EVOLUTION

INTERNATIONAL JOURNAL OF ORGANIC EVOLUTION

PUBLISHED BY

THE SOCIETY FOR THE STUDY OF EVOLUTION

Vol. 59 **June 2005** **No. 6**

Evolution, 59(6), 2005, pp. 1165–1174

PERSPECTIVE: SIGN EPISTASIS AND GENETIC CONSTRAINT ON EVOLUTIONARY TRAJECTORIES

DANIEL M. WEINREICH,[1,2] RICHARD A. WATSON,[1,3] AND LIN CHAO[4]

[1]*Department of Organismic and Evolutionary Biology, Harvard University, 16 Divinity Avenue, Cambridge, Massachusetts 02138*

[2]*E-mail: dmw@post.harvard.edu*

[4]*Division of Biology, University of California at San Diego, 9500 Gilman Drive, La Jolla, California 92093*

Abstract.—Epistasis for fitness means that the selective effect of a mutation is conditional on the genetic background in which it appears. Although epistasis is widely observed in nature, our understanding of its consequences for evolution by natural selection remains incomplete. In particular, much attention focuses only on its influence on the instantaneous rate of changes in frequency of selected alleles via epistatic contribution to the additive genetic variance for fitness. Thus, in this framework epistasis only has evolutionary importance if the interacting loci are simultaneously segregating in the population. However, the selective accessibility of mutational trajectories to high fitness genotypes may depend on the genetic background in which novel mutations appear, and this effect is independent of population polymorphism at other loci. Here we explore this second influence of epistasis on evolution by natural selection. We show that it is the consequence of a particular form of epistasis, which we designate sign epistasis. Sign epistasis means that the sign of the fitness effect of a mutation is under epistatic control; thus, such a mutation is beneficial on some genetic backgrounds and deleterious on others. Recent experimental innovations in microbial systems now permit assessment of the fitness effects of individual mutations on multiple genetic backgrounds. We review this literature and identify many examples of sign epistasis, and we suggest that the implications of these results may generalize to other organisms. These theoretical and empirical considerations imply that strong genetic constraint on the selective accessibility of trajectories to high fitness genotypes may exist and suggest specific areas of investigation for future research.

Key words.—Compensatory mutations, Fisher, functional epistasis, genetic recombination, statistical epistasis, Wright's fitness landscape.

Received April 28, 2004. Accepted March 12, 2005.

Evolutionary genetics seeks to understand how mutations affect organismal phenotype and ultimately their influence on fitness, or lifetime reproductive success. Genetic loci (i.e., units of Mendelian inheritance) commonly exhibit functional interactions, implying that the effect of a mutation on fitness (or any other phenotype) may be dependent on the alleles present at other loci in the genome. Such interactions are called epistasis and are an intrinsic property of the organism's mapping from genotype to phenotype. Epistasis is widely observed in nature, but to date the evolutionary implications of epistasis for fitness have received an incomplete theoretical analysis owing in part to the way the problem was first formulated in the modern synthesis between Darwinism and Mendelism.

R. A. Fisher (1930) felt that because organisms are generally ''marvelously and intricately adapted'' to their surroundings, natural selection's primary effect was the short-term response to continual biotic and abiotic deterioration in the environment. He asked how natural selection would cause allele frequencies at polymorphic loci to change to offset such environmental decline, and his fundamental theorem of natural selection gives the answer: each generation, population mean fitness increases by an amount exactly equal to what he called the additive component of standing genetic variance for fitness. Because Fisher's fundamental theorem only addresses the instantaneous selective response driven by genetic variation now segregating in an evolving population, in this framework epistasis only has evolutionary relevance if the functionally interacting loci are simultaneously polymorphic. This evolutionary influence of epistasis, dependent on the population's current allelic composition, is often called ''statistical epistasis'' (Fenster et al. 1997). Moreover, under the common assumption that alleles are in linkage equilibrium, even statistical epistasis among segregating alleles does not contribute to additive genetic variance and so only constitutes a form of statistical noise in the system.

Sewall Wright corroborated the mathematics underlying Fisher's fundamental theorem (Wright 1930; Provine 1986),

[3] Present address: Electronics and Computer Science, Southampton University, Highfield, Southampton SO17 1BJ, United Kingdom.

and both workers recognized the evolutionary importance of occasional novel beneficial mutations in addition to the significance of changes in allele frequency at segregating loci. Unlike Fisher, however, Wright was also interested in the evolutionary consequences if the fitness effect of a novel mutation varied as a function of the genetic background on which it appeared, whether that background was polymorphic or not. In the evolutionary genetics literature, this more inclusive concept of epistasis (i.e., independent of allele frequency at the interacting loci) is sometimes referred to as ''physiological'' (Cheverud and Routman 1995) or ''functional'' (Hansen and Wagner 2001) epistasis to distinguish it from statistical epistasis (with its dependence on simultaneous segregation of interacting loci), and this contrast has been highlighted by many authors (e.g., Wade 1992; Whitlock et al. 1995; Fenster et al. 1997; Phillips 1998; Brodie 2000). In particular, unlike statistical epistasis, which itself changes with changing allele frequencies and thus is predictive only of short-term evolutionary dynamics, functional epistasis can also be important on the much longer time-scale of the mutational process, over which a succession of novel beneficial mutations may appear and reach fixation.

Wright (1932) drew particular attention to the possibility that functional epistasis might give rise to genotypes in which individual mutations at all loci are deleterious in spite of the existence of higher fitness genotypes differing by mutations at multiple loci. A hypothetical population fixed for such a genotype would then be stuck on a local peak on the fitness landscape because no mutational trajectory (i.e., succession of individual mutations) would be selectively favored, and Wright felt that under these circumstances natural selection alone could not increase fitness (but see Carter and Wagner 2002; Iwasa et al. 2004; Weinreich and Chao 2005). Note that this observation is outside the scope of statistical epistasis because it does not depend on the interacting loci being simultaneously polymorphic. Yet, it is functional epistasis itself that is directly responsible for the lack of selectively accessible mutational trajectories to higher fitness genotypes.

Thus, statistical epistasis only captures some of the evolutionary importance of functional interactions between loci. However, statistical epistasis can readily be measured in natural populations (e.g., Falconer 1994). In contrast, the existence in nature of functional epistasis sufficient to give rise to local peaks on the fitness landscape is technically nearly impossible to show (Whitlock et al. 1995) because it requires examination of all genotypes in the mutational neighborhood of a putative peak, many of which are unlikely to be present in any population. Additionally, the existence of such peaks is theoretically controversial (Gavrilets 2003, 2004), and the problem of evolutionary escape by populations stuck at such a peak have not been satisfactorily resolved (e.g., Coyne et al. 1997, 2000; Wade and Goodnight 1998; Goodnight and Wade 2000). These empirical and theoretical considerations at least in part explain why many workers have followed Fisher and focused on the evolutionary consequences of statistical epistasis among segregating alleles.

To summarize, within the framework of the modern synthesis our understanding of the evolutionary implications of epistasis for fitness is incomplete. One school of thought focuses on statistical epistasis, asserting that only fitness interactions among segregating alleles matter, but this view excludes the important possibility of epistatic effects between new mutations and a monomorphic genetic background. The alternative approach addresses these cases but historically has focused on the existence of multiple peaks on the fitness landscape, a hypothesis not easily tested empirically.

Here we explore the broader theoretical question of whether and how functional epistasis for fitness can constrain the selective accessibility of mutational trajectories to higher fitness genotypes. After carefully defining Wright's fitness landscape and the model of evolution by natural selection best suited for this problem, we discuss the circumstances under which one or more mutational trajectories leading to the genotype of highest fitness might be selectively inaccessible (of which Wright's notion of local fitness peaks is a special case in which none are accessible). We show that such limitations on selectively accessible trajectories arise only as a consequence of a particular form of functional epistasis perhaps not widely appreciated, which we denote sign epistasis. Under sign epistasis, mutations are beneficial on some genetic backgrounds and deleterious on others; in other words, the sign of the fitness effect of such a mutation is conditional on genetic background.

Additionally we show that sign epistasis exists in nature. Circumstantial evidence exists for sign epistasis in a wide variety of organisms and much explicit data now exist as a consequence of recent technical innovations in microbial experimental systems that permit more exhaustive empirical exploration of fitness values in modest mutational neighborhoods. We briefly review this literature and use it to demonstrate that sign epistasis is common in these organisms. We suggest further that the implications of empirical results derived largely from microbes may nevertheless be quite general. Finally, we touch on the theoretical implications of genetic recombination for evolutionary trajectories. We conclude with a summary of open questions and prospects suggested by these facts.

Evolution by Natural Selection on Wright's Fitness Landscapes

Wright (1932) introduced the fitness landscape to represent evolution by natural selection in the presence of arbitrary epistatic interactions between loci. In fact, Wright (e.g., Wright 1982) advocated two related conceptions of the fitness landscape (Provine 1986), which we denote the genotypic fitness and population mean fitness landscapes. We briefly describe both to motivate our focus on the former, to illustrate that modeling evolution by natural selection on these two landscapes recapitulates the contrast between Fisher's and Wright's views of evolution outlined above, and to clarify the very close connection between the two.

Every possible haploid genotype containing L biallelic loci can be uniquely represented by a point in a discrete space of dimensionality L in which each dimension represents a different locus and may assume two values corresponding to the two alleles defined for that locus. We denote this the genotype sequence space and for each point, adjacent points in space represent genotypes that differ by a single point mutation (Maynard Smith 1970). We shall disregard gross

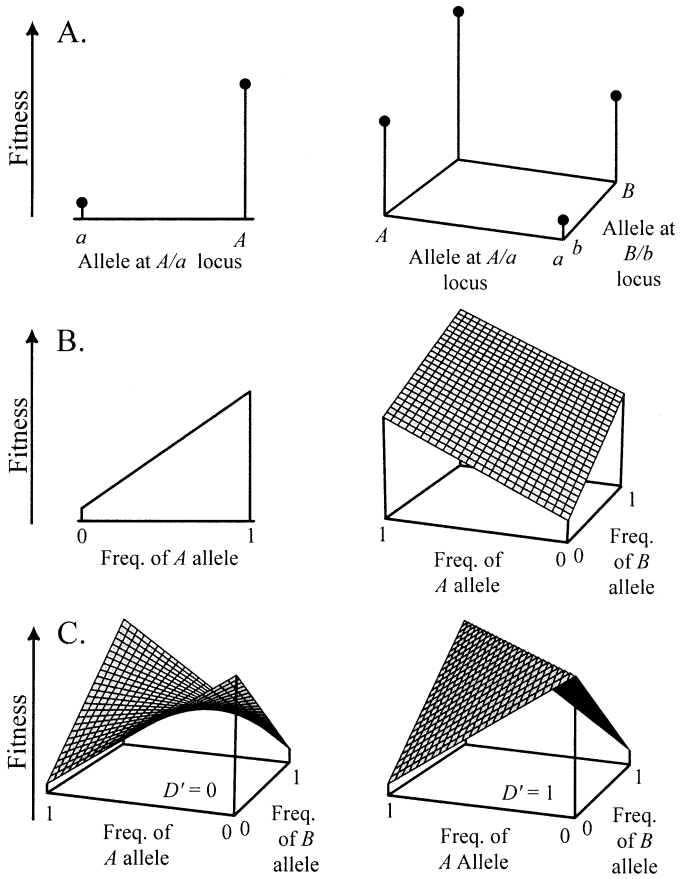

FIG. 1. Fitness landscapes for haploid genomes comprising L biallelic loci are (L + 1)-dimensional projections of fitness values. (A) Genotypic fitness landscape, created by projecting the fitness value of each genotype arranged in L-dimensional genotype sequence space into the (L + 1)th dimension. Left: L = 1; right: L = 2. In each case, the genotype sequence space is represented below and the genotypic fitness values are projected upward. Because genotype sequence space is discrete, the genotypic fitness landscape is discrete. (B) Population mean fitness landscape defined for genotypic fitness values in panel A, created by plotting population mean fitness values as a function of allele frequencies. Left: L = 1; right: L = 2. In each case, the allele frequency space is represented below and the population mean fitness values are projected upward. (C) Population mean fitness in the presence of epistasis. Two values of D', a measure of linkage disequilibrium, yield two population mean fitness landscapes for the same genotypic fitness landscape.

mutational processes such as insertion/deletions and inversions, whose adjacencies are not easily represented in this space. (Formally genotype sequence space is an L-dimensional hypercube whose vertices represent genotypes and whose edges represent point mutations.) The genotypic fitness landscape is created by projecting the fitness value of each possible genotype represented in genotype sequence space into the (L + 1)th dimension (Fig. 1A). We assume that genotypic fitness values are constant, and therefore disregard frequency dependent selection, dominance, and temporally changing environment throughout. Because genotype sequence space is discrete, the genotypic fitness landscape is also discrete.

To explore the selective accessibility of mutational trajectories through genotype sequence space, we follow Gillespie (1984) and employ the strong selection/weak mutation (SSWM) model. Here, populations are regarded as genetically monomorphic and so can be represented by a single point in genotype sequence space. Occasional point mutations generate novel progeny corresponding to adjacent points in sequence space, and natural selection and genetic drift are assumed to instantly fix or eliminate the mutant genotype. Because natural selection favors higher fitness genotypes, an evolving population will tend to follow the local gradient on the genotypic fitness landscape through sequence space to mutationally nearby regions of higher fitness. Thus, the genotypic fitness landscape is well suited to our (and Wright's) interest in exploring constraints on the selective accessibility of evolutionary trajectories.

The SSWM model represents populations as single points in genotype sequence space and thus is unable to model either genetic recombination (which requires polymorphism to generate novelty) or changing allele frequencies. A polymorphic population occupies a distribution of points in this space, and in a later section we briefly review a treatment that explores the consequences of recombination in this framework. Wright's own approach to modeling changes in allele frequency was to represent a population by the point corresponding to its center of mass in genotype sequence space. Note that this point lies in another L-dimensional space, where each dimension now represents population allele frequency at a different biallelic locus. We denote this the allele frequency space and projection of the mean fitness over this space into the (L + 1)th dimension yields the population mean fitness landscape (Fig. 1B). Here, beginning from some particular population composition, natural selection will adjust allele frequencies in such a manner as to increase population mean fitness, and so an evolving population will follow the local gradient on the population mean fitness landscape to regions of higher fitness. Thus, the population mean fitness landscape readily addresses the questions framed by Fisher's fundamental theorem about selective changes to allele frequencies and resultant increases in population mean fitness.

The close connection between Wright's two fitness landscapes is apparent when one notes that, although the population mean fitness landscape is continuous when population size is infinite, it is otherwise discrete and indeed converges to the genotypic fitness landscape as population size goes to one. Thus, the population mean fitness landscape can be seen simply to fill the interstices of the genotypic fitness landscape (cf. Figs. 1A and 1B) at a granularity equal to the reciprocal of population size. However, for fixed genotypic fitness values the mapping from allele frequency space to mean fitness landscape is not unique in the presence of epistasis (Fig. 1C). This one-to-many mapping from allele frequency space to population mean fitness landscape reflects the fact that, while all information about linkage disequilibrium between alleles across loci is present in the representation of a population as a distribution of points in genotype sequence space, it is lost in the transformation to a single point in allele frequency space. Therefore, long-term evolutionary prediction based on the population mean fitness landscape is difficult in the presence of epistasis (Weinreich and Chao 2005). Recall, how-

ever, that this is precisely the possibility that motivates our interest in fitness landscapes.

Genetic Constraints on Evolutionary Trajectories and Epistasis

Natural selection causes evolving populations to follow mutational trajectories that move them upward on the genotypic fitness landscape. However, as recognized early by Wright (Provine 1986), epistasis can cause the fitness landscape to possess ridges and valleys that constrain the ability of evolving populations to reach the genotype of highest fitness. We now rigorously develop this connection.

Consider arbitrary genotypes x and y in genotype sequence space. Without regard to fitness values, we can define a mutational trajectory through sequence space from x to y as mutations at a succession of loci that carry a population at x first to some mutationally adjacent genotype x', thence to x'', and so on, eventually reaching genotype y. Writing $D(x, y)$ to represent the set of loci at which genotypes x and y differ in allelic state, the shortest trajectories between x to y are those in which mutations occur exactly once at each locus in $D(x, y)$ and at no others. (Allowing mutational reversions gives rise to longer trajectories and may include mutations at loci not in $D(x, y)$). If $|D(x, y)|$ represents the number of loci in $D(x, y)$, these shortest trajectories are thus $|D(x, y)|$ mutations long; moreover, because mutations at the members of $D(x, y)$ may occur in any order, there are $|D(x, y)|!$ shortest trajectories between x and y.

A fitness landscape specifies the fitness value for each possible genotype. Given an arbitrary fitness landscape f, let $P(f)$ represent the peak or highest fitness genotype on f. What is the capacity of natural selection to move a population from arbitrary genotype x along some mutational trajectory to $P(f)$? We designate a trajectory as selectively accessible if and only if each successive mutation along the trajectory is beneficial, that is, genotypes on this trajectory are monotonically increasing in fitness (but see Carter and Wagner 2002; Iwasa et al. 2004; Weinreich and Chao 2005). If $B(x, f)$ represents the set of loci at which beneficial mutations are possible in genotype x on fitness landscape f, then a selectively accessible trajectory between x and $P(f)$ consists first of a mutation at some locus in $B(x, f)$ whose fixation moves the population to genotype x', then by a mutation at a locus in $B(x', f)$ and so on, and our interest is in the abundance and length of selectively accessible mutational trajectories between arbitrary x and $P(f)$.

All shortest mutational trajectories between arbitrary genotype x and $P(f)$ are selectively accessible if and only if all members of $D(y, P(f))$ are also members of $B(y, f)$ for all y written $D(y, P(f)) \subseteq B(y, f)$. To see this, consider arbitrary genotype x; the condition $D(y, P(f)) \subseteq B(y, f)$ for all genotypes y implies that all first mutational steps along all shortest trajectories between x and $P(f)$ are beneficial. Moreover, because the condition applies equally to each genotype visited along each shortest mutational trajectory, all such trajectories between x and $P(f)$ are selectively accessible. Thus, if $D(y, P(f)) \subseteq B(y, f)$ for all genotypes y, then all shortest mutational trajectories from all genotypes to $P(f)$ are selectively accessible, and the logical converse follows trivially from the definitions of $D(x, y)$ and $B(x, f)$.

Assuming only that $D(y, P(f)) \subseteq B(y, f)$ for all genotypes y, one might suppose that for some genotype x there could also exist loci in $B(x, f)$ which are not in $D(x, P(f))$, and thus that additionally, some longer mutational trajectories from x to $P(f)$ may also be selectively accessible. We now show that this is not possible by demonstrating that if such a locus exists for one genotype, it necessarily comes at the expense of the selective accessibility of some shortest mutational trajectories from another genotype. Put another way, if on landscape f all shortest mutational trajectories from all genotypes to $P(f)$ are selectively accessible, then no additional, longer mutational trajectories to $P(f)$ can be selectively accessible.

To see this, assume that $D(y, P(f)) \subseteq B(y, f)$ for all y, and suppose further that for genotype x there additionally exists a locus l that is a not member of $D(x, P(f))$ but is a member of $B(x, f)$. Let x' be the genotype produced by mutation at locus l in genotype x. Because l is not in $D(x, P(f))$, it must be in $D(x', P(f))$, since the allele present at l in $P(f)$ cannot be present in both x and x'. Furthermore because l is in $B(x, f)$, it cannot be in $B(x', f)$, since x and x' cannot both be fitter than the other. Therefore, because l is a member of $B(x, f)$ but not of $D(x, P(f))$, it must be a member of $D(x', P(f))$ but not of $B(x', f)$, contradicting the assumption that $D(y, P(f)) \subseteq B(y, f)$ for all y. This demonstrates that under the assumption that $D(y, P(f)) \subseteq B(y, f)$ for all y, no locus l can exist which is a member of $B(x, f)$ but not of $D(x, P(f))$ for any genotype x, or algebraically, that if $D(y, P(f)) \subseteq B(y, f)$ for all y then $D(y, P(f)) = B(y, f)$ for all y. (The logical converse of this statement flows immediately from set theory: if two sets are equal then each is also a subset of the other.)

To summarize, all shortest mutational trajectories from all genotypes to $P(f)$ will be selectively accessible if and only if $D(y, P(f)) = B(y, f)$ for all y, and this condition has the corollary implication that from all genotypes only the shortest mutational trajectories to the peak can be selectively accessible. On such a landscape, all mutations that take any genotype closer in sequence space to the highest fitness genotype are beneficial, and all beneficial mutations take a genotype closer to the peak. We adopt this as a natural and intuitive definition for a fitness landscape lacking genetic constraint on selectively accessible mutational trajectories.

Given these facts, at the scale of individual mutational effects there are logically only two ways that the selective accessibility of mutational trajectories to the peak may be constrained. First, for genotype x there may exist a locus l which is a member of $D(x, P(f))$ but not of $B(x, f)$, reducing the number of selectively accessible trajectories to $P(f)$. Alternatively, for genotype x there may exist a locus l which is a member of $B(x, f)$ but not of $D(x, P(f))$. In this case, natural selection will favor a mutation that nevertheless moves the population farther from $P(f)$. Although this second possibility may increase the number of selectively accessible mutational trajectories from x to $P(f)$, such additional trajectories are necessarily two mutations longer than $|D(x, P(f)|$. This follows because the mutation at l will require eventual reversion before the trajectory reaches $P(f)$, since l is not in $D(x, P(f))$. Moreover, as we have just shown, such additional selectively accessible trajectories from x come only at the

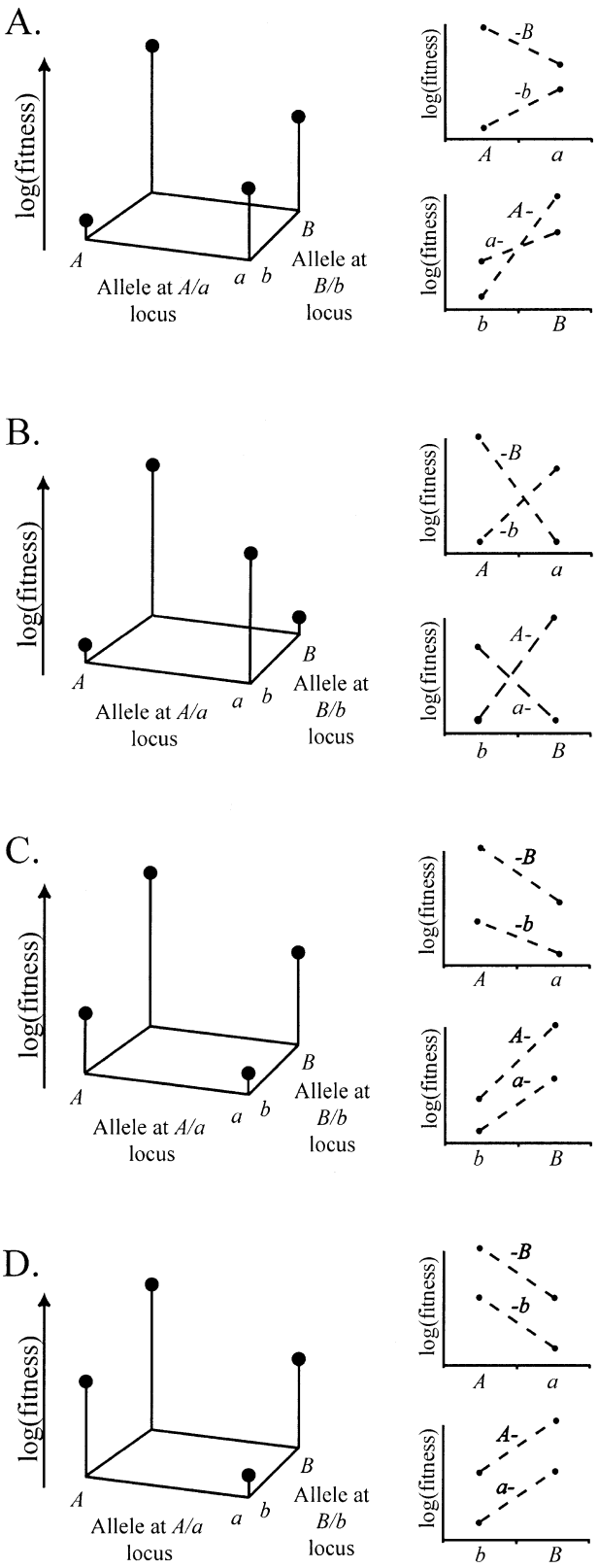

FIG. 2. Sign and magnitude epistasis. To the left in each panel, genotype sequence space is represented on the x-y plane and the log of each genotype's fitness is represented on the z-axis. To the right of each genotypic fitness landscape, two cross-sections through the landscape are shown to illustrate the fitness effect of mutation at each locus on each background: mutations at the *A*/*a* locus on each background appear above; mutations at the *B*/*b* locus on each background, below. Because log fitness values are used, the slope of the fitness landscape cross-sections represents the selection coefficient acting on the underlying mutation. (A) Sign epistasis: the $a \rightarrow A$ mutation is beneficial on the *-B* background but deleterious on the *-b* background. The $b \rightarrow B$ mutation is beneficial on both backgrounds (i.e., on *a-* and *A-*) but necessarily exhibits magnitude epistasis. (B) Multiple peaks: the $a \rightarrow A$ mutation is beneficial on the *-B* background and deleterious on the *-b* background; the $b \rightarrow B$ mutation is beneficial on the *A-* background and deleterious on the *a-* background. (C) Magnitude epistasis: the selection coefficient acting on the $a \rightarrow A$ mutation is larger on the *-B* background than on the *-b* background, and the selection coefficient acting on the $b \rightarrow B$ mutation is larger on the *A-* background than on the *a-* background. However, both mutations are beneficial on each background. (D) No epistasis: the selection coefficient acting on the $a \rightarrow A$ mutation is independent of the allele present at the *B*/*b* locus, and the selection coefficient acting on the $b \rightarrow B$ mutation is independent of the allele present at the *A*/*a* locus. Note that only a single fitness value on this landscape differs from those in panel C, yet magnitude epistasis is now absent at both loci.

expense of selectively accessible trajectories to $P(f)$ from neighboring genotypes in sequence space.

The connection between this definition of genetic constraint and epistasis may be seen as follows. The condition that $B(y, f) = D(y, P(f))$ for all genotypes y is equivalent to saying that on all genotypes y the identity of the fitter allele at every locus is the allele found at that locus in $P(f)$. Thus, the selective accessibility of mutational trajectories on fitness landscape f from all genotypes to $P(f)$ is unconstrained if and only if the identity of the fitter allele at every locus is unconditioned on genotype. Note that this does not require that all epistasis be absent: if only the magnitude of the fitness effect of a mutation varies with genetic background, membership in $B(y, f)$ will remain unchanged for all genotypes y and so no constraint on selectively accessible trajectories to $P(f)$ is introduced. But for the landscape to lack genetic constraint on selectively accessible trajectories, we require that the sign of the fitness effect of all mutations be unconditioned on genetic background.

It is convenient to formalize this central contrast between forms of epistasis, and we designate the alternatives sign and magnitude epistasis. Sign epistasis at a locus means that a mutation there is beneficial on some genotypic backgrounds and deleterious on others, or equivalently that the sign of the fitness effect of a mutation is under epistatic control (Figs. 2A, B). We contrast sign epistasis with magnitude epistasis, in which mutations are unconditionally beneficial or unconditionally deleterious, although the magnitude of their effect may depend on genotypic background (cf. Figs. 2C and 2D). We have shown that genetic constraint on selectively accessible mutational trajectories to $P(f)$ exists if and only if sign epistasis is present on f. The dual manifestations of constraint introduced by sign epistasis (the loss of selectively accessible shortest trajectories to the peak and gain of selectively accessible longer trajectories) are illustrated in the contrast between Figure 2C (which lacks sign epistasis) and Figure 2A (in which sign epistasis is present at the *A*/*a* locus). The sign epistasis in Figure 2A eliminates one of the selectively accessible shortest trajectories from the *ab* genotype but si-

multaneously introduces a novel, albeit longer selectively accessible trajectory from *Ab* via *ab*.

A brief examination of previous classifications of epistasis illustrates the novelty of the distinction between sign and magnitude epistasis. Positive and negative epistasis, defined in models used to explore the evolutionary consequence of genetic recombination (reviewed in Kondrashov 1988; Peters and Lively 2000) are forms of magnitude epistasis. Unidimensional and multidimensional epistasis (Kondrashov and Kondrashov 2001) can have magnitude epistasis alone or can have both, although models of unidimensional epistasis commonly assume only magnitude epistasis (Kondrashov 1988) and the exemplar model of multidimensional epistasis (Kondrashov and Kondrashov 2001) also possesses sign epistasis. Finally, in the traditional two-locus quantitative genetic parameterization, if the additive genetic variance at each locus is sufficiently small, then additive-by-additive variance gives rise to sign epistasis. Additive-by-additive variance together with relatively larger additive effects gives rise only to magnitude epistasis between loci.

However, the consequences of sign epistasis on the selective accessibility of the shortest mutational trajectories to $P(f)$ have been implicitly exploited in the theoretical literature. For example, the reduction in the number of selectively accessible trajectories from some genotypes as a consequence of sign epistasis underlies the fitness landscape presented by Kondrashov and Kondrashov (2001), and the capacity of sign epistasis to lengthen trajectories is emphasized in the root2path landscape of Horn et al. (1994; a rendering of this landscape appears on the cover of this issue). Additionally, although interest in neutral network fitness landscapes (Lipman and Wilbur 1991; Schuster et al. 1994; Huynen et al. 1996; van Nimwegen and Crutchfield 2000) traditionally focuses on the existence of neutral mutational pathways, the rarity of beneficial mutations on such fitness landscapes appears to be a consequence of sign epistasis (not shown). Indeed, on an exemplar neutral network (the ''royal staircase with ditches'' fitness landscape of van Nimwegen and Crutchfield 2000) beneficial mutations are beneficial only on a single genotypic background and deleterious on all other backgrounds.

Finally, the most well-known theoretical example of genetic constraint on the selective accessibility of mutational trajectories, the case of fitness landscapes possessing multiple, mutationally isolated fitness peaks (e.g., Wright 1932; Kauffman 1993), depends on sign epistasis. To see this, note that in order for genotype x to lack beneficial mutations in spite of the existence of higher fitness genotypes elsewhere on f, it must be that $|B(x, f)| = 0$ while $|D(x, P(f))| > 0$. Because we have seen that for all fitness landscapes on which sign epistasis is absent $B(y, f) = D(y, P(f))$ for all y, it follows that f possesses sign epistasis. In this case again, sign epistasis also increases the number of beneficial mutations available to some mutational neighbors of local peak x, but because these novel beneficial mutations lead to x, no additional selectively accessible trajectories to $P(f)$ are created. For example, in Figure 2B, sign epistasis increases the number of beneficial mutations available to genotypes *aB* and *Ab* when compared to Figure 2C; however, this does not give rise to novel selectively accessible trajectories to *AB*, the highest fitness genotype.

To summarize, genetic constraint on the selective accessibility of mutational trajectories to high-fitness genotypes arises only as a consequence of sign epistasis, in which case on some genetic background(s) mutation to an allele present in a high-fitness genotype is deleterious. At the same time, on other background(s) mutation at a locus already carrying the allele present in a high-fitness genotype is beneficial. Put another way, in the presence of sign epistasis the local fitness gradient along some mutational trajectories is inconsistent with the gradient at larger mutational scales; this is necessary for the existence of ridges and valleys on the genotypic fitness landscape and represents the root source of such genetic constraint on evolution by natural selection.

Sign Epistasis in Nature

Sign epistasis constrains the ability of evolution by natural selection to carry an evolving population to the genotype of highest fitness and can even give rise to multiple peaks on the fitness landscape. This motivates an examination of its prevalence in nature. As noted above, the question of multiple peaks on the fitness landscape of any organism has been a difficult one to address experimentally because of the requirement to exhaustively survey a local region of genotype sequence space (to show that $|B(x, f)| = 0$ for some genotype x). Fortunately, demonstration of sign epistasis in an organism requires only that for genotypes x and y, where the fitness of y exceeds that of x, there exists some locus l that is a member of $B(x, f)$ but not of $D(x, y)$ or visa versa.

Compensatory mutations fix in evolving populations in response to fitness loss due to a mutation elsewhere in the genome. Mutations are regarded as being conditionally beneficial and thus compensatory if they do not appear in control populations evolving in the absence of the changed genetic background. This interpretation relies on the population genetic hypothesis that, were the mutation unconditionally beneficial, the control population would have sampled enough mutants to have found and fixed it (Burch and Chao 1999). Thus, unless such mutations are strictly neutral (or possibly only very slightly beneficial) in the control population, they are deleterious and so exhibit sign epistasis. Several examples of such conditional fitness improvements have been reported, including the bacteriophages ϕ6 (Burch and Chao 1999) and ϕX174 (Poon and Chao 2005), HIV-1 (reviewed in Quiñones-Mateu and Arts 2001), humans (Kondrashov et al. 2002; Gao and Zhang 2003; Kern and Kondrashov 2004), and *Drosophila* (Kulathinal et al. 2004). In the metazoan cases, a phylogenetic approach was employed, and an appreciable fraction of deleterious point mutations in the focal species were found to be the wild-type allele in one or more other species. Here the inference is that in the other species second-site mutations ameliorate the deleterious effect observed in the focal species and make it beneficial (Kondrashov et al. 2002). Thus, they exhibit sign epistasis. Finally, in a recent maximum-likelihood-based meta-analysis of the suppressor mutation literature, Poon et al. (2005) estimated that there are an average of 11.8 compensatory mutations per deleterious mutation examined. Importantly, this figure differs only

modestly among data from viruses, prokaryotes, and eukaryotes.

A more explicit approach to detecting sign epistasis is to employ molecular manipulations such as site-directed mutagenesis to construct genotypes occupying a number of adjacent points in genotype sequence space. Recent technical innovations in microbial systems have made this approach practical and permit the assessment of fitness effects of individual mutations on two or more genetic backgrounds. Several such studies are reviewed next. Moreover, in our view, the fact that these data are derived from microbes does not undermine their generality. Rather, we suggest that the structure of epistasis acting to constrain the evolutionary trajectories of molecules encoded in microbial genomes may at least provisionally be taken as representative of such effects in any organism.

5S RNAs are structural components of ribosomes and are found in nearly all organisms. In the bacterial species *Vibrio alginolyticus* and *V. proteolyticus*, the 5S RNA genes differ at four nucleotides. Lee et al. (1997) used site-directed mutagenesis to construct all $2^4 - 2 = 14$ mutational intermediates between the 5S RNA genes of these species (Fig. 3A) and found that only five of $6! = 24$ possible mutational pathways between species did not require passage through a saddle corresponding to a sequence of reduced physiological activity (Fig. 3A). Thus, sign epistasis must exist among these mutations, since we have shown that peaks and saddles in the genotypic fitness landscape cannot exist in its absence. Similar results showing sign epistasis were observed in the analysis of the mutational pathways between the 5S RNA genes in *V. alginolyticus* and another related bacterium, *V. nervis* (Lee et al. 1997).

Molecular sign epistasis among mutations in the protease gene of HIV-1 has also been detected using site-directed mutagenesis. For example, the 10th amino acid in the HIV-1 protease is a leucine, and replacing it with an isoleucine (a mutation denoted Leu10Ile) reduces viral resistance to the protease inhibitor saquinavir. However, the same mutation increases saquinavir resistance in the presence of the Gly48Val and Leu90Met mutations, either individually or together (Mammano et al. 2000).

Both *Escherichia coli* (Schrag et al. 1997) and *Salmonella typhimurium* (Maisnier-Patin et al. 2002) readily evolve resistance to the antibiotic streptomycin. In both species, streptomycin resistance mutations slow protein synthesis, thereby lowering organismal fitness in the absence of the drug. However, this fitness loss can be offset by one of several compensatory second-site mutations (Schrag and Perrot 1996; Björkman et al. 1999). Importantly, when many of these second-site mutations were placed on their respective wild-type streptomycin-sensitive genetic backgrounds, they were found to be deleterious (Schrag et al. 1997; Maisnier-Patin et al. 2002), explicitly showing that these compensatory second-site mutations exhibit sign epistasis.

In both *E. coli* and *S. typhimurium*, the genotypic fitness landscape in the absence of streptomycin may be approximated by Figure 2B, in which the wild-type genotype is represented by *AB*, the $A \rightarrow a$ mutation represents antibiotic resistance and the $B \rightarrow b$ mutation represents the second-site mutation. It is worth noting that, although both studies

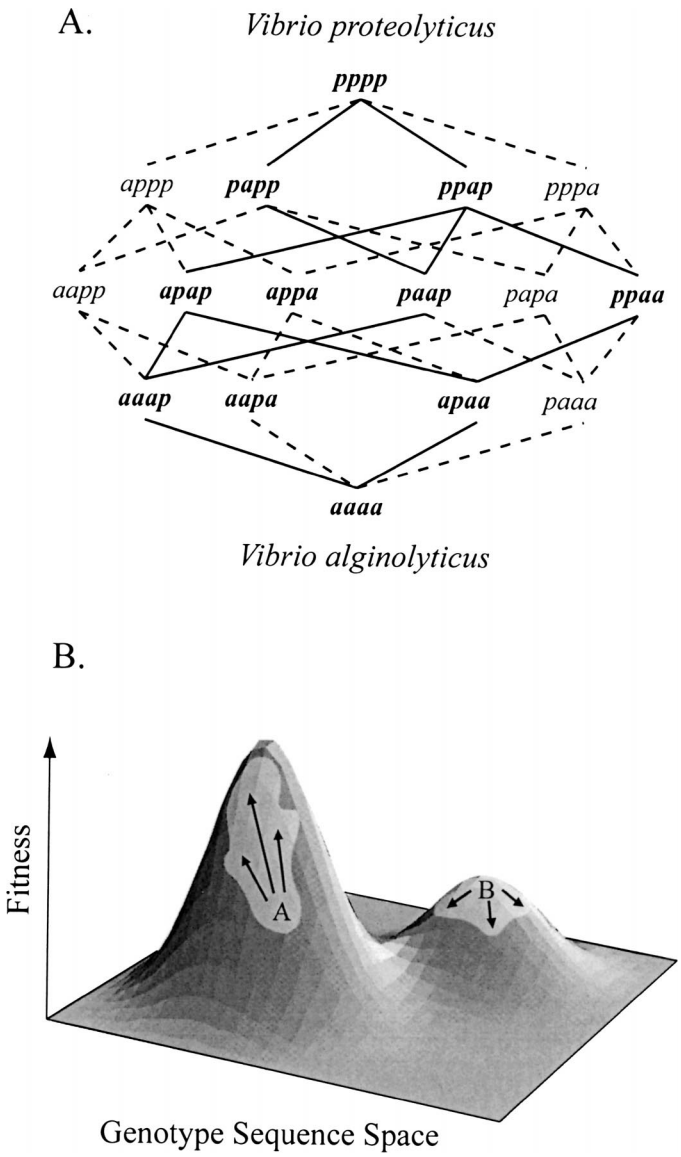

FIG. 3. Empirical genotypic fitness landscapes. (A) The 5S RNA genes of *Vibrio proteolyticus* and *V. alginolyticus* differ at four nucleotides and are otherwise identical. The two naturally occurring genes and all mutational intermediates may be represented by a string of four characters chosen from *p* and *a*, corresponding to the nucleotides observed in the *V. proteolyticus* and *V. alginolyticus* genes, respectively. Nodes in the figure represent all $2^4 = 16$ possible 5S RNA genes, with the *V. proteolyticus* gene at top and the *V. alginolyticus* gene at the bottom. Lines connect pairs of mutationally adjacent genes differing at one nucleotide. Physiologically active genes are shown in bold, and selectively accessible mutational pathways connecting the *V. proteolyticus* and *V. alginolyticus* genes are represented by solid lines. Many pathways are selectively inaccessible, demonstrating the action of sign epistasis (see text). (Reprinted with permission from Lee et al. 1997.) (B) An illustration of a portion of the genotypic fitness landscape inferred for the bacteriophage ϕ6. Axes as in Figure 1A, except that true adjacencies underlying genotype sequence space cannot be accurately represented in two dimensions. Replicate lineages founded from clone A repeatedly increased in fitness, while those founded from the closely related clone B repeatedly declined in fitness. (Reprinted with permission from Burch and Chao 2000.)

(Schrag et al. 1997; Maisnier-Patin et al. 2002) demonstrate sign epistasis, formally they cannot demonstrate the existence of isolated fitness peaks (Whitlock et al. 1995). This point is easily overlooked: mutations at additional loci define genotypes whose fitness values are in principal entirely unconstrained by the data in hand and that may thus span the putative fitness valley.

This limitation of inference was overcome recently by Burch and Chao (2000), who showed that genetically different clones of $\phi 6$ reside on different fitness peaks. Specifically, they showed that the clones evolved toward different equilibrium fitness values (Fig. 3B). Five lineages derived from clone A all increased in fitness, demonstrating that higher fitness genotypes exist. Nevertheless, five lineages derived from clone B were unable to attain this fitness (and actually declined in fitness, presumably owing to the action of genetic drift). This demonstrates that no beneficial mutations were available to clone B and suggests that it resides on a mutational peak, the existence of which implies the action of sign epistasis in that genome as shown above. Importantly, the fact that all five replicates failed to increase in fitness argues that it is unlikely that the experiment failed to sample mutations that would bridge the putative fitness saddle.

Recombination, Sign Epistasis, and Evolutionary Trajectories

We have drawn attention to the influence of epistasis on the selective availability of mutational trajectories through genotype sequence space. For these purposes, we have adopted the SSWM assumptions (Gillespie 1984) and disregarded population polymorphism. We have thereby also disregarded the effects of recombination, which requires polymorphism to generate novelty. One might ask how this approach differs from modeling the entire genome as a single locus with a very great number of alternate alleles (as in the infinite alleles model; Kimura and Crow 1964). However, we recommend the representation based on sequence space advanced here because it makes explicit the mutational adjacency between genotypes absent in the alternative, and no simple representation of evolutionary trajectories appears possible in the single-locus model. Similarly, the single-locus model cannot represent epistasis because the underlying combinations of alleles have been subsumed into single alleles at the metalocus.

How can the accessibility of evolutionary trajectories in the presence of genetic recombination be explored? To be of evolutionary importance, recombination requires linkage disequilibrium, since otherwise it will not change the frequencies at which alleles coexist in the population. However, representation of linkage disequilibrium is problematic in allele frequency space as seen in Figure 1C. If instead populations are represented as a distribution of points in genotype sequence space, this difficulty is avoided because the frequency of each genotype in the population is preserved in the model. In this manner, one can begin to address the influence of genetic recombination on the selective accessibility of evolutionary trajectories in the presence of arbitrary forms of epistasis.

This is a long-standing problem in the computer science literature (Holland 1975), and some results are known. Recombination samples genotype sequence space qualitatively differently than does mutation (Gitchoff and Wagner 1996; Wagner and Stadler 1998) and may produce progeny that are not adjacent to either parental genotype in sequence space (Watson 2005), permitting evolutionary jumps through sequence space. Consequently, the action of recombination may profoundly affect the interplay between epistasis, fitness landscape topography, and selectively accessible trajectories. For example, fitness landscapes have been defined that lack any selectively accessible mutational trajectories (as a consequence of sign epistasis giving rise to multiple peaks) but that have one or more accessible trajectories in the presence of genetic recombination (Jansen and Wegener 2001; Watson 2001, 2004). The absence of any selectively accessible mutational trajectories can mean that the expected time for natural selection to carry an asexual evolving population to the maximum fitness genotype grows exponentially in genome size, essentially because sign epistasis can render the local fitness gradient along mutational trajectories entirely uninformative at longer scales. However, because recombination may create one or more selectively accessible trajectories to the peak genotype, the expected time to the fittest peak on these landscapes is only linear in genome size for a sexual population.

Much work remains to better understand the effect of sign epistasis on the evolution of recombining populations, but what is clear is that these results are qualitatively different than the interaction between recombination and magnitude epistasis previously described (e.g., Peters and Lively 2000). First, recombination can generate selectively accessible trajectories on landscapes where none are available to an evolving asexual population. In contrast, we have shown that in models assuming only magnitude epistasis mutation will never fail to yield selectively accessible trajectories. Additionally, unlike mutation-mediated perturbations in sequence space, recombinational perturbations are not necessarily random with respect to fitness because the genetic material brought together by sex has been subject to natural selection in other individuals in previous generations (Watson 2005). Finally, some results highlight the significance of loci that are epistatically dependent also being physically linked on the chromosome and show a benefit for recombination that disappears when the physical position of loci on the chromosome are reordered (Watson 2004, 2005).

Prospects

Many questions remain. How common is sign epistasis in nature? Compensatory mutations are recognized only after becoming beneficial and fixing by natural selection. What is the frequency of mutations that are only cryptically beneficial? Site-directed mutagenesis in microbial genes such as the *Vibrio* 5S RNA (Lee et al. 1997) and TEM-1 in *E. coli* (Hall 2002) appear to offer an excellent opportunity to address these questions. In these systems all possible combinations of alleles at a small number of loci can be generated by site-directed mutagenesis, their fitness (or other phenotype) measured, and thus the frequency of sign epistasis in an unbiased sample of mutations may be determined. In this

manner, the selective accessibility of all mutational pathways through a portion of genotype sequence space can be characterized. Comparable experimental work in multicellular organisms is more difficult, although we suggest that it is at present unclear whether the structure of epistatic interactions among loci in microbes should be qualitatively different than that in multicellular organisms (Poon et al. 2005). Moreover, indirect evidence of sign epistasis from phylogenetic comparisons in metazoans is compelling (Kondrashov et al. 2002; Gao and Zhang 2003; Kulathinal et al. 2004), and this latter work offers the opportunity to localize and identify the functionally interacting loci that give rise to the phenomenon in these organisms. There is theoretical reason to suppose that if such interactions are tightly genetically linked they are most likely to be selectively favored (Kimura 1985; Stephan 1996; Carter and Wagner 2002; Weinreich and Chao 2005), and interactions within genes appear to be more likely than between genes (Poon et al. 2005), as one might intuitively suspect. Transgenics could thus be employed in *Drosophila* (e.g., Siegal and Hartl 1996) to test the hypothesis that the ameliorating second-site mutation is in the same gene as the deleterious mutation, the effect of which it compensates.

Theoretical progress is also possible. Classifying genotypic fitness landscapes according to the presence or absence of sign epistasis may be more instructive than asking whether they have multiple peaks. But still richer characterization of the epistatic structure defined by a fitness landscape is possible. For example, the implications for the selective accessibility of trajectories may differ in the case of a mutation that is beneficial on all but one possible genetic background compared to the case of a mutation that is deleterious on all but one background. A still more sophisticated approach might build on the fact that an evolving population is likely to visit fitter genotypes more often than less-fit genotypes. Thus, the evolutionary importance of sign epistasis may be reduced if mutation to the allele present in the highest fitness genotype is deleterious only on genotypes of very low fitness. Moreover, although we have shown that magnitude epistasis has no influence on the selective accessibility of alternative mutational trajectories, if the magnitude of the selective effect of a mutation varies as a function of genetic background, the rate at (and probability with) which alternative mutational trajectories are likely to be followed by evolving populations may be influenced by such epistasis. Finally, for simplicity we have here focused on mutations that have fitness effects in all genetic backgrounds, and have not explicitly addressed conditionally neutral mutations. Work exploring this possibility can also be illuminating (e.g. van Nimwegen and Crutchfield 2000).

In summary, prior work has recognized two mechanisms by which epistasis for fitness can influence evolution by natural selection. Statistical epistasis among currently segregating alleles can affect additive genetic variance in the presence of linkage disequilibrium and thereby influence the instantaneous rate of allele frequency change. And regardless of current population polymorphism, functional epistasis can constrain the selective availability of mutational trajectories to genotypes of high fitness. Previous attention to the latter point has been focused on the possibility of mutationally isolated fitness peaks: genotypes on which no mutations are beneficial. Here we have emphasized that this is but one example of a broader class of effects, in which sign epistasis reduces the number of selectively accessible trajectories from some genotypes. More specifically, we have shown that unless a system exhibits sign epistasis the fitness landscape will be unconstrained with respect to the selective availability of mutational trajectories to the peak genotype. Sign epistasis is becoming experimentally tractable and we are optimistic that attention to this phenomenon will lead to a clearer and more sophisticated understanding of the forces that shape evolution by natural selection.

Acknowledgments

We thank present and former members of the J. Wakeley, D. Hartl, and L. Chao labs and B. Kerr, R. Harrison, and anonymous reviewers for numerous contributions to the development of this work. The figure on the cover of this issue was prepared with the kind assistance of C. A. Muirhead.

Literature Cited

Björkman, J., P. Samuelsson, D. I. Andersson, and D. Hughes. 1999. Novel ribosomal mutations affecting translational accuracy, antibiotic resistance and virulence of *Salmonella typhimurium*. Mol. Microbiol. 31:53–58.

Brodie, E. D., III. 2000. Why evolutionary genetics does not always add up. Pp. 3–20 *in* J. B. Wolf, E. D. Brodie III, and M. J. Wade, eds. Epistasis and the evolutionary process. Oxford Univ. Press, Oxford, U.K.

Burch, C. L., and L. Chao. 1999. Evolution by small steps and rugged landscapes in the RNA virus ϕ6. Genetics 151:921–927.

———. 2000. Evolvability of an RNA virus is determined by its mutational neighborhood. Nature 406:625–628.

Carter, A. J. R., and G. P. Wagner. 2002. Evolution of functionally conserved enhancers can be accelerated in large populations: a population-genetic model. Proc. R. Soc. Lond. B 269:953–960.

Cheverud, J., and E. J. Routman. 1995. Epistasis and its contribution to genetic variance components. Genetics 139:1455–1461.

Coyne, J. A., N. H. Barton, and M. Turelli. 1997. Perspective: a critique of Sewall Wright's shifting balance theory of evolution. Evolution 51:645–671.

———. 2000. Is Wright's shifting balance process important in evolution? Evolution 54:306–317.

Falconer, D. S. 1994. Introduction to quantitative genetics. Longman Scientific and Technical, Essex, England.

Fenster, C. B., L. F. Galloway, and L. Chao. 1997. Epistasis and its consequences for the evolution of natural populations. Trends Ecol. Evol. 12:282–286.

Fisher, R. A. 1930. The genetical theory of natural selection. Clarendon Press, Oxford, U.K.

Gao, L., and J. Zhang. 2003. Why are some human disease-associated mutations fixed in mice? Trends Genet. 19:678–681.

Gavrilets, S. 2003. Perspective: Models of speciation: What have we learned in 40 years? Evolution 57:2197–2215.

———. 2004. Fitness landscapes and the origin of species. Princeton Univ. Press, Princeton, NJ.

Gillespie, J. H. 1984. Molecular evolution over the mutational landscape. Evolution 38:1116–1129.

Gitchoff, P., and G. P. Wagner. 1996. Recombination induced hypergraphs: a new approach to mutation-recombination isomorphism. Complexity 2:37–43.

Goodnight, C. J., and M. J. Wade. 2000. The ongoing synthesis: a reply to Coyne, Barton and Turelli. Evolution 54:317–324.

Hall, B. G. 2002. Predicting evolution by in vitro evolution requires determining evolutionary pathways. Antimicrob. Agents Chemother. 46:3035–3038.

Hansen, T. F., and G. P. Wagner. 2001. Modeling genetic archi-

tecture: a multilinear theory of gene interaction. Theor. Popul. Biol. 59:61–86.
Holland, J. H. 1975. Adaptation in natural and artificial systems. Univ. of Michigan Press, Ann Arbor, MI.
Horn, J., D. E. Goldberg, and K. Deb. 1994. Long path problems. Lect. Notes Comput. Sci. 866:149–158.
Huynen, M. A., P. F. Stadler, and W. Fontana. 1996. Smoothness within ruggedness: the role of neutrality in adaptation. Proc. Natl. Acad. Sci. USA. 93:397–401.
Iwasa, Y., F. Michor, and M. A. Nowak. 2004. Stochastic tunnels in evolutionary dynamics. Genetics 166:1571–1579.
Jansen, T., and I. Wegener. 2001. Real Royal Road functions: where crossover provability is essential. Pp. 1034–1041 *in* L. E. Spector and E. D. Goodman, eds. Proceedings of the genetic and evolutionary computation conference, GECCO-2001. Morgan Kaufmann, San Francisco, CA.
Kauffman, S. A. 1993. The origins of order. Oxford Univ. Press, Oxford, U.K.
Kern, A. D., and F. A. Kondrashov. 2004. Mechanisms and convergence of compensatory evolution in mammalian mitochondrial tRNAs. Nat. Genet. 36:1207–1212.
Kimura, M. 1985. The role of compensatory neutral mutations in molecular evolution. J. Genet. 64:7–19.
Kimura, M., and J. F. Crow. 1964. The number of alleles that can be maintained in a finite population. Genetics 49:725–738.
Kondrashov, A. S. 1988. Deleterious mutations and the evolution of sex. Nature 336:435–440.
Kondrashov, F. A., and A. S. Kondrashov. 2001. Multidimensional epistasis and the disadvantage of sex. Proc. Natl. Acad. Sci. USA 98:12089–12092.
Kondrashov, A. S., S. Sunyaev, and F. A. Kondrashov. 2002. Dobzhansky-Muller incompatabilities in protein evolution. Proc. Natl. Acad. Sci. USA 99:14878–14883.
Kulathinal, R. J., B. R. Bettencourt, and D. L. Hartl. 2004. Compensated deleterious mutations in insect genomes. Science 306: 1553–1554.
Lee, Y.-H., L. M. D'Souza, and G. E. Fox. 1997. Equally parsimonious pathways through an RNA sequence space are not equally likely. J. Mol. Evol. 45:278–284.
Lipman, D. J., and W. J. Wilbur. 1991. Modelling neutral and selective evolution of protein folding. Proc. R. Soc. Lond. B 245: 7–11.
Maisnier-Patin, S., O. G. Berg, L. Lijas, and D. I. Andersson. 2002. Compensatory adaptation to the deleterious effect of antibiotic resistance in *Salmonella typhimurium*. Mol. Microbiol. 46: 355–366.
Mammano, F., V. Trouplin, V. Zennou, and F. Clavel. 2000. Retracing the evolutionary pathways of human immunodeficiency virus type 1 resistance to protease inhibitors: virus fitness in the absence and in the presence of drugs. J. Virol. 74:8524–8531.
Maynard Smith, J. 1970. Natural selection and the concept of a protein space. Nature 225:563–565.
Peters, A. D., and C. M. Lively. 2000. Epistasis and the maintenance of sex. Pp. 99–112 *in* J. B. Wolf, E. D. Brodie III, and M. J. Wade, eds. Epistasis and the evolutionary process. Oxford Univ. Press, Oxford, U.K.
Phillips, P. C. 1998. The language of gene interaction. Genetics 149:1167–1171.
Poon, A., and L. Chao. 2005. The rate of compensatory mutation in the DNA bacteriophage ϕX174. Genetics. *In press.*
Poon, A., B. H. Davis, and L. Chao. 2005. The coupon collector and the supressor mutation: estimating the number of compensatory mutations by maximum likelihood. Genetics. *In press.*
Provine, W. B. 1986. Sewall Wright and evolutionary biology. Univ. of Chicago Press, Chicago.
Quiñones-Mateu, M., and E. J. Arts. 2001. HIV-1 fitness: implications for drug resistance, disease progression, and global epidemic evolution. Pp. 134–170 *in* C. Kuiken, B. Foley, B. Hahn, B. Marx, F. McCutchan, J. Mellors, S. Wolinsky and B. Korber, eds. HIV sequence compendium. Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, NM.
Schrag, S. J., and V. Perrot. 1996. Reducing antibiotic resistance. Nature 381:120–121.
Schrag, S. J., V. Perrot, and B. R. Levin. 1997. Adaptation to the fitness cost of antibiotic resistance in *E. coli*. Proc. R. Soc. Lond. B 264:1287–1291.
Schuster, P., W. Fontana, P. F. Stadler, and I. L. Hofacker. 1994. From sequence space to shape and back: a case study in RNA secondary structure. Proc. R. Soc. Lond. B 255:279–284.
Siegal, M. L., and D. L. Hartl. 1996. Transgene coplacement and high efficiency site-specific recombination with the Cre/loxP system in *Drosophila*. Genetics 144:715–726.
Stephan, W. 1996. The rate of compensatory evolution. Genetics 144:419–426.
van Nimwegen, E., and J. P. Crutchfield. 2000. Metastable evolutionary dynamics: Crossing fitness barriers or escaping via neutral paths? Bull. Math. Biol. 62:799–848.
Wade, M. J. 1992. Epistasis. Pp. 87–91 *in* E. F. Keller and E. A. Lloyd, eds. Keywords in evolutionary biology. Harvard Univ. Press, Cambridge, MA.
Wade, M. J., and C. J. Goodnight. 1998. Perspective: The theories of Fisher and Wright in the context of metapopulations: when nature does many small experiments. Evolution 52:1537–1553.
Wagner, G. P., and P. F. Stadler. 1998. Complex adaptations and the structure of recombining spaces. Pp. 151–170 *in* C. Nehaniv and M. Ito, eds. Proceedings of the conference on semi-groups and algebraic engineering. World Scientific, Singapore.
Watson, R. A. 2001. Analysis of recombinative algorithms on a non-separable building-block problem. Pp. 69–90 *in* W. N. Martin and W. M. Spears, eds. Foundations of genetic algorithms. Morgan Kaufmann, San Francisco, CA.
———. 2004. A simple two-module problem to exemplify building-block assembly under crossover. Pp. 160–169 *in* X. Yao, E. Burke, J. A. Lozano, J. Smith, J. J. Merelo-Guerós, J. A. Bullinaria, J. Row, P. Tino, A. Kabán and H.-P. Schwefel, eds. Parallel problem solving from nature, PPSN VIII. Springer-Verlag, New York.
———. 2005. Compositional evolution: the impact of sex, symbiosis and modularity on the gradualist framework of evolution. MIT Press, Cambridge, MA.
Weinreich, D. M., and L. Chao. 2005. Rapid evolutionary escape by large populations from local fitness peaks is likely in nature. Evolution *In press*.
Whitlock, M. C., P. C. Phillips, F. B.-G. Moore, and S. J. Tonsor. 1995. Multiple fitness peaks and epistasis. Annu. Rev. Ecol. Syst. 26:601–629.
Wright, S. 1930. Review of the genetical theory of natural selection. J. Hered. 21:349–356.
———. 1932. The roles of mutation, inbreeding, crossbreeding and selection in evolution. Pp. 356–366 *in* D. F. Jones, ed. Proceedings of the sixth international congress of genetics. Brooklyn Botanic Garden, Menasha, WI.
———. 1982. Character change, speciation and the higher taxa. Evolution 36:427–443.

Corresponding Editor: R. Harrison