inbreeding on fitness components in *Mimulus guttatus*. *Evolution* 47, 864–876
38 Wimsatt, W.C. (1984) Reductionistic research strategies and their biases in the units of selection controversy. In *Conceptual Issues in Evolutionary Biology* (1st edn) (Sober, E., ed.), pp. 142–183, Massachusetts Institute of Technology Press
39 Cartwright, N. (1983) *How the Laws of Physics Lie*, Oxford University Press
40 Williams, G.C. (1966) *Adaptation and Natural Selection: A Critique of Some Current Evolutionary Thought*, Princeton University Press
41 Wade, M.J. (1992) Sewall Wright: gene interaction and the Shifting Balance Theory. In *Oxford Surveys in Evolutionary Biology* (Vol. 8) (Futuyma, D. and Antonovics, J., eds), pp. 35–62, Oxford University Press
42 Griesemer, J.R. (1994) Tools for talking. Human nature, Weismannism, and the interpretation of genetic information. In *Are Genes Us? The Social Consequences of the New Genetics* (Cranor, C.F., ed.), pp. 83–88, Rutgers University Press
43 Oyama, S. (1985) *The Ontogeny of Information: Developmental Systems and Evolution*, Cambridge University Press (2000 reprint, Duke University Press)
44 Wimsatt, W.C. (1986) Forms of aggregativity. In *Human Nature and Natural Knowledge* (Donogan, A. *et al.*, eds), pp. 259–291, Reidel Publishing Company
45 Wimsatt, W.C. (1997) Functional organization, functional analogy and functional inference. *Evol. Cognit.* 3, 102–132
46 Buss, L.W. (1985). The uniqueness of the individual revisited. In *Population Biology and Evolution of Clonal Organisms* (Jackson, J.B.C. *et al.*, eds), pp. 467–505, Yale University Press
47 Buss, L.W. (1987) *The Evolution of Individuality*, Princeton University Press
48 Gilbert, S. *et al.* (1996) Resynthesizing evolutionary and developmental biology. *Dev. Biol.* 173, 357–372
49 Raff, R.A. (1996) *The Shape of Life. Genes, Development and the Evolution of Animal Form*, University of Chicago Press