Perspective: A Critique of Sewall Wright's Shifting Balance Theory of Evolution

Jerry A. Coyne; Nicholas H. Barton; Michael Turelli

Evolution, Volume 51, Issue 3 (Jun., 1997), 643-671.

Stable URL:
http://links.jstor.org/sici?sici=0014-3820%28199706%2951%3A3%3C643%3APACOSW%3E2.0.CO%3B2-L

EVOLUTION
INTERNATIONAL JOURNAL OF ORGANIC EVOLUTION
PUBLISHED BY
THE SOCIETY FOR THE STUDY OF EVOLUTION

Vol. 51 June 1997 No. 3

Evolution, 51(3), 1997, pp. 643–671

PERSPECTIVE: A CRITIQUE OF SEWALL WRIGHT'S SHIFTING BALANCE THEORY OF EVOLUTION

JERRY A. COYNE,[1] NICHOLAS H. BARTON,[2] AND MICHAEL TURELLI[3]
[1]*Department of Ecology and Evolution, The University of Chicago, 1101 East 57th Street, Chicago, Illinois 60637*
E-mail: j-coyne@uchicago.edu
[2]*I.C.A.P.B., Division of Biological Sciences, University of Edinburgh, Edinburgh EH9 3JT, UK*
E-mail: n.barton@ed.ac.uk
[3]*Section of Evolution and Ecology and Center for Population Biology, University of California, Davis, California 95616*
E-mail: mturelli@ucdavis.edu

Abstract.—We evaluate Sewall Wright's three-phase "shifting balance" theory of evolution, examining both the theoretical issues and the relevant data from nature and the laboratory. We conclude that while phases I and II of Wright's theory (the movement of populations from one "adaptive peak" to another via drift and selection) can occur under some conditions, genetic drift is often unnecessary for movement between peaks. Phase III of the shifting balance, in which adaptations spread from particular populations to the entire species, faces two major theoretical obstacles: (1) unlike adaptations favored by simple directional selection, adaptations whose fixation requires some genetic drift are often prevented from spreading by barriers to gene flow; and (2) it is difficult to assemble complex adaptations whose constituent parts arise via peak shifts in different demes. Our review of the data from nature shows that although there is some evidence for individual phases of the shifting balance process, there are few empirical observations explained better by Wright's three-phase mechanism than by simple mass selection. Similarly, artificial selection experiments fail to show that selection in subdivided populations produces greater response than does mass selection in large populations. The complexity of the shifting balance process and the difficulty of establishing that adaptive valleys have been crossed by genetic drift make it impossible to test Wright's claim that adaptations commonly originate by this process. In view of these problems, it seems unreasonable to consider the shifting balance process as an important explanation for the evolution of adaptations.

Key words.—Adaptation, genetic drift, natural selection, peak shift, population structure, shifting balance.

Received May 29, 1996. Accepted December 19, 1996.

> The problem of evolution as I see it is that of a mechanism by which the species may continually find its way from lower to higher peaks. . . . In order that this may occur, there must be some trial and error mechanism on a grand scale by which the species may explore the regions surrounding the small portion of the field which it occupies. To evolve, the species must not be under strict control of natural selection. Is there such a trial and error mechanism? (Wright 1932, pp. 358–359)

During the Modern Synthesis, evolutionary views of adaptation coalesced into two competing schools. The first, often called Fisherian after its most famous proponent (Fisher 1930), holds that the bulk of adaptive evolution results simply from Darwinian mass selection. Some, on the other hand, feel that adaptation cannot be explained by selection alone, and that stochastic processes such as genetic drift often play an important role. The major vehicle for the latter view is Sewall Wright's famous "shifting balance" theory of evolution. Proposed in mathematical form in 1931, described nonmathematically in 1932, and then explained and defended by Wright in a lifelong series of papers, the shifting balance theory has been regarded by many—including Wright himself—as his most important contribution to biology. Its influence has permeated evolutionary theory, population genetics, and animal breeding.

Although the mathematics of the shifting balance theory (henceforth SBT) is complicated, its essence is simple. Wright proposed that adaptation involved a shifting balance between evolutionary forces, resulting in a three-phase process:

Phase I: Genetic drift causes local populations (demes) to temporarily lose fitness, shifting across "adaptive valleys" toward new "adaptive peaks."

Phase II: Selection within demes places them atop these new peaks.

Phase III: Different adaptive peaks compete with each other, causing fitter peaks to spread through the entire species. (Wright believed that populations occupying higher adaptive peaks would send out more migrants, ultimately driving other populations to the highest peak.)

There is thus a clear distinction between the Fisherian and Wrightian views of evolution: the former requires only that populations be larger than the reciprocal of the selective coefficient acting on a genotype, and the latter requires subdivided populations, particular forms of epistasis, genetic drift that counteracts selection, and differential migration between populations based on their genetic constitution.

Although the SBT has received occasional criticism (e.g., Haldane 1959; Kimura 1983; Provine 1986; Hill 1989; Slatkin 1989), a comprehensive evaluation has never been published. Here we critique the SBT, discussing its role in evolutionary biology, its mathematical basis, Wright's reasons for proposing it, empirical evidence for its assumptions and conclusions, and attempts to test it in nature and the laboratory.

Influence of the Theory.—Approbation for the SBT has been nearly universal. Soon after 1931, it made its way into the founding manifestos of the Modern Synthesis, reaching many biologists who were baffled by Wright's complicated mathematics. Dobzhansky, for example, incorporated a 12-page discussion of the SBT in his *Genetics and the Origin of Species* (1937, pp. 180–191), reproducing three figures from Wright's 1932 paper, including the famous diagrams of adaptive landscapes. Praising the SBT, Dobzhansky declared that "Wright argues very convincingly that . . . a differentiation into numerous semi-isolated colonies is the most favorable one for a progressive evolution" (p. 190). Wright's theory motivated much of Dobzhansky's later work on population structure and migration in natural populations of *Drosophila* (Lewontin et al. 1981; Provine 1986).

In his influential *Tempo and Mode in Evolution* (1944), G. G. Simpson also accepted Wright's claim that Fisherian mass selection could not explain major adaptive shifts in the fossil record, and suggested that the SBT might explain "punctuated" patterns of evolution (e.g., pp. 67–69, 89–96, 121–122). (Simpson was, however, somewhat less laudatory in *The Major Features of Evolution* [1953, p. 123], noting that many species did not possess population structures or migration rates conducive to operation of the SBT.) Wright's views on stasis and the hierarchical nature of selection have also inspired modern proponents of "punctuated equilibrium," including S. J. Gould (1980) and N. Eldredge (1995, ch. 3). Finally, in *Systematics and the Origin of Species*, Mayr (1942, p. 236) cited Wright's 1931 and 1932 papers in support of evolutionary inertia in large populations: "The consequence of this consideration is that evolution should proceed more rapidly in small populations than in large ones, and this is exactly what we find." The interaction between selection and genetic drift and the supposed insensitivity of large populations to selection are important parts of Mayr's famous "founder-effect" theory of speciation (1963), and of other recent theories of speciation in colonizing populations (see Provine 1989).

In his biography of Wright, Provine (1986, chs. 9–12) documents the influence of the SBT–and its attendant metaphor of adaptive landscapes–on evolution, genetics, and animal breeding. Wright's 1931 paper was cited 90 times between 1985 and 1995 in the journals *Evolution* and *Genetics* alone; of these citations, nearly half (41) were in experimental or theoretical analyses of genetic differentiation among subpopulations. An additional 15 papers dealt with other aspects of the SBT, such as epistasis and the efficacy of phase III. As a major paradigm of evolution, the SBT is discussed in many textbooks, usually uncritically (a notable exception is Ridley 1993, pp. 205–208).

Why has Wright's theory proven so attractive? In part, it is because the shifting balance (henceforth SB) process provides a fusion of three disparate elements: complex mathematics, an experimentally informed concern for the origin of complex adaptations, and especially an attractive graphic metaphor—the "adaptive landscape." The alternative mathematical versions of simple Darwinism elaborated by Fisher (1930) and Haldane (1932) emphasize the effects of individual alleles in isolation, the brand of "beanbag genetics" so effectively demonized by Mayr (1959). Moreover, because the SBT encompasses a diversity of evolutionary elements and processes (epistasis, pleiotropy, population structure, selection, and genetic drift), some may consider it more comprehensive than simple Fisherian selection in populations of reasonable size.

Scope of the Theory.—Wright clearly thought that the SBT explained not just an occasional adaptation, but most of them, including complex ones. These claims for generality are evident from his early work:

> It has been pointed out, however, that the most favorable conditions for a continuing evolutionary process are those in which there is, to a first order, balanced action of all of the statistical evolutionary factors. It is consequently to be expected that in most actual cases indications can be found of simultaneous action of all of them (Wright 1940, p. 181).

In volumes 3 and 4 of his *magnum opus*, Wright (1977, 1978) interpreted a wide variety of observations from nature and the laboratory in light of the SBT. Indeed, in these four books it is hard to find a single experiment or observation that is not construed as supporting this theory. In the peroration, Wright (1978, p. 524) declared:

> It is maintained here that the continued operation of this shifting balance process, involving joint action of all evolutionary factors—mutations and immigration pressures, mass selection, random drift of all sorts, and interdeme selection—has been the principal basis for evolution under exclusively biparental reproduction.

Although Wright occasionally admitted that Fisherian mass selection might play some role in evolution (e.g., 1977, p. 451), he always added that the SBT permitted more rapid adaptation.

We begin our analysis with an examination of the theory itself and then discuss the data offered in its support by Wright and others. We will conclude that (1) many of Wright's motivations for the SBT were based on the problems he perceived with the alternative process of mass selection, but these problems are largely illusory; (2) although, as Wright

postulated, alternative adaptive peaks separated by adaptive valleys clearly exist, there is little evidence for the assumption that movement between peaks involves a temporary loss of fitness; (3) although phases I and II of the theory may be at least theoretically plausible, there is little theoretical support for phase III of the shifting balance, in which adaptations spread from particular populations to the entire species; (4) the few possible examples of the SB process do not increase adaptation in the way envisioned by Wright; (5) there are almost no empirical observations that are better explained by Wright's mechanism than by mass selection; and (6) because of the complexity of the SBT, it is impossible to test Wright's claim that it is a common evolutionary process. In view of these problems, we think that it is unreasonable to consider the SBT an important explanation of adaptation in nature.

Our criticisms of the SBT are not meant to belittle Wright's numerous other contributions to population genetics, including his fundamental work on inbreeding, the theory of genetic drift and its interaction with other evolutionary forces, and polygenic evolution. Here we address only his ambitious attempt to meld these factors into a grand theory of adaptation.

Our discussion of the theoretical issues becomes quite technical at times. To readers who wish to avoid these technicalities, or who are interested only in empirical work on the SBT, we suggest reading the section "Wright's Motivations," the first four paragraphs and last six paragraphs of the subsequent "Critique of the Theory" section, and then the sections entitled "The Evidence" and "Conclusions."

Wright's Motivations.—The SBT was not proposed to account for some new observations about nature that defied explanation by simple Darwinism. Rather, Wright wanted to combine his work on epistasis, drift, and animal breeding into a new explanation for facts that were already known (Wright 1978, 1982; Provine 1986, pp. 233–237). A perusal of his early papers, however, also reveals motivations deriving more from deduction than from observation. Underlying all of Wright's work on the SBT, and made explicit in his early publications, was his belief that mass selection was powerless to solve the main problem of evolution, framed as the "mechanism by which the species may continually find its way from lower to higher peaks" (Wright 1932, pp. 358–359):

> Since gene frequencies tend to approach equilibrium under long-persisting conditions, continuing evolution by individual selection depends either on continually changing conditions or on the occurrence of novel favorable mutations. Conditions are, indeed, continually changing in terms of geologic time, but the ensuing evolution is somewhat like a treadmill, old adaptations continually being undone in favor of new ones. A more serious limitation is the restriction of selection to the net effects of genes (except where two or more are strongly linked), because of the rapid breaking up of gene combinations in meiosis.
>
> Novel favorable mutations no doubt occur from time to time, but so rarely that evolution on this basis must be enormously slower than from changing conditions. Even under changing conditions, however, the rate of gene substitution is necessarily slow because the rate of selective change of frequency at any locus is limited by the amount of reproductive excess . . . the so-called cost of evolution (Wright 1977, pp. 561–562).

Wright claimed that this cost was reduced by the process of interdeme selection (with Wright, we use the term "interdeme selection" to denote a form of selection that depends on population subdivision into demes):

> [Haldane] held that even geologic time might not be sufficient for many observed evolutionary changes to have been brought about by an accumulation of minor mutations. This, however, applies only to panmictic populations. "Cost" is not an important consideration if subdivided among many local populations, as occurs under the shifting balance process (Wright 1982, p. 8).

Finally, Wright believed that low heritabilities would make mass selection ineffective, even over long periods of time. Referring to his work on guinea pigs, he noted that:

> About 70 percent of the individual variation in resistance to tuberculosis and over 90 percent of that in the rate of gain, and size of litter is determined by external conditions. Progress by ordinary selection of individuals would thus be very slow or nil. A single unfortunate selection of a sire, good as an individual, but inferior in heredity, is likely at any time to undo all past progress. On the other hand, by starting a large number of inbred lines, important hereditary differences in these respects are brought clearly to light and fixed (Wright 1922, p. 49).

Wright therefore felt that the SBT produced adaptation much more quickly than did mass selection because the latter process was limited by (1) the rate of favorable mutations, which he thought was very low; (2) the "cost of selection," which imposed an upper bound on the rate of adaptive evolution; (3) the low heritability of characters; and (4) its inability to fix combinations of genes that were individually deleterious but adaptive in concert.

When examined closely, these arguments against mass selection lose their force. First, we know of no evidence that mass selection is too slow to explain either adaptation or biological diversity. Such a claim would have been stronger had Wright presented any calculations or data suggesting that Fisherian selection could not explain these phenomena. Wright is surely correct in noting that continuing evolution depends on changing conditions or new favorable mutations, but we have no empirical evidence that natural populations adapt faster under conditions favoring the SBT than under those favoring Fisherian selection. For instance, Simpson's claim in *Tempo and Mode in Evolution* (1944, p. 95) that "large populations tend to a genetic equilibrium relatively unfavorable for rapid and sustained progressive evolution" was clearly influenced by Wright's assertions; yet Simpson's own analysis of the fossil record led him to suggest that phyletic evolution (i.e., large-scale anagenic changes such as the increase of hypsodonty in horse evolution) was "most clearly seen in groups with continuously large to moderate breeding populations" (Simpson 1944, p. 205 and table 19). Modern quantitative-genetic analyses (e.g., Hill 1982; Lynch

and Lande 1993) suggest that mutation creates additive variance at a rate sufficient to fuel rapid adaptation; this is reflected in the observed cumulative response to selection starting from highly inbred base populations (e.g., Mackay et al. 1994). Theory also predicts that the rate of advance should increase with population size, as observed, for instance, by Weber and Diggins (1990).

Second, it is not clear that the cost of substitutions constrains the rate of adaptation (Felsenstein 1971), or that the cost would be reduced under the SBT. Haldane's cost argument would impose a serious limit on higher organisms only if most substitutions were selected; even then, epistasis can greatly reduce the cost (Gillespie 1991, p. 272). With additive selection (no dominance or epistasis for fitness), subdivision increases the substitution load because it delays the spread of the favorable allele through the population (Barton, unpubl.); no analytical results are available concerning the joint effects of epistasis and population subdivision on the substitutional load. Under the SBT, gene combinations are established within demes by random drift opposing selection, and then spread across populations by some kind of interdemic selection; both processes require that some individuals leave more offspring than others. It is therefore implausible that the net "cost of evolution" should be lower with the SB than with the Fisherian process. We note, however, that evolution need not proceed in the least costly manner.

Third, Wright's early view that broad-sense heritabilities on the order of 10–30% would limit adaptation by mass selection is a misconception that may be related to the view (prevalent at the time) that selection on polygenic traits was complicated and possibly insufficient to account for adaptive change. (Although Wright never repeated this argument, he never repudiated it either, and as noted below, it probably influenced his view that the SBT explained selection response in organisms such as shorthorn cattle.) The traditional equation of response to polygenic selection, $R = h^2S$, did not appear until 1937 (Lush, p. 84); and, in *The Causes of Evolution*, Haldane (1932, p. 196) asserted that the response of a polygenic character to selection could not be predicted because "the gene ratios for a number of different genes will be varying at once, and hence the mean will vary in an unpredictable manner." We now have extensive evidence of heritable variation for most traits in most populations, and some confidence that the "breeder's equation" usefully predicts at least the short-term response to selection (Hill and Caballero 1992; Turelli and Barton 1994). Moreover, numerous experiments show that mass artificial selection can produce rates of phenotypic change far exceeding any observed in natural populations (Gingerich 1983; Falconer 1989, ch. 12).

Finally, with the possible exception of molecular "covarions" in messenger RNA (see below), there is little evidence that adaptations frequently involve changes in the frequencies of alleles that are individually deleterious but adaptive together, or that phenotypic evolution involves the appearance of maladaptive intermediates (cf. Darwin 1859, ch. 6). While such changes may theoretically occur, we cannot point to a single empirical example that demands such an explanation.

When discussing the superior speed and efficiency of the SBT, Wright's arguments often sound vaguely teleological. Even if the SB process were really more efficient than mass selection, there is no reason to believe that evolution works in the most efficient way. We therefore bypass the question of which evolutionary process would in principle be optimal, and concentrate on the more concrete problem of whether a subdivided population adapts more rapidly than a panmictic one.

Critique of the Theory

Wright's proposed SB between different evolutionary forces is a complex process that raises two sets of theoretical questions. The first concerns phases I and II. Do populations of a species often occupy alternative adaptive peaks? How often do evolutionary transitions correspond to movements between peaks on an adaptive landscape of mean fitnesses? Under what circumstances will populations switch peaks by crossing adaptive valleys against the force of selection? Is drift necessary for movement between peaks? Are fitness landscapes sufficiently constant that populations are likely to be near equilibria? The second set concerns the third phase. Are the conditions that allow phases I and II to operate consistent with the action of phase III? What happens when populations near different peaks exchange migrants; in particular, do genotypes corresponding to fitter peaks tend to spread? These two sets of questions are largely independent. For example, selection among divergent gene combinations fixed in different demes may allow the species to either adapt faster or reach a higher fitness peak than would be possible with panmixis, regardless of whether divergence was due to drift (as postulated by the SBT) or to selection.

Wright developed fundamental mathematical methods to represent different portions of the process: for example, the island model of population subdivision (Wright 1931), the calculation of the probability of shifts between adaptive peaks (Wright 1941), and the distribution of allele frequencies under selection, mutation, migration, and drift (Wright 1937a,b). However, he never produced a single mathematical analysis that included more than one phase of the SBT. In the Appendix, we adapt Wright's methods, producing a model of the SB process that includes migration among demes, but not the fitness-based differential migration postulated by Wright. Our model builds on Wright's (1935a,b) first explicit multilocus example of a complex adaptive landscape and uses his island model, in which a large number of equal-sized demes exchange migrants. This analysis quantifies the interaction of selection, drift, and migration and shows that, as Wright suspected, intermediate levels of migration may maximize mean fitness. Even in our simple model, however, the region of parameter space producing this effect is limited. Moreover, the main weaknesses of phase III of the SBT are not apparent from our island model of population structure, which ignores variation in population density and the isolating effects of distance. In the following discussion, we separate the various components of the SBT and then assess the theoretical plausibility of the process as a whole.

What Is the Adaptive Landscape?

Because the SBT has been closely associated with Wright's (1931, 1932) idea of an adaptive landscape, we must consider

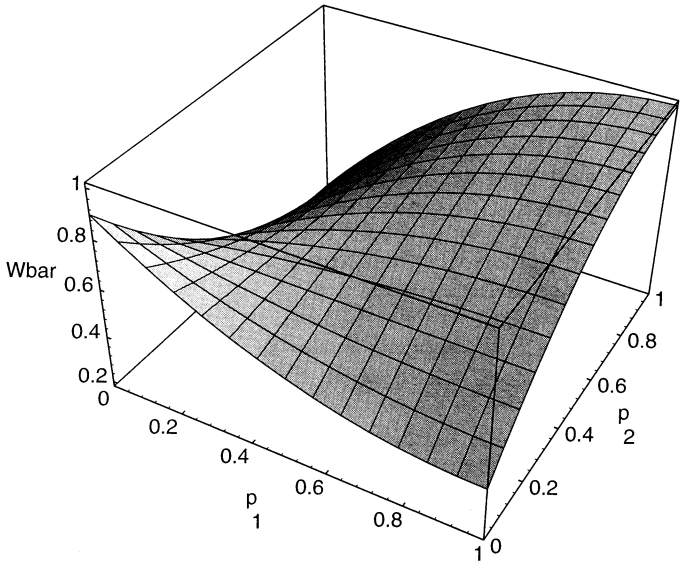

FIG. 1. A simple "adaptive landscape" for two loci. Mean fitness is computed assuming no linkage disequilibrium and complete dominance at each of two diallelic loci. The dominant allele frequencies are denoted p_1 and p_2. The "double dominant" (dd) and double recessive (rr) genotypes are assumed to be most fit, with $w_{dd} = 1$, $w_{rr} = 0.8$, $w_{dr} = 0.4$, and $w_{rd} = 0.2$.

the utility of this metaphor. Confusingly, Wright used two quite different definitions, apparently without realizing the fundamental distinction between them (Provine 1986; Wright 1988). The first describes the relation between *the genotypes of individuals* and their fitness, while the second describes the relation between *the allele frequencies in a population* and its mean fitness. The latter is similar to Lande's (1976, 1979) mean fitness surfaces, in which the fitness of a population is a function of the means of a set of quantitative traits with fixed variances and covariances. We use the second definition throughout because it lends itself to a quantitative description of the evolutionary process.

Figure 1 illustrates a simple two-locus example in which two homozygous genotypes are more fit than all other genotypes (see the legend for parameter values). For simplicity, we assume dominance at each locus. Populations fixed for either the *AABB* or *aabb* genotypes form alternative adaptive peaks that differ in net fitness. These two genotypes could, for example, produce coat colors that are both cryptic but have different pleiotropic effects on fertility, while genotypes having dominant alleles at only one of the two loci would have conspicuous coloration. Under a regime of visual predation, an adaptive landscape in which population fitness was plotted against the frequency of alleles at the two loci would form a "saddle," as shown in Figure 1. If the adaptive landscape remains constant, polymorphic populations anywhere on it will be forced by selection to the nearest peak, which may not be the highest one. Once on the lower peak (fixation for *aabb*), a population cannot then reach the higher peak by selection alone. Wright's SBT was meant to explain how genetic drift could drive such a population through the adaptive valley, allowing selection to draw it up the higher peak. The genotype corresponding to this peak could then spread through the entire species by interdeme selection. In this way, a combination of evolutionary forces could allow populations to move to ever higher peaks (and fitnesses) on the adaptive landscape.

Wright usually envisioned such peaks as resulting from selection for a single optimum phenotype that could be produced by various genotypes differing in their pleiotropic effects on fitness. In such a case, the SB process might not produce a *phenotype* different from that achievable by mass selection, but would still produce a population with higher mean fitness. It is also possible, however, to imagine the SBT operating on phenotypes that differ in fitness, as in directional artificial selection experiments. (Using the two-locus model, for example, *AABB* platypuses might produce more eggs than would *aabb* platypuses.) In this situation, one might expect the SB process to produce phenotypes different from those attainable by mass selection. This latter view is implicit in the discussions of Hill (1989, 1996) and Weber (1996) about testing the SBT using artificial selection experiments. Similarly, the SBT is often invoked to explain the transition between different highly adapted phenotypes separated by less-fit intermediates (e.g., Simpson 1944). Wright himself occasionally conceived of the SBT in this way, as in his assertions that the SBT was superior to mass selection in producing biological diversity.

The landscape of mean population fitness is a useful metaphor because selection causes populations to evolve uphill on this landscape at a rate proportional to its gradient. (This requires the approximations that fitnesses are independent of genotype frequencies and that selection is weak enough relative to recombination that linkage disequilibrium is negligible.) If the landscape stays constant, a local maximum of mean fitness (an adaptive peak) will eventually be reached (at least approximately—mutation and recombination usually decrease fitness). Wright (1937a,b) exploited this property to approximate the simultaneous effects of drift, selection, and migration on the joint distribution of allele frequencies. Using the island model, he considered n diallelic loci, with alleles P_i and Q_i at locus i having frequencies p_i and $q_i = 1 - p_i$ in a specific deme; demes are assumed to be so numerous that even though allele frequencies in each deme fluctuate randomly, the average frequency in the pool of migrants (denoted $\bar{p}_i$) changes deterministically. Each deme is assumed to have N diploid individuals, and each is assumed to receive a fraction m of its zygotes each generation from a common migrant pool to which each deme contributes equally. Under these assumptions, Wright (1937a,b) showed that the joint distribution of allele frequencies across loci, $\mathbf{p} = (p_1, p_2, \ldots, p_n)$ in each deme can be approximated by

$$\Psi(\mathbf{p}) = C\bar{W}^{2N} \prod_i p_i^{4Nm\bar{p}_i + 4N\mu - 1} q_i^{4Nm\bar{q}_i + 4N\nu - 1}, \qquad (1)$$

where m denotes the fraction of new immigrants (from the migrant pool with allele frequencies $\bar{p}_i$ at locus i) in each deme each generation; $\bar{W}$ is the mean fitness, which depends on the local allele frequencies; C is a normalizing constant; p_i and $q_i = 1 - p_i$ are the allele frequencies at the ith locus, which vary randomly across demes; $\bar{p}_i$ and $\bar{q}_i$ are the average allele frequencies across all the demes; μ is the mutation rate from allele Q_i to P_i, and ν the rate from P_i to Q_i. Equilibria for the average allele frequencies across demes are deter-

mined by solving the equations $\bar{p}_i = \int \bar{p}_i \Psi(\mathbf{p}) d\mathbf{p}$, with $\Psi(\mathbf{p})$ given by equation (1) (see Barton and Rouhani 1993). We apply equation (1) in our analysis of polygenic selection given in the Appendix.

Wright's formula (1) approximates the frequency distribution of an ensemble of demes across the adaptive landscape. The leading term $\bar{W}^{2N}$ describes the force of selection pushing allele frequencies toward peaks. Of course, evolution does not exactly satisfy these assumptions. In general, mean fitness does not necessarily increase under recombination, mutation, or migration; and the stable equilibrium genotype frequencies will often not correspond precisely to fitness maxima (Wright 1967; Nagylaki 1976, ch. 8; Ewens 1979, ch. 6). If fitnesses depend on genotype frequencies, the change due to selection will still be proportional to the selection gradient (defined as changes of log mean fitness with allele—or more generally, haploid genotype—frequencies) in each generation, but the adaptive landscape will change as the genotype frequencies change. For most forms of multilocus selection, it is impossible to define any function (such as mean fitness) that increases each generation (Wright 1967; Akin 1979). Although these problems complicate mathematical analyses and limit the heuristic value of the adaptive-landscape metaphor, they do not damage the SBT. Even if no quantity is being maximized, allopatric populations will typically evolve toward alternative stable states, and the rate of adaptation might depend on the rate of drift-induced transitions between these states. We will use "adaptive peak" as shorthand for a stable state (for examples in which the stable states are not fixed frequencies of genotypes, see Boerlijst and Hogeweg 1995).

Selection is not the only factor besides drift that influences genotype frequencies. Equation (1) is the product of the factor $\bar{W}^{2N}$ describing selection and the factors $p_i^{4Nm\bar{p}_i+4N\mu-1}$ and $q_i^{4Nm\bar{q}_i+4N\nu-1}$ describing mutation and migration. In more complex situations, the pattern of mutation may also determine which genotypes are available, and hence which adaptive peak is approached. The adaptive landscape metaphor treats the available alleles as given, but chance may often determine the genetic constitution of the landscape on which evolutionary forces act. Consider, for example, haploid selection at one locus with three alleles (*P, Q, R*), whose fitnesses are $(1{:}1 + S{:}1 + T)$, with $0 < S < T$. Here, the complete adaptive landscape is a surface that increases monotonically to a single peak, corresponding to the fixation of *R*. However, if mutation is impossible between *Q* and *R* (which precludes application of the multiallelic version of eq. 1), and mutations are rare, a population that starts with allele *P* may establish either *Q* or *R* (Johnson et al. 1995). Hence, under any model of evolution, chance and historical constraints are likely to help determine the route traversed on the "mutational landscape," and the highest fitness peaks may remain inaccessible. This is most obvious for protein or nucleic acid sequences, where each sequence can mutate to only a tiny fraction of possible sequences (Maynard Smith 1970; Gillespie 1984; Eigen 1986).

Although the SBT focuses on the mean fitness landscape, discussions of evolutionary transitions often concern the *individual* fitness surface, defined by assigning fitnesses to phenotypes described by several traits. This view is embodied in Simpson's (1944, Ch. VI) metaphor of transitions between "adaptive zones," the different ways of life that separate major groups of organisms. Part of the appeal of Wright's theory may result from a confusion between mean-fitness and individual-fitness landscapes, resulting in a failure to appreciate the changes in these surfaces that are caused by intra- and interspecific competition. In particular, populations may often move between peaks on the *individual* fitness surface, which corresponds to a transition between alternative highly adapted morphologies separated by less fit intermediates, without a decrease in mean *population* fitness.

Such transitions between peaks in the individual fitness surface can occur in two ways. The first relies on the fact that the mean fitness surface of a population, which involves averaging over the fitnesses of all genotypes present, is always smoother than the individual fitness surface. For instance, consider stabilizing selection on phenotypes described by $w(z) = \exp[-z^2/(2\omega^2)]$. If z is normally distributed with mean $\bar{z}$ and variance V_z, the mean fitness surface is smoother because it is proportional to $\exp[-\bar{z}^2/(2V)]$, with $V = \omega^2 + V_z > \omega^2$. In biological terms, "optimal" phenotypes will occur even when the population mean is away from the optimum, and peaks in the surface of individual fitness need not produce peaks in the surface of mean population fitness (Felsenstein 1979; Kirkpatrick 1982). Consider, for instance, a two-peaked fitness surface for bill size in birds, with a lower peak at a small bill size and a higher peak at a large size. This might be produced by a bimodal distribution of seed sizes, with larger seeds more abundant. An individual of intermediate bill morphology may be less fit because of its inability to exploit either seed size efficiently. Nevertheless, a *population* whose mean is intermediate need not reside in a fitness valley, because it may contain a mixture of individuals that can efficiently use both resources. This could lead to evolution of a larger bill size by selection alone.

The second way in which selection can cause such transitions involves competition. As in the bill size example, adaptive peaks may correspond to morphologies adapted to alternative limiting resources (e.g., Simpson 1944, ch. VI; Schlefer et al. 1986; Grant 1986; Smith 1993; Schluter 1996). Although intermediates may be less fit when both better-adapted specialists are present, the same intermediates may be at a selective advantage when one of the specialists is absent (Wilson and Turelli 1986), allowing transitions between alternative peaks in the individual fitness surface by selection alone. This process can also produce stable polymorphisms as well as multiple equilibria (Maynard Smith 1966; Wilson and Turelli 1986). This example shows that Wright's metaphor of an extrinsically determined, static adaptive landscape cannot accommodate fluctuations in mean fitness produced by density and frequency dependence.

Models of selection on many loci typically have very many stable equilibria separated by unstable states with lower fitness (Hastings 1989). Simple examples include stabilizing selection on an additive polygenic trait (Wright 1935a; appendix), selection against heterozygous chromosome arrangements, frequency-dependent selection favoring the commoner type (as with Müllerian mimicry; Mallet 1986), and disruptive selection on one or more quantitative traits (e.g., Brodie 1991). While multiple peaks on the adaptive

landscape are easily visualized when just a few genes or traits are involved, the issues become much more complex when considering many loci or traits (Provine 1986, p. 274). Under such conditions, are there peaks around which fitness falls off in all directions? If so, how does the fitness of the peak attainable by mass selection compare with the fitness that might be attained by the SB? As demonstrated in the appendix, the SB process does not ensure that the most-fit peak is always reached. The interaction of selection, population subdivision, drift, and migration can lead to either higher or lower mean fitness than mass selection alone, depending on the parameter values and initial conditions considered.

Fisher questioned the existence of true peaks, arguing that as the number of traits considered increases, the chance that a given equilibrium would be stable to perturbations in all directions decreases (cf. Provine 1986, ch. 8). Hence, a population would be likely to find some direction in which natural selection could increase mean fitness, so that no valleys need be crossed. This view becomes more plausible if one considers that each locus may produce a wide range of allelic effects rather than the fixed effects often assigned by population geneticists.

A potential flaw in Fisher's argument, however, is that although the conditions for stability of equilibria become increasingly restrictive with increasing numbers of traits or loci, the *number* of equilibria may increase so rapidly that the number of stable equilibria actually increases. For instance, Kauffman and Levin (1987) showed that if fitnesses are assigned entirely randomly to genotypes, then the number of locally optimal genotypes rises exponentially with the number of loci, so that a random genotype is typically only a few steps away from the nearest local optimum. A Fisherian evolutionist might respond by asserting that adaptive landscapes are constantly changing and also that adjacent genotypes have similar fitnesses—not the random fitnesses considered by Kauffman and Levin (1987). Both considerations may convert local peaks into ridges, allowing adaptive advance by mass selection.

Kauffman and Levin (1987) relaxed their unrealistic fitness assumption by constructing models that assign similar fitnesses to similar genotypes. In their "*NK* model," the fitness contribution of each of N loci depends in a random way on K other loci. Thus, the parameter K describes the degree of epistasis. With additive fitness effects across loci ($K = 0$) and directional selection at each locus, selection inevitably leads to the global optimum, irrespective of population structure. As K increases, however, the number of local optima increases, and the mean fitness of the nearest optimum attainable by selection decreases toward that of an entirely random genotype. Fontana et al. (1988, 1991, 1993a,b) have developed models with a stronger link to biology by simulating selection for stable RNA sequences. Their results correspond to the *NK* model, with $K \approx 7$–8.

Conclusions.—The existence of reproductively isolated species is strong evidence for alternative adaptive peaks separated by deep valleys. There are many reasons, however, to doubt whether transitions between these peaks often require that populations cross adaptive valleys with the help of genetic drift.

Phases I and II: Transitions between Adaptive Peaks

Peak Shifts Involving Genetic Drift.—There have been numerous calculations of the chance that random drift will cause a shift to a new adaptive peak. Though many of these calculations have been motivated by an interest in chromosomal evolution and founder-effect speciation, they provide a fairly complete analysis of phase I of the SBT. The simplest case is that of heterozygotes at a single locus having relative fitness $1 - s$ while both homozygotes have fitness 1; the problem is to find the chance that a single mutation will be fixed in a single deme of N diploid individuals. Wright (1941) argued that this is proportional to the probability that the population will reach the saddle point that separates the two adaptive peaks; once past this point, selection will pull it to the new peak (phase II). Wright further argued that this probability is given by equation (1), and hence depends primarily on the mean fitness of the population at the saddle point, relative to the original peak. With symmetrical selection against heterozygotes, the critical point is at $p = 0.5$, where the mean fitness is $\bar{W} = 1 - s/2$, and the probability of reaching this point is proportional to $\bar{W}^{2N} = (1 - s/2)^{2N} \approx \exp(-Ns)$. This implies that fixation is likely only with very weak selection or very small deme size. Subsequent work has shown the accuracy of Wright's (1941) analysis; the key conclusion is that the rate of peak shifts decreases exponentially with Ns, and numerical examples show that it is typically extremely small for $Ns > 10$ (Lande 1979; Hedrick and Levin 1984).

Equation (1) implies that at equilibrium the chance that a deme is near some point, $\mathbf{p}$, is proportional to $\bar{W}(\mathbf{p})^{2N}$; hence, as expected, demes are most likely to be near adaptive peaks. Wright's argument and its generalizations demonstrate that the rate of transitions between peaks 1 and 2 is proportional to $(\bar{W}_0/\bar{W}_1)^{2N}$, where $\bar{W}_1$ is the mean fitness at peak 1 and $\bar{W}_0$ is the mean fitness at the saddle point separating the peaks. The rate of transition in the opposite direction is proportional to $\bar{W}_0/\bar{W}_2)^{2N}$, and so the relative times spent on peak 1 versus peak 2 is proportional to $(\bar{W}_1/\bar{W}_2)^{2N}$ (Barton and Charlesworth 1984). If demes are large, then a slight difference between $\bar{W}_1$ and $\bar{W}_2$ leads to a very high probability of being near the fitter peak. The rate of approach to this equilibrium, however, becomes extremely slow as N increases because peak shifts become increasingly unlikely.

Wright's (1941) argument that the chance of a shift to a new adaptive peak can be approximated from the equilibrium probability density (eq. 1) applies quite generally: for example, to models of disruptive selection on quantitative traits, and to models of discrete loci with arbitrary epistasis (Ludwig 1975; Gillespie 1983; Barton and Charlesworth 1984; Lande 1985a; Barton and Rouhani 1987; Mangel 1994). The key assumptions for extending Wright's approximation to such models are: (1) the dynamics must be described by the adaptive landscape, or some analogous potential function; (2) evolution must be slow enough that the diffusion approximation applies; and (3) population size must be constant. Wagner et al. (1994), on the other hand, claim that epistasis per se may facilitate peak shifts. In their model of stabilizing selection on a quantitative trait, peak shifts occur more rapidly if loci interact epistatically rather than additively in determining the

trait value. In their specific epistatic model, however, epistasis reduces the depth of the adaptive valley, and this is the real reason that peak shifts become more frequent. The reduced valley depth in Wagner et al.'s (1994) model is due to the fitness asymmetries induced by epistasis: this increases the variation in valley depth along the alternative routes connecting the adaptive peaks, and thereby makes shifts more likely (see the two-locus analysis of Michalakis and Slatkin 1996). The general theory describing the probability of peak shifts in terms of mean fitnesses applies irrespective of epistasis and predicts that the primary factor determining the rate of shifts is valley depth.

These analyses emphasize the cumulative effects of relatively small population size in producing peak shifts. Alternatively, drift might be most effective during the founding of new colonies. The chance that a severe population bottleneck will cause a peak shift has been studied as a way of understanding founder-effect speciation (e.g., Templeton 1980; Charlesworth and Smith 1982; Barton and Charlesworth 1984; Rouhani and Barton 1987a; Charlesworth and Rouhani 1988; Gavrilets and Hastings 1996), but bottlenecks may also contribute to the SB. If the founding population is small enough ($Ns \ll 1$), selection will be negligible during the bottleneck, and so the chance of a peak shift is just the chance that the population will drift into the domain of attraction of the new peak during the bottleneck (Rouhani and Barton 1987a). For example, a bottleneck that reduced heterozygosity by a factor $1 - F$ will cause the mean of an additive quantitative trait to drift randomly in each generation, in a distribution with variance $2FV_g$ (Barton and Charlesworth 1984). Thus, even in the most severe bottleneck, there is only a 5% chance that the mean will drift by more than $2\sqrt{2}$ genetic standard deviations (i.e., by $2.8\sqrt{V_g}$ for $F = 1$). For a peak shift to be probable, there must be enough standing variation, the bottleneck must be severe, and the domain of attraction of the new peak must be sufficiently close. If the shift involves strong selection, this implies that there must be a substantial standing genetic load due to variation around the original optimum that can contribute to peak shifts (Barton 1989a,b).

In contrast, Gavrilets and Hastings (1996) argue that founder-effect speciation may be much more plausible if there are ridges in the adaptive landscape (as proposed by Dobzhansky 1937, p. 256), which make it unnecessary to cross adaptive valleys. If many such shifts accumulate in allopatry, then the F_1 and F_2 hybrids may be so unfit as to give strong reproductive isolation. However, when demes that differ by only one or a few such peak shifts meet (as is required in the SBT), the relatively fit intermediate state will be reconstructed by recombination, and the effect is to give a shallow valley—just as with, say, weak underdominance. The point here is that although individual changes may face only weak counterselection, and in the short term lead to weak reproductive isolation, they may accumulate to give strong isolation, and the erroneous suggestion that a deep valley has been crossed. The importance of ridges in the landscape can be seen in Phillips' (1993) analysis of the evolution of compensatory mutations. If two substitutions are individually deleterious, but give higher fitness when together, then the rate of shifts is proportional to the *square* of the mutation rate, and hence is extremely slow. Phillips (1993) argues that peak shifts occur at a reasonable rate only if selection against single changes is sufficiently weak.

The chance of a peak shift for a quantitative trait may also be enhanced by drift-induced fluctuations in genetic variance that can flatten the adaptive landscape and temporarily eliminate an adaptive valley. Although the genetic variance may *on average* decrease during such a bottleneck, it may often increase substantially by chance (Whitlock 1995). Moreover, the chance increase of rare recessives (Willis and Orr 1993), or epistatically interacting gene combinations (Goodnight 1988), can increase the average genetic variance, an effect seen in some experiments (Bryant et al. 1986, 1988). One can think of these examples as involving a transition through an adaptive valley that is made shallow (or nonexistent) by a transient increase in genetic variance that raises the frequency of fit individuals.

The above theory applies only to isolated populations. Can the SBT operate in a network of interconnected demes, as Wright envisaged? Though Wright did much to develop the theory of isolation by distance for neutral alleles, he never extended this to shifts between adaptive peaks. As discussed below, the chance that a new adaptive peak spreads through the whole population depends on the mechanism of that spread.

Peak Shifts Involving Other Mechanisms of Divergence.—Discussion of the SBT has been dominated by the question of whether genetic drift, rather than other processes such as fluctuating selection, is likely to produce shifts between peaks in the surface of mean fitnesses. There are several reasons for this emphasis. First, drift is ubiquitous and described by straightforward theory. Second, the prominent controversy between Fisher and Wright identified Wright as an advocate for the importance of drift (Provine 1986, ch. 8). Finally, the first genetic examination of population structure in nature, the collaboration between Wright and Dobzhansky (see Lewontin et al. 1981), was an explicit attempt to see whether the structure of natural populations (especially values of Nm, which play a critical role in determining the ability of drift to produce phase I) could sustain the SBT. This still motivates many studies of F-statistics (Slatkin 1987).

Nevertheless, there are many processes besides genetic drift that can move populations to different adaptive peaks on either the individual or mean fitness surface. As emphasized by Fisher (1930) and Merrell (1994), environmental conditions may change fitnesses so that a population can reach a new adaptive peak by selection alone (e.g., Gibbs and Grant 1987). This may involve a change in direct selection on the trait (see the example of shell coiling in snails discussed in the "Evidence" section), or selection on a correlated trait (Price et al. 1993). Wright proposed a similar process (1931, 1940), noting that the peak shifts could be driven by random fluctuations in selection, which he regarded as equivalent to sampling drift (cf. Gillespie 1991, sec. 4.7). However, if fluctuating selection were allowed to drive the SB, Wright's theory would become indistinguishable from Fisher's view that populations are continually evolving under changing selective conditions. (See the correspondence on this issue between Fisher and Wright excerpted in Provine 1986, p. 272.)

Several alternative mechanisms for peak shifts have been analyzed. Whitlock (in press) reviews these, and makes the general point that, for parameter values giving an appreciable chance of a shift due to random drift, extremely small changes will trigger a purely deterministic shift. For example, Kirkpatrick (1982) found that the mean of a quantitative trait under disruptive selection may shift if its phenotypic variance increases. If the change in phenotypic variance or the fitness landscape is due to a change in extrinsic conditions (Kirkpatrick 1982; Petry 1982; Price et al. 1993), the shift is being driven purely by selection. Similarly, selection favoring the new homozygote can overwhelm selection against heterozygotes in models of chromosome evolution (Lande 1979; Hedrick 1981; Hedrick and Levin 1984; Spirito 1992). In two-locus models, a change in recombination rate or assortative mating can also trigger deterministic peak shifts (Williams and Sarkar 1994).

Populations can also come to occupy different peaks when different mutations arise in different places, are fixed by either drift or selection, and prove incompatible with each other. This scenario may involve randomness, but does not require small, isolated demes or passage through fitness valleys. The simplest case is where two alternative alleles at a locus are each favorable, but cannot mutate directly to each other (Johnson et al. 1995; see above). Other examples include stabilizing selection on a polygenic trait (Mani and Clarke 1990), and chromosomal fusions involving overlapping arms (Baker and Bickham 1986). Dobzhansky was one of the first to recognize this possibility. In a famous passage in *Genetics and the Origin of Species* (Dobzhansky 1937, p. 256), he noted that the sterility or inviability of hybrids can result from adaptive divergence alone. For example, an ancestral species may have the two-locus genotype *aabb*, which changes via selection to *AAbb* in one isolate and to *aaBB* in another. If the *A* and *B* alleles produce developmental problems when they co-occur in a hybrid, the interpopulation *AaBb* hybrids would be sterile or inviable. In these examples, populations evolve to fitter states without passing through unfit intermediates, always climbing uphill on the adaptive landscape.

If different peaks may frequently be reached by deterministic processes rather than drift, Wright's evolutionary views converge on those of Fisher. Nevertheless, although populations may diverge by processes other than those envisioned by Wright, it is still possible that population subdivision might allow fitter gene combinations to spread, leading to faster adaptation of the whole species through the third phase of the SBT.

The problem of finding the highest of many alternative peaks arises routinely in optimization theory. Following Holland (1975), there has been much interest in a class of computer routines known as "genetic algorithms" that mimic various features of evolution (Sumida et al. 1989; Forrest 1993; Paton 1994). These rely mainly on processes analogous to mutation and recombination to bridge unfit intermediates, rather than relying on drift in a subdivided population (for exceptions, see Sumida and Hamilton 1994; Wilson et al. 1994; Bergman et al. 1995). The closest analogy with the SB are "simulated annealing" algorithms (Kirkpatrick et al. 1983), which involve assigning "fitnesses" to trial solutions according to some optimization criterion. Random changes are made to the current trial solution having fitness W, leading to a new trial solution with fitness W^*. If $W^* > W$, the new solution is accepted, while if $W^* < W$, the new solution is accepted with probability $(W^*/W)^{2N}$. This stochastic process leads to alternative solutions with probabilities proportional to W^{2N}, which is analogous to Wright's equation (1). If the parameter $2N$ (analogous to deme size) is large, solutions are trapped at local optima; whereas if it is small, solutions are chosen nearly at random over the whole range of possibilities. The "simulated annealing" algorithm searches the solution space by increasing $2N$ from some low value; if the increase is sufficiently slow, convergence to the global optimum is guaranteed. The importance of local optima in optimization problems, and the close analogy with the simulated annealing algorithm that is used to escape them, supports the *potential* importance of Wright's theory in evolution. However, there is no reason to believe that natural populations behave like these search algorithms (e.g., real populations, unlike simulated annealing algorithms, do not suddenly crash and then grow slowly to achieve every adaptive advance) or that real fitness landscapes are as constrained by invariant local optima as the abiotic systems typically analyzed with these algorithms.

Conclusions.—Wright's formula (1) shows that small population sizes are required for the first two phases of the SBT. However, genetic drift is not required for populations to move between different adaptive peaks; in particular, many peaks will often be connected by ridges that require no loss of fitness. In addition, temporal and/or spatially varying selection may often be more plausible than genetic drift as mechanisms for switching adaptive peaks.

Phase III: Spread of New Adaptive Peaks

Thus far we have considered whether and how populations come to be subdivided into regions near different adaptive peaks (or more generally, diverge along incompatible paths). As noted, it is often hard to determine whether such divergence occurs by random drift, changing selection, random accumulation of different mutations, or a combination thereof. However, unless a species is effectively a single panmictic unit, we expect that to some degree its populations must form a mosaic of incompatible gene combinations; these combinations may be the raw material for the formation of new races or species. In this section, we consider how these gene combinations compete with each other during phase III of the SBT. If the SB makes a significant contribution to adaptation, as Wright envisaged, it is crucial that "fitter" combinations should tend to spread. Haldane (1959, pp. 140–141) judged this to be the weakest part of the theory, yet until recently it has received the least theoretical or empirical attention. Here, we show that adaptive peaks might spread through structured populations in several ways, but argue that fitter peaks need *not* spread and that alternative adaptations in different subpopulations might not be easily combined. Although several theoretical results (reviewed below) seem to favor the SBT, we believe that the problems of phase III constitute a central weakness of the theory.

Deterministic Mechanisms.—Wright thought of phase III

as a deterministic process, in which the fitter peak would inevitably spread because increased fitness would lead to increased density and increased emigration. In fact, such spread can occur deterministically without changes in population density or migration rate, and can favor the fitter peak simply by individual selection. In a one- or two-dimensional habitat, two adaptive peaks will meet in a cline with stable shape, whose width is maintained by a balance between disruptive selection and migration. This cline, termed a "tension zone" (Key 1968; Barton and Hewitt 1985), is not tied to any particular environmental gradient, and can in principle form anywhere. If one adaptive peak is fitter than the other, the cline will tend to move in its favor, so that (in the absence of barriers to gene flow) it may spread throughout the entire species (Bazykin 1969; Barton 1979). Though such models were first formulated as explanations of hybrid zones, they also describe phase III of the SB process.

Rouhani and Barton (1987b) present a model of tension zones and stochastic peak shifts that incorporate all three phases of the SBT. They model a species distributed continuously in two dimensions subject to disruptive selection on a quantitative trait having constant variance. Their model of peak shifts involves two stages. First, there are stochastic transitions to the higher peak in local neighborhoods that when established in a sufficiently large region, form a "critical bubble" that corresponds to the unstable equilibrium allele frequency in simple one-population underdominance models. Selection then drives an inexorable expansion as the tension zone that separates the two peaks moves outward.

In two dimensions, the probability of peak shifts in the first stage declines exponentially with the neighborhood size, N_b, and so local shifts become very unlikely for Nb greater than, say, 10. (As defined by Wright [1943], $N_b = 4\pi\rho\sigma^2$, where ρ is the effective population density and σ^2 is the variance in distance between parent and offspring at the time each reproduces.) The probability also depends very strongly on the difference in mean fitness between the two peaks. An inferior peak can never be established, while a superior peak has fixation probability near $\exp(-N_b/\alpha)$, where $\alpha = (\bar{W}_1 - \bar{W}_2)/(\bar{W}_1 - \bar{W}_v)$ is the difference in fitness between the peaks, relative to the depth of the valley. Surprisingly, this probability does not depend strongly on the overall strength of selection in the model. This is because although peak shifts through deep adaptive valleys are less likely to occur, they need only be established in a smaller area to expand. This analysis also applies to discrete demes, provided migration is faster than selection ($m \gtrsim s$), and extends to underdominant alleles (Barton and Rouhani 1991).

Lande (1989) noted that, in contrast to the Barton and Rouhani (1991) prediction, chromosome rearrangements that greatly reduce heterozygote fertility are fixed far less often than rearrangements that are less underdominant (even though they arise at comparable rates). This suggests that local variation in population density may make extreme fitness peak transitions much less probable. In particular, rearrangements with large fitness effects are less likely to meet the criterion $m > s$ needed to apply the theory, which assumes a continuous rather than subdivided population.

The probabilities of peak shifts are also influenced by factors that need not foster increased adaptation. For example, if frequency-dependent or epistatic selection favors the more common morph, the morph determined by dominant alleles will tend to spread (Mallet 1986; Phillips 1993), even when it reduces fitness. The reason is that in places where the alleles are equally frequent, the morph determined by the dominant allele will be more common, and hence will tend to increase. Other asymmetries in the genetic basis of the selected traits (e.g., epistasis) will cause tension zones to move. In general, mean fitness may *decrease* under the joint effects of selection and gene flow (Barton 1979).

The spread of a new adaptive peak may be aided if it is associated with increased population density or dispersal: this was the mechanism emphasized by Wright. However, contrary to Wright's belief that increases in mean fitness would generally increase population density, explicit models of density regulation show that population densities can remain constant or even decrease with increases in fitness (e.g., Prout 1980). Even when the population size does increase with $\bar{W}$, the effects of $\bar{W}$ on density are likely to be minor compared with its effects on transition probabilities between peaks. This has been demonstrated in recent analyses of phase III. Crow et al. (1990) considered selection favoring one particular gene combination, and acting against all recombinants (cf. Kondrashov 1992; Phillips 1993). In this model, a low rate of migration from a deme fixed for the fitter combination enables the combination to spread through a neighboring population initially fixed for a less fit combination; they interpreted this as supporting Wright's view of phase III. However, there are two difficulties with this interpretation. First, a low rate of one-way migration can establish uniformly *deleterious* alleles despite selection against them: the analysis of Crow et al. (1990) really shows the power of migration to swamp local selection (Barton 1992). Crow et al.'s (1990) model is especially sensitive to swamping, because selection favors only one particular gene combination, and so is ineffective when the novel genes are rare.

Second, Crow et al. (1990) assumed that gene flow would be predominantly from the fitter deme to the less fit and that emigration is proportional to the *excess* mean fitness, $\bar{W} - 1$. While this may occur in nature (e.g., in species having a fixed number of available territories), it seems implausible in general that emigration will occur only if there is a population "excess" or that emigration rates will be proportional to the population excess. In explicit models of density dependence, population density often varies with mean *relative* fitness (Roughgarden 1979, ch. 17; but see Prout 1980). Gavrilets (1996) sets out one- and two-locus deterministic models that show the difficulty in establishing a new adaptive peak (initially present in a single deme), even when the shift raises mean fitness and hence population density. Given that the rate of deterministic spread of higher peaks is roughly proportional to the fitness advantage of the new peak (say, s), the *additional* effect of increased migration attributable to raising densities would generally be negligible (of order s^2) unless selection is very strong. Thus, although migration can have a profound effect on mean fitness, as discussed above, fitness-dependent changes in migration rates will often have little additional effect. The argument that effects of mean fitness on population density are negligible also applies to movement of tension zones (Barton 1980; Barton and Clark

1990) and spread by successive peak shifts (Rouhani and Barton 1993; see below).

Spread of Adaptations by Single versus Successive Peak Shifts.—Wright envisioned a sequential process in which a new adaptive peak is established in one deme, which then sends out migrants to neighboring demes, triggering further peak shifts. Several mathematical analyses support this "chain reaction" effect. Lande (1985b) considered the spread of an underdominant chromosome arrangement through an array of effectively monomorphic demes ($Nm \ll 1$). He found that this process could give a strong advantage to the fitter homozygote, because the probabilities of migration-triggered shifts in either direction are in the ratio $(\bar{W}_1/\bar{W}_2)^{2N}$. For appropriate parameter values, slight differences in mean fitness between peaks can be amplified into large differences in the chances that they will invade neighboring demes.

Barton and Rouhani (1993) generalized Lande's (1985b) analysis by considering the full range of migration rates and analyzing disruptive selection on a quantitative trait as well as selection against heterozygotes. Peak shifts are most likely to favor the high-fitness genotypes that are commonest in the migrant pool, and each shift to this new peak makes it more likely that migrants will trigger further shifts. This positive feedback can give a very strong advantage to a slightly fitter adaptive peak. The advantage initially increases with the number of migrants (Nm). Above a critical Nm, however, the fitter peak cannot establish itself when the constituent genotypes are rare, because the most-fit genotype is swamped by immigration. Thus, the population may be trapped at a suboptimal peak (see appendix). Below the critical Nm, the bias toward the fitter peak increases with deme size and with the strength of selection (Ns); however, for large Ns, the rate of peak shifts becomes so slow that the process is ineffective. Overall, the process of successive peak shifts is most effective for intermediate values of Ns and Nm.

Although these analyses support Wright's SB process for some parameter values, they consider only the spread of a single allele or trait via a single peak shift. It is much harder to see how a complex adaptation could spread by this process. If such an adaptation requires favorable combinations of genes or traits from different demes (fixed in each by a peak shift), the components of the adaptation would be broken up by recombination when they coexist at low frequency in a single deme (Haldane 1959). Gavrilets (1996) gives a two-locus example in which a pair of genes are favorable in the double homozygote, but reduce fitness in other combinations. These can spread through a one- or two-dimensional array of demes, provided that migration is sufficiently high, and recombination sufficiently low. New adaptive peaks are not vulnerable to being broken up by recombination if spread is through colonization of empty sites by individuals that all come from a *single* deme, or by the movement of a multitrait tension zone through a continuous population, with no barriers to dispersal. However, these mechanisms of spread give a smaller advantage to the fitter peak than does spread by successive peak shifts (cf. Rouhani and Barton 1987a, 1993). The weaknesses of phase III become even more pronounced when one allows for realistic variation in population structure.

Extinction-Recolonization and Random Variation in Population Structure.—The simple models considered above assume a stable and relatively uniform population structure: either a uniform continuous density, a stepping-stone model, or an island model. Slatkin (1989) has argued that demographic conditions are not likely to remain constant long enough for the SB process to work as Wright envisioned. Moreover, when variation in population density is considered, adaptations favored by mass selection have a distinct advantage over those involving peak shifts. Although unconditionally favorable alleles can easily spread through a continuous species range that has variable population density, slight variations in population density or dispersal rate suffice to trap tension zones, thereby reducing the probability that phase III of the SBT will work (Barton 1979; Barton and Hewitt 1985). Similarly, in a species comprising discrete demes, adaptive peaks cannot spread deterministically if migration is weaker than selection ($m \ll s$) (Barton 1979; Lande 1985a); but initially rare, advantageous genes can spread easily across geographic barriers. Thus, a new adaptive peak may be trapped by local barriers to gene flow, and can escape only by substantial changes in population structure.

An extreme example of demographic instability is represented by local extinction and recolonization, a situation that Wright (1941) conjectured might favor the SBT. Lande (1979) found that if a species consists of a set of isolated demes that go extinct and are recolonized at random, the chance that an underdominant allele will become fixed in the whole species is equal to its initial frequency, irrespective of its effects on fitness. This model assumed that extinction and recolonization rates are independent of genotype, so that there is no tendency for the fitter adaptive peaks to persist or spread. Lande (1985b) further showed that differential extinction and recolonization are likely to significantly aid the fitter peak only if stochastic peak shifts are very much rarer than random extinctions. This would limit the time scale of adaptation by the SBT to the time scale of local population extinctions and recolonizations, which is clearly much slower than the time scale on which mass selection operates. Moreover, if large regions go extinct rather than individual demes, the variation in success of different regions is likely to be determined largely by extrinsic factors such as climate, leaving even less scope for the differential spread of particular gene combinations that increase fitness. It has often been noted that group selection of the sort postulated by Wright is likely to be ineffective relative to individual selection, simply because there are fewer selective events involving differential group versus individual performance (Maynard Smith 1976; Leigh 1983).

Summary of the Theory

In our discussion of the theory, we have distinguished three main issues: whether alternative adaptive peaks exist, whether and how populations shift between them, and how fitter peaks might spread through the whole species. The first issue is hard to address in general, since it requires an understanding of the relation between fitness and genotype at very many loci. At this stage, empirical evidence (see below) may be more helpful than are theoretical models. Nevertheless, all but the simplest models of evolution imply multiple equilib-

ria. Indeed, the existence of distinct species requires that such equilibria exist. A much thornier question is whether populations typically diverge by passing through unfit intermediate states (roughly speaking, by crossing adaptive valleys), or whether they usually evolve under selection alone, but diverge because they accumulate different mutations or because a changing environment provides "uphill" paths between alternative peaks. Theory shows that random sampling drift is a plausible cause of peak shifts in isolated demes only if the effective population size is fairly small relative to the depth of the adaptive valley. In two-dimensional habitats, peak shifts are not unlikely provided that neighborhood size or number of migrants are not too large and that the new peak also has a substantial advantage relative to the depth of the valley. The most critical theoretical problems are associated with phase III. In general, peak shifts may occur only in relatively sparsely populated parts of a species' range. The difficulty is then to see how novel peaks could spread from these regions through the whole species. Such spread often seems implausible because genes are likely to flow predominantly from central into marginal populations, and because empty patches are more likely to be recolonized from abundant regions. Hence, although peak shifts caused by genetic drift may occur in nature, they seem unlikely to form the basis of most adaptation

More generally, adaptation may occur through competition between alternative highly adapted genotypes established in different parts of a species' range. Whether this competition contributes to adaptation depends on whether fitter genotypes are more likely to spread. Wright's preferred mechanism—increased emigration from fitter demes—depends on a form of group selection, and is unlikely to be important relative to individual selection within demes. Although certain classes of models demonstrate large advantages for higher peaks, the same models show that factors not involved with adaptation, such as dominance, can critically influence spread (Mallet 1985; Phillips 1993). More important, heterogeneity in population structure, driven by extrinsic factors such as changing climate or availability of suitable habitat, is likely to critically affect the geographical distribution of alternative forms. Slight intrinsic fitness differences between adaptive peaks may be negligible in comparison with such random factors, as demonstrated by models of extinction-recolonization.

We have mainly considered shifts between particular *pairs* of adaptive peaks. There may in fact be very many possible peaks that increase fitness more or less independently. For example, a species may vary for alternative chromosome arrangements, for various traits under disruptive or stabilizing selection, and so on. The fitter peaks in different systems are likely to establish in different places, and so if the SB is to be effective, there must be some means by which these adaptive advances can be combined with each other. This may happen when the different gene combinations spread out and meet each other, as suggested by Wright (1964, fig. 12). However, this requires that the recombinant genotypes increase and spread from the region of hybridization, which can happen only under restrictive conditions. The recombinant genotype must be fitter than either parental type, and even then, can only be established above a critical threshold frequency. In two-locus models, where selection acts against heterozygotes but favors a recombinant homozygote, establishment of the favored double homozygote (combining two adaptive peaks) is certain if recombination is higher than a critical value, of the same order as the selection coefficients (Pialek and Barton, in press). In contrast, when individual peak shifts are based on interactions among several loci, low recombination rates are necessary for them to spread, as shown by Gavrilets (1996), because recombination breaks up the adaptive complex. When recombination is too frequent, the fitter recombinant will remain at low frequency, and might only be established by random drift (McCarthy et al. 1995). Hence, if we consider combining *two* multilocus adaptations from different demes, the conditions on recombination that would allow each to spread will often not allow them to be combined. Though more theoretical work would be welcome, there seems to us to be a considerable difficulty in establishing one particular combination of adaptive peaks, especially when each peak is vulnerable to being broken up by recombination.

A more serious difficulty in combining adaptive peaks is that if peaks spread mainly by extinction and recolonization, or by less drastic changes in population structure, then these forces affect all genes in the same way, and will therefore tend to keep sets of genes together. Successive expansion and contraction of the species' range will tend to bring genetic differences together in complex hybrid zones, as is observed (see below and Barton and Hewitt 1985). Thus, the instability in population structure that is required for adaptive peaks to overcome local barriers to gene flow also interferes with the interdemic selection of phase III that favors fitter peaks. Wright saw the SB as combining the advantages of sexual and asexual reproduction, by allowing demes to compete without recombining the well-adapted sets of genes they contain (Wright 1964, fig. 12). However, the difficulty of merging adaptations based on different sets of genes clearly saddles the SB with the many disadvantages of asexual reproduction.

The SBT can, in principle, operate: random drift can cause demes to move between adaptive peaks, the differential spread of these peaks can give a substantial advantage to fitter gene combinations for at least some parameter values, and adaptation can therefore be more efficient in a subdivided population (Lande 1985b; Rouhani and Barton 1987b; Moore and Tonsor 1994; Bergman et al. 1995). However, it seems implausible that this complex process dominates adaptation. Fitter peaks are unlikely to spread from the marginal populations where they are most likely to occur; the increased emigration caused by increased mean fitness has generally negligible effects; spread is most likely to be through random changes in population structure, determined by extrinsic conditions; and independent peak shifts cannot easily be combined. Nevertheless, genetic differentiation of local populations might contribute to the evolution of reproductive isolation, and migration between populations at different adaptive peaks can contribute to the maintenance of genetic variation (Goldstein and Holsinger 1992). Moreover, even if adaptive valleys are not commonly crossed, a subdivided population may respond more rapidly to selection for recessive alleles (Caballero et al. 1991) and favorable gene combinations (Goodnight 1995).

We conclude that adaptation via Wright's SB is theoreti-

cally implausible. Even if the theory were plausible, however, its ultimate value depends on whether it is more consistent with the evidence from nature than is the alternative of simple Fisherian mass selection.

The Evidence

For many, the main value of the SBT has been the metaphor of the adaptive landscape, in which evolution is seen as a form of hill climbing. If the theory is to have more than metaphorical value, however, it must be subject to scientific tests. According to Crow (1991, p. 973), such testability was of minor concern to Wright:

> Wright never showed much interest in experimental tests of his theory; his arguments were based on plausibility and analogy. He thought that much of evolution, the steady improvement of adaptation, could happen by mass selection acting on the additive component of the genetic variance, as Fisher said, but he thought that evolutionary creativity demanded something more. This might not happen often, and hence would be difficult to test in nature, but would be important when it did happen.

We do not agree that the SBT was meant to explain only rare evolutionary novelties. Even a cursory look at the four volumes of *Evolution and the Genetics of Populations* shows that Wright regarded his theory as an actual description of evolution in many if not most species. Moreover, we feel that, like all theories, the SBT *must* be testable if it is to be considered an important explanation of evolution. Ultimately, its validity must depend on finding adaptations that are explained more plausibly by Wright's mechanism than by Fisherian mass selection.

There are several ways to "test" the SBT. First, we can ask whether, as Wright envisioned, it accounts for most adaptations. To perform this strong test, one must show either that many adaptations chosen randomly from nature are more likely to have evolved by the SB process than by mass selection, or that subdivided populations are in some sense better adapted than panmictic populations and that this difference involves drift and selection among adaptive peaks. Given the difficulties of finding suitable material on which to test the theory, and our inability to reconstruct the path of evolution, such a test seems impossible. We will never have even a rough idea of how often the SBT explains adaptation.

Second, we can ask whether *individual* cases of evolution in nature are better explained by the SBT than by mass selection. Some positive results would support the SBT, but not Wright's assertion that it is ubiquitous. We discuss several such empirical studies below. Although one or more phases of the SBT may have operated in some of these examples (particularly those involving chromosome evolution and cytoplasmic incompatibility), none provides a convincing example of adaptation via the complete SB process.

Third, we can determine whether the SBT can be demonstrated in the laboratory or in artificial selection experiments. This would at least show that the theory was of *potential* importance in explaining adaptations in nature. We discuss the relevant evidence below and argue that the few existing tests do not support the SBT.

Finally, even if we cannot show that the SBT as a whole has operated in nature or the laboratory, we may be able to show that the *conditions* necessary for its operation sometimes occur in nature. Can we find evidence, for example, that genetic drift has successfully opposed natural selection? Is there substantial epistasis for fitness among segregating alleles? Is a species subdivided into partially isolated populations? We must remember, of course, that demonstrating some of these conditions does not provide particularly strong support for the SBT, which requires that epistasis, drift, population structure, and differential migration operate *together* in the evolution of a single adaptation.

Nevertheless, operation of individual components of the SBT is still of potential importance. We may find evidence (and we do) that genetic drift can successfully oppose natural selection. This demonstrates that drift may play a role in adaptation, an idea for which we have surprisingly little evidence. The finding of important epistasis for fitness has implications for the reversibility of evolutionary change, and observations of population structure have implications for the possibility of local adaptation and sympatric speciation. While not completely supporting the SBT, such findings could nevertheless expand our view of evolutionary change beyond Fisher's scheme of directional selection operating in very large populations. In the following discussion we will call attention to cases in which *components* of the SBT have been found by observation and experiment.

When considering whether observation or experiment support the entire SB process, we take a fairly rigorous stance, demanding that all the phases of the theory be demonstrated in any given case. That is, we require evidence that an adaptation involves crossing adaptive valleys against the force of selection, a subsequent climb to a new and higher adaptive peak, and a spread through the population by Wrightian interdeme selection. Moreover, as there is nearly always an alternative explanation involving Fisherian mass selection, we also require that that explanation be falsified. This does not mean that we consider the Fisherian alternative the correct explanation, only that we embrace it as the null hypothesis. Some readers may consider these criteria too stringent and our parsimony argument unconvincing. Recognizing this disagreement, we nevertheless regard Fisherian mass selection as the proper null hypothesis because there is ample evidence for its operation in nature and the laboratory, and, as shown in the previous sections, theory demonstrates that it is effective over a much wider range of conditions than is the SB process.

Adaptive Peaks and Transitions among Them

Climbing Ridges versus Crossing Valleys.—When considering whether the SBT can operate in nature, the first question is whether fitness landscapes have multiple peaks. Simple models, such as stabilizing selection on an additive polygenic trait, do indeed produce such peaks. To assess the SBT, however, one must determine whether populations actually occupy isolated peaks, whether these peaks *remain* isolated despite changing conditions, and whether the transitions be-

tween peaks involve temporary decreases in population fitness. Two high-fitness genotypes might, for example, be connected by some chain of fit intermediates that would allow transitions due to natural selection, perhaps aided by changes in the environment that modify the adaptive landscape. Although it may be hard to imagine "fit" intermediates between radically different morphologies, such as terrestrial mammals versus cetaceans or the jaw-joint differences distinguishing mammals and reptiles, the occurrence of such intermediates in the fossil record demonstrates the reality of "implausible" ridges (e.g., Crompton 1963; Carroll 1988, pp. 391–407; Thewissen et al. 1994).

A microevolutionary example of such ridge-climbing is the divergence of warning color in the burnet moth, *Zygaena ephialtes*. In central Europe, this moth is a Müllerian mimic of *Z. filipendulae*. In Italy, however, it mimics another moth, *Amata phegaea*. The different pattern and color of the southern race is controlled by dominant alleles at two loci. The two races meet in a narrow hybrid zone in Austria. The narrowness of the hybrid zone implies that in this location there are distinct adaptive peaks. However, one recombinant morph is much more common than the other, presumably because it has a fair resemblance to the southern model, *A. phegaea*. Where *A. phegaea* is sufficiently abundant relative to the model, this resemblance could increase the fitness of the recombinant morph enough to remove the valley separating the peaks, and allow divergence by natural selection alone (Turner 1971).

The existence of species that remain distinct despite gene exchange, as well as the reduced fitness of interspecific hybrids, is sometimes taken as evidence for phases I and II of the SBT. However, these phenomena show only that an adaptive valley *currently* separates the alternative genotypes, not that the valley was actually crossed (Muller 1940; Lande 1989; Orr 1995). As shown by Dobzhansky's two-locus model described above, it is entirely plausible for populations to diverge by natural selection alone, creating an adaptive valley between them that no population ever crossed. After reviewing these considerations and the extensive literature on gene interactions, Whitlock et al. (1995) concluded that although there is ample evidence for epistasis, it does not necessarily produce adaptive landscapes in which peaks must be reached by crossing adaptive valleys.

Population crosses therefore provide no critical evidence for the traversing of adaptive valleys unless one can reconstruct the evolutionary pathway and show that it involves a loss of fitness. If, however, an adaptation is thought to have evolved by the SB process, then hybrids between individuals having and lacking that trait should show reduced fitness. Negative results of such a comparison would therefore militate against the SBT.

Inferences from F_{ST} values.—One of the main outcomes of the SBT has been to stimulate empirical studies of population structure. The goal of such studies is usually to determine whether natural populations are sufficiently subdivided to permit substantial genetic drift (Slatkin 1987). In Wright's view, "The importance of estimating inbreeding coefficients of the type F_{ST} for demes within more comprehensive natural populations is in judging the likelihood of evolution of these populations by the shifting balance process" (Wright 1978, p. 61). There have indeed been many such estimates, though we are not aware of a comprehensive review of F_{ST}-statistics from natural populations. A rough impression would be that such values are typically around 0.05–0.1, but can be as high as 0.7 in species such as salamanders that have limited dispersal (see Slatkin 1981; McCauley and Eanes 1987). The theoretical analysis given above, however, shows that many of these attempts may be misguided. The SB process does not require that a species meet the criterion that $Nm \approx 1$ throughout its range. It might suffice instead for a *subset* of the range to be subdivided, provided that adaptations could spread from these populations through the species as a whole. The only empirical finding that could invalidate the SBT would be the observation that there is either a large number of migrants per deme (Nm) or a large neighborhood size (N_b) *everywhere* within the range of a continuously distributed population.

It is also worth noting that most efforts to interpret F_{ST} values are based on the assumption that geographical variation in allele frequencies is determined largely by genetic drift. This approach is valid only if sampling drift is responsible for both the observed geographic patterns and for the shifts between the adaptive peaks that are of interest.

How Important Is Genetic Drift?

The SBT requires that genetic drift be sufficiently pervasive and strong to influence the evolution of adaptations. In its textbook characterization, genetic drift stands beside natural selection as one of the two great forces of evolutionary change. However, the evidence for selection, particularly that affecting morphological characters (Endler 1986), is far stronger than the evidence for drift, whose evolutionary role is still debated. There is no question that drift dominates the fate of new mutations and other rare alleles (Haldane 1932, p. 200; Gillespie 1983; Kimura 1983), and it may also explain certain aspects of molecular evolution, such as the consistently high polymorphism and rates of evolution of third positions and pseudogenes. Other molecular evidence for drift includes the existence of reduced heterozygosity in small island populations (Berry 1986; Caccone et al. 1996) or in widespread species whose relatives are genetically quite variable (e.g., Menken 1987); although such observations, like those of high frequencies of deleterious alleles in human isolates (e.g., Risch et al. 1995) might reflect a limited number of initial colonists rather than sampling events occurring after colonization. Our inability to cite many cases in which drift has affected characters beyond protein and DNA sequence does not mean that sampling events are unimportant, only that they are difficult to document.

Wright often cited two lines of evidence for the importance of genetic drift: the nonadaptive nature of characters distinguishing related taxa, and the random geographic distribution of flower-color morphs in *Linanthus parryae*. Although neither was considered as strong support for the entire SBT, they constituted for Wright powerful evidence for phase I.

Diagnostic Characters.—In his early writings on the SBT, Wright often discussed the work of systematists, particularly the British zoologists O. W. Richards and G. C. Robson (1926; see also Robson 1928; Robson and Richards 1936).

After critically analyzing experiments purporting to show selection on morphological characters, Robson and Richards concluded that the majority of phenotypic differences between populations, subspecies, and closely related species were nonadaptive. Wright (e.g., 1931, 1932, 1939, 1940) concluded that these nonadaptive differences were fixed by genetic drift. While Robson and Richards' arguments against selection were of some value—many of the early "proofs" of adaptation were indeed quite weak—they fail to show that most diagnostic differences were unaffected by selection. Robson and Richards did not conduct their own experiments on these characters, but simply concluded that species differences were nonadaptive because they could imagine no selective explanation. A failure of imagination does not, of course, constitute strong evidence for neutrality. Moreover, some of the traits discussed by Robson and Richards, such as mimicry and genital morphology, are now known to be affected by selection (e.g., see Eberhard 1985 on genitalia). In view of our continuing ignorance about the selective significance of most diagnostic characters, the arguments of Robson, Richards, and their fellow systematists can no longer be taken to support the ubiquity of genetic drift.

Color Polymorphism in Linanthus parryae.—Wright's prime example of genetic drift affecting morphology, an example that he cited throughout his life, was flower-color polymorphism in the desert plant *L. parryae*. The geographic distribution of the two color morphs (white vs. blue, based on two alleles of a single locus) was the subject of intensive investigation by Carl Epling, Theodosius Dobzhansky, and others. Their empirical work and Wright's analysis of the data are recounted in Provine (1986, ch. 11). We summarize this work briefly and then describe new data bearing on its interpretation.

Linanthus parryae is an obligate outcrossing annual pollinated by beetles. It is distributed not in discrete demes that correspond to the classic population structure of the SBT, but as a quasi-continuous population stretching across the California desert. There are occasional bad years in which no plants appear, but in other years plant density can be as high as several hundred individuals per square foot. Wright's analysis of geographic variation in flower color was designed to show that random drift could cause spatial differentiation of a morphological character across the range of a species, even if the species did not comprise discrete demes.

As originally described by Epling and Dobzhansky (1942), population samples from California showed a U-shaped distribution of color morphs, so that areas of polymorphism were rarer than those fixed or nearly fixed for one or the other color morph. Analyzing these data using *F*-statistics (and assuming that the alleles were neutral), Wright (1943) concluded that this pattern was consistent with genetic drift operating within populations having a neighborhood size between 10 and 30 plants. The differentiation between larger areas was, however, too large to be accounted for by pure drift, and for this Wright (1943, p. 155) invoked regional selection in favor of multilocus "adaptive systems," in which flower color played a role.

Epling and his colleagues continued work on this species for 20 years, and in 1960 published new observations contradicting Wright's conclusions. First, Epling et al. (1960) showed that allele frequencies in small areas were quite stable over time, and concluded that only selection and not drift could explain such stability. Second, Epling et al. found that *L. parryae* had an extensive bank of dormant seeds that could remain viable for at least six years and probably far longer. Such a dense seed store would certainly reduce the effects of drift by limiting the effects of drought on population size.

In 1978 (ch. 6), Wright reanalyzed these newer results, concluding again that allele-frequency changes between 1944 and 1966, as well as the geographic variation among small areas, were due primarily to drift. Responding to the argument that seed banks would increase the effective population size, Wright (1978, p. 222) postulated long periods of drought that might deplete the seed store and increase the role of drift.

There is, however, no evidence for such bad periods, or for Wright's conjecture that the seed bank would be depleted after even 10 years of drought. In the last 50 years, for example, there has been no six-year period in which rainfall was so low as to prevent the appearance of *L. parryae*, and plant density did not significantly decrease following the severe droughts of the late 1980s and early 1990s in California (D. Schemske, pers. comm.)

Recent studies by Schemske and Bierzychudek (pers. comm.) have again raised the possibility of selection on flower-color morphs. First, their reanalysis of 22 years of data by Epling et al. showed that the frequency of blue flowers within transects decreased significantly more often than expected by chance. In addition, transplant experiments, studies of clines over short distances, and direct measurement of seed production all implied fairly strong selection between morphs, probably due to differences in soil moisture. Finally, at one site there was evidence for fluctuating selection on flower color, raising the possibility that the polymorphism could be maintained by temporal variation.

Clearly, the last word has not been said about geographic differentiation of flower color in *L. parryae*. However, in view of the increasing evidence for selection among color morphs, and for persistent seed banks, local geographic differentiation in *L. parryae* cannot be taken as convincing evidence for either the SBT or the effects of genetic drift on a morphological character.

Possible Examples of the Shifting Balance Process

The SBT requires more than simple differentiation of populations by drift. To produce new adaptations, drift must *counteract* selection, driving populations into adaptive valleys until natural selection draws them up a different peak. Most of the evidence that drift can overcome selection comes from the few known cases in which unambiguously deleterious alleles have increased in frequency, including the classic examples of small human isolates with high frequencies of genetic diseases or deformities (e.g., McKusick 1978). These examples generally involve founder effects, rather than repeated sampling, but the consequences are the same.

We are aware of only a few traits whose evolution may have involved the passage through an adaptive valley. We discuss below six of the best-documented cases. Although several of these probably do not involve increased adaptation, and none strongly support all three phases of the SBT, all

show the potential importance of genetic drift in influencing the evolution of adaptations.

Chromosome Evolution.—Perhaps the most widely cited example of the SB process is the evolution of new chromosome arrangements, particularly the fixation of those for which the heterokaryotype is less fit than either homokaryotype. This underdominance can result from either improper segregation (as with translocations or centric fusions) or the production of aneuploid gametes through crossing-over (as with pericentric inversions). There are many cases of populations and closely related species fixed for different inversions or fusions, and one sometimes sees abnormal meiosis and partial sterility in interpopulation hybrids (White 1978; King 1993). As with chirality in snails (discussed below), the transition between the two homokaryotypic states would seem to involve a loss of fitness and passage through an adaptive valley. Such chromosome evolution is widely accepted as a one-factor peak shift. Maynard Smith (1989, pp. 181–183) observed, for example, that "the best reason for thinking that [adaptive] valleys are sometimes crossed is that some kinds of structural changes in chromosomes do seem to imply intermediates of lower fertility." Wright repeatedly cited chromosome evolution as demonstrating the joint action of drift, selection, and population size that was necessary for the SB process (e.g., 1977, p. 473; 1982, p. 441).

There is, however, surprisingly little evidence that chromosome rearrangements are fixed when genetic drift overcomes the underdominance of heterozygotes. The existing data have two problems. First, many heterokaryotypes may not be less fit than the corresponding homokaryotypes. Heterozygotes for different chromosome arrangements vary widely in fitness (Ruiz and Alverola 1983; Sites and Moritz 1987; John 1981; Coyne et al. 1991, 1993; Nachman and Searle 1995). Within each class of arrangement, one finds some examples in which the expected underdominance simply does not occur, probably because recombination in heterozygotes is inhibited or segregation is regular. Reed et al. (1995), for example, reported almost no meiotic problems in chromosomal heterozygotes between two races of the lizard *Sceloporus grammicus*. Heterozygotes for fissions showed balanced segregation, and heterozygotes for a pericentric inversion underwent heterosynaptic (straight) pairing, eliminating the recombination that yields aneuploid gametes. Likewise, surveys of artificially induced pericentric inversions in *D. melanogaster* (Coyne et al. 1991, 1993) showed that while many were semisterile when heterozygous against the Standard sequence, nearly 40% of them showed no loss of fitness within the limits of detection (roughly a 1% difference in female fertility).

Such variation in fitness implies that the chromosome arrangements most likely to be fixed among populations or species are those with the least underdominance. In accordance with this view, Lande (1984, 1989) noted that classes of arrangements most underdominant as heterozygotes (e.g., translocations) are fixed much less often than those that are more benign (e.g., inversions and centric fusions). Without direct evidence for underdominance, then, the observation of populations fixed for different chromosome arrangements tells us little about the role of genetic drift or its opposition to selection. In fact, we know of no cases in which single rearrangements fixed among populations or species have been shown to be underdominant when heterozygous with the ancestral form.

In addition, even if chromosomal rearrangements are involved in hybrid sterility, this may result from heterozygosis for *multiple* rearrangements that individually are not underdominant. The classic example of this is monobrachial fusion, in which a single chromosome fuses with two different chromosomes in two isolated populations (e.g., chromosome A fuses with chromosome B in population 1, but with chromosome C in population 2). Meiotic problems, which may not occur in heterozygotes for a single fusion, may be profound in hybrids when four different chromosomes try to pair (Baker and Bickham 1986). The individual fusions may thus be fixed as neutral or adaptive traits, and the underdominance of interpopulation hybrids would occur without either population having crossed an adaptive valley. This resembles Dobzhansky's (1937) scenario for the evolution of hybrid sterility or inviability as a result of adaptive divergence. As Bickham and Baker (1980), Baker and Bickham (1986), Sites and Moritz (1987), and Nachman and Searle (1995) note, most of the well-known examples of chromosome evolution in vertebrates do indeed involve monobrachial fusions, including the famous case of chromosome races in *Mus domesticus* (Nachman and Searle 1995). In *Mus*, there is no convincing evidence that single fusions are underdominant, but much evidence that heterozygotes for multiple monobrachial fusions are deleterious. In other cases, such as the pericentric inversions in grasshoppers made famous by the work of M. J. D. White (1973, 1978), heterozygous chromosomes all pair without recombination, so that there is no loss of fitness (John 1981).

Two other pieces of evidence, however, show that some chromosome rearrangements fixed in nature may have produced unfit heterozygotes. First, among vertebrate taxa there is a significant negative correlation between the rate of chromosome evolution and allozyme heterozygosity (Coyne 1984). Because allozyme heterozygosity probably reflects the effective size of a species, Coyne concluded that species with fewer individuals undergo more rapid chromosome evolution, as one might expect if the new karyotypes are underdominant and must be fixed with the help of drift. This argument, however, is not completely convincing, as such a correlation might be expected even if the arrangements are not underdominant. Because nearly all chromosomes carry deleterious recessive alleles (Crumpacker 1967; Lewontin 1974; Simmons and Crow 1977), fixing a unique new arrangement almost inevitably means fixing a chromosome that is deleterious in comparison to other karyotypes in the population. Genetic drift would aid this fixation, and might account for the correlation observed by Coyne (1984) without the occurrence of any underdominance. Thus, although the correlation probably supports the role of drift in chromosome evolution, and may show that drift can counteract selection, it does not necessarily show that adaptive valleys have been crossed.

The strongest evidence for the underdominance of some chromosome rearrangements is their geographic distribution: in most cases, populations are monomorphic over large areas

and chromosome races are separated by narrow clines (e.g., see White 1978). However, these clines are often broad relative to the dispersal range, implying that they are maintained by only weak selection (Barton and Hewitt, 1985, Fig. 1). It is thus plausible that chromosome rearrangements are established when mild meiotic problems in the heterozygote are overcome by some force besides drift, such as direct selection or indirect selection through hitchhiking (Charlesworth and Charlesworth 1980; Nachman and Searle 1995). Moreover, there is no evidence that the fitter karyotype tends to spread after fixation, as required by the SBT. Chromosomal races are usually distinguished by many other genetic differences (Barton and Hewitt 1985), so that the fate of these races may well be determined by the net effects of all these differences, as well as by random expansions and contractions of population size. Reviewing all of these explanations for rapid chromosome evolution in mice, Nachman and Searle (1995) conclude that we cannot distinguish among them, or say with any certainty that single heterozygotes suffer a loss of fitness.

Finally, even if drift does fix new arrangements that form deleterious heterozygotes, this will produce increased adaptation only when the new homokaryotype is fitter than its ancestor. This is has not been demonstrated in any case of chromosome evolution. Nevertheless, even if we know nothing about the relative heights of chromosome peaks, the observation that a new chromosome arrangement has become fixed despite net underdominance would give evidence for the operation of phases I and II of the SBT. To document phase III, one must also show that the new arrangement spread among populations by either interdeme selection or a wave of advance, as a result of some adaptive advantage.

Despite the difficulties of demonstrating all three phases of the SBT in chromosome evolution, there is suggestive evidence that genetic drift may occasionally be involved, and that it might be strong enough to counteract underdominance. It is important to establish how often single rearrangements fixed in nature are underdominant as heterozygotes, as this could provide fairly strong evidence for single-factor peak shifts.

Müllerian Mimicry in Heliconius Butterflies.—One of the best understood examples of how natural selection can maintain alternative stable equilibria is found in Müllerian mimicry. Here, two or more distasteful species evolve similar warning patterns that advertise their unpalatability. Individuals deviating from the common pattern suffer increased predation because they are not recognized by predators as distasteful. Such frequency-dependent selection maintains homogeneous warning patterns both within species and between species that are part of the same "mimicry ring."

Given such strong stabilizing selection, it is hard to see why a given species should show different patterns in different places. In the best studied example, *Heliconius erato* and *H. melpomene* are divided into a mosaic of parallel geographic races that are separated by narrow hybrid zones (Turner 1971, 1977; Turner and Mallet 1996). Two explanations for such geographic divergence have been proposed. Brown et al. (1974) argue that random changes in the relative abundance of alternative model species (biotic drift) can allow a species to shift between mimicry rings. Mallet (1986), on the other hand, invokes the SB process, suggesting that random genetic drift establishes a novel pattern in a small region. Once this new pattern becomes sufficiently abundant, it may spread via a moving tension zone. Both models may be facilitated by the contraction of *Heliconius*' range into small forest refugia, a biogeographic scenario that is quite controversial. However, both can also operate within a continuous distribution, so that the argument over the origins of geographic variation does not depend on whether refugia existed.

Turner and Mallet (1996) review the evidence for each hypothesis. There are clear cases in which species switch between alternative patterns as a result of changes in the abundance of those patterns (e.g., the *Zygaena* example discussed above). Strong evidence for biotic drift also comes from the fact that pattern races differ by a few major genes that are subject to strong selection. A switch between rings is most likely to occur by a single mutation that bridges the gap between alternative phenotypes, whereas genetic drift is most unlikely to establish an allele that has a strong selective disadvantage when rare. Finally, there is no evidence for drift and geographic subdivision in *Heliconius*: allozyme studies show $F_{ST} \approx 0.012$–0.025 (Turner and Mallet 1996). The strongest argument against the biotic drift hypothesis is that it cannot account for the origin of new patterns, but only for switches between existing patterns. Thus, some explanation is required for at least the initial diversification, and indeed for the origin of aposematism itself. These phenomena might involve genetic drift, but might also be due to fluctuations in predator abundance, sexual selection, or selection on thermoregulation.

It is instructive to compare Müllerian mimicry in *Heliconius* with chromosome evolution. In *Heliconius*, selection on pattern is apparently stronger than on most chromosome rearrangements; for example, Mallet and Barton (1989) and Mallet et al. (1990) estimated that rare color patterns suffer a selection intensity of roughly 50%. This makes it harder to envisage *any* forces that could overcome the deep adaptive valley, including both random sampling drift and fluctuating selection. Indirect evidence for some role of drift in chromosome evolution comes from correlations with allozyme heterozygosity; moreover, many of the most striking examples of chromosome races come from sparsely distributed species with limited dispersal. In contrast, *Heliconius* is highly vagile, abundant, and shows little genetic divergence among populations.

This discussion has concentrated on the initial establishment of divergence (phase I). It is plausible that, once established, pattern races might spread through a continuous habitat behind moving zones of hybridization (phase III; Mallet 1986; Turner and Mallet 1996). Indeed, there is some direct evidence of movement of hybrid zones in *Heliconius* (Mallet 1993). However, there is little evidence that such spread has been adaptive; indeed, Mallet (1986) has shown that those patterns determined by dominant alleles will tend to spread, even if they have lower mean fitness. Thus, although Turner and Mallet (1996) suggest that the SBT is the most probable explanation for Müllerian mimicry in *Heliconius*, we believe that there is not yet enough evidence to strongly support any explanation.

Tristyly in Eichhornia paniculata.—A well-documented case of genetic drift opposing selection is the destabilization

of the genetic polymorphism for tristyly in small populations of several angiosperm species (reviewed in Barrett 1993). In all known cases, the polymorphism involves three morphs differing in style length and anther height (hereafter called L, M, and S morphs, standing for "long," "mid," and "short"). The phenotypic differences are controlled by two diallelic loci that interact epistatically, and the polymorphism at each locus is maintained by frequency-dependent selection promoting efficient pollen transfer (Lloyd and Webb 1992) and/or the avoidance of inbreeding (Charlesworth and Charlesworth 1979). At equilibrium in large populations, all three morphs are present at equal frequencies (Heuch and Lie 1985). Given the dominance and epistasis involved in this genetic system, however, the *alleles* at each locus are not present in equal frequency, and theory indicates that small populations will often experience stochastic loss of the rarer dominant allele governing the S morph. This produces populations dimorphic for the L and M forms (Barrett et al. 1989).

Observations of the aquatic annual *Eichhornia paniculata* in northeastern Brazil confirm that very small populations do indeed frequently lack the S morph. The loss of these morphs has been observed directly from one year to the next in association with random fluctuations in population size (Husband and Barrett 1992a,b). While the dissolution of tristyly may occasionally result from selection caused by the absence of long-tongued bees that pollinate the S morph (Barrett et al. 1989), the correlation of morph number with population size argues that breakdown of the polymorphism involves genetic drift.

Because tristyly is a stable equilibrium, the stochastic loss of morphs represents the displacement of a population from its adaptive peak. While gene flow could restore missing alleles and reestablish tristyly, there is good evidence that once populations lose the S morph, they are vulnerable to invasion by genetic modifiers of the M morph. These modifiers produce a "modified M" morph having altered stamens that allow self-pollination, and these modified morphs are fixed by selection. The evolutionary breakdown of tristyly in *E. paniculata* can therefore be viewed as a shift from one adaptive peak (outcrossing) to another (selfing) through an unstable state in which populations contain both outcrossing and selfing variants.

Is this then a peak shift leading to greater population fitness, that is, an example of the full SB process? The answer is not clear. First, the SBT requires changes in the frequencies of already existing genes, whose evolution drives the population first into an adaptive valley then up a new adaptive peak. In the case of *E. paniculata*, the movement up the second peak requires the fixation of alleles not involved in the original tristyly, and which are not subject to selection until the S morph is lost. More important, selection for selfing results from new ecological pressures that act in small populations, including selection for reproductive assurance (based on the scarcity of pollinators) and the purging of the genetic load that normally blocks the evolution of selfing. This scenario of new selection regimes based on population size is not an explicit part of the SBT. In addition, we do not know whether the transition to monomorphism associated with the fixation of the selfing morph leads to a net increase in population fitness. While the selfing variants are higher in fitness than other morphs in dimorphic and monomorphic populations (Barrett et al. 1989), in tristylous populations they are no more fit than the outcrossing morphs; and there is no evidence that they have spread among tristylous populations, as is required by phase III of the SBT (Kohn and Barrett 1994).

Nevertheless, *E. paniculata* provides a fairly convincing example of genetic drift displacing populations from their adaptive peaks and paving the way for a new adaptation that would not be reached by large panmictic populations. This certainly demonstrates phase I of the SBT, but the occurrence of phase II is questionable, and that of phase III unlikely.

Coiling in Snails.—Gittenberger (1988) and Orr (1991) reviewed the literature on the direction of coiling (chirality) in terrestrial snails. Some species possess both right- and left-handed morphs, a polymorphism typically caused by two alleles of a single maternally acting gene. The two coiling morphs are somewhat reproductively isolated because of their great difficulty in copulating, which causes strong selection against rare morphs. Polymorphic populations are hence genetically unstable, explaining why they are hardly ever observed. Each coiling type represents a stable equilibrium, and transitions between them are opposed by selection. Nevertheless, this evolutionary transition has occurred repeatedly, both among closely related species and among populations of single species, in which isolated "inverse populations" have occasionally been described. Because of the reproductive isolation between morphs, these transitions have been considered as examples of both sympatric speciation (Gittenberger 1988) and single-gene speciation (Orr 1991). The theoretical work of Orr (1991) and Van Batenburg and Gittenberger (1996) shows that the critical factors permitting change in coiling are maternal inheritance (which allows the new allele to attain a high frequency before many individuals show the corresponding phenotype), the dominance of the new chirality allele, and small population size ($N_e < 20$), an assumption not unrealistic for snails.

While the transition between coiling types would seem to be a *prima facie* case of phases I and II of the SB process, there are some potential problems. First, in some cases the change in coiling might actually be caused by selection alone. If a species is sympatric with a related species having the same direction of coiling, selection to reduce maladaptive hybridization might operate to reverse the chirality in the area of overlap (assuming that hybrids are somewhat sterile or inviable and there are no strong premating isolating mechanisms). The fixation of a new chirality would thus occur via reinforcement or reproductive character displacement, not by genetic drift. In these cases there may be no adaptive valley. Such an adaptive transition has apparently occurred in the Tahitian snail *Partula suturalis*, in which the narrow border separating the coiling morphs corresponds to the boundary of a related species (Johnson 1982). While there have been some reports of isolated "reversed" populations of species in the absence of a sympatric congener (Gittenberger 1988), these lack adequate documentation.

Finally, there is no evidence that the transition between coiling types leads to a net increase in fitness, as required by Wright's theory. However, such an increase is at least

possible, particularly if different coiling alleles have pleiotropic effects on fitness (Gould et al. 1985). It is therefore possible that chirality switches represent examples of single-locus peak shifts, particularly if these switches occur in the absence of selection for reproductive isolation from another species.

Paired Substitutions in RNA.—Interspecific comparisons of molecules show that substitutions affecting protein or mRNA sequences sometimes occur in pairs. It has been suggested that the first of these mutations might be deleterious and the second advantageous because it removes the deleterious effects of the first (Kimura 1985a,b; Stephan and Kirby 1993). Empirical support for this scenario is strongest for some mRNAs that have reverse-complementary sequences occurring near each other, possibly forming internal-pairing structures that stabilize the molecule. (Although there is no direct evidence for such pairing, the length of some complementary sequences does support its likelihood, as does the linkage evidence described below.) A single base-pair change might disrupt this pairing and destabilize the mRNA, but a second substitution in the complementary position could restore proper pairing. A recent analysis of pre-mRNA from the alcohol dehydrogenase (*Adh*) locus in *Drosophila* (Stephan and Kirby 1993; Kirby et al. 1995) showed many possible "compensatory mutations."

Although the substitution of two complementary mutations could have occurred via an SB process, this would be true only if the second substitution were fixed after the first. If both nucleotides were substituted together, however, the entire process could occur through directional selection, and drift would not be required. Such joint substitution indeed seems to have happened in the evolution of *Drosophila Adh* pre-mRNA. Stephan and Kirby (1993) found a strong negative correlation between the degree of covariation between sites and the number of bases that separate them in the molecule. Linkage is required only if mutations are being substituted simultaneously, for if the second mutation arose after the first was fixed, there would be no requirement for close linkage between the sites. Although we have no empirical information about the population frequency of the first mutation at the time when the second arose, Stephan (1996) has shown theoretically that this correlation occurs most readily when the first mutation is initially at a very low frequency. These observations suggest a scenario in which the initial mutation was kept at low frequency by counterselection, and then a second "compensatory" mutation arose nearby, allowing both to be fixed together.

In the best-studied case, then, "compensatory" substitutions appear to have been fixed not by an SB process, but by directional selection. Similar studies of linkage would be welcome for other molecules showing "covarying" substitutions.

Cytoplasmic Incompatibility in Drosophila simulans.—Cytoplasmic incompatibility in *D. simulans* and many other insects is caused by a maternally inherited microbe related to *Wolbachia pipientis* (Hoffmann et al. 1986; Turelli and Hoffmann 1995). Horizontal transmission of the microbe is vanishingly rare, so it spreads through a species only by an increase in the number of infected mothers. These infected females produce a comparable number of progeny whether they mate with infected or uninfected males. Uninfected females, however, produce significantly fewer progeny when they mate with infected males than with uninfected males. This produces a systematic tendency for the frequency of infection to increase within and among populations. However, Turelli and Hoffmann (1995) have argued that the spread of the infection through *D. simulans* populations in California also requires that some local populations must rise above an unstable equilibrium frequency of infection that is produced by a balance between imperfect maternal transmission (which tends to eliminate the infection when it is rare) and the frequency-dependent advantage of the infection caused by cytoplasmic incompatibility. Although an increase in frequency might result from various combinations of a fecundity advantage of infected females, paternal transmission, or perfect maternal transmission, none of these has been found in nature. A more likely explanation is that a local founder event raises the frequency of infected flies above the critical value of about 15%. Once established locally, the infection can then spread in a wave through the whole species (Turelli and Hoffmann 1991, 1995).

Although the spread of *Wolbachia* in this species involves an increase in the frequency of a deleterious trait and not of deleterious alleles, it does seem to show that a sampling process can counteract natural selection, and that the trait can spread from one population to others (phase III). Moreover, this process has occurred in a populous and cosmopolitan species of Drosophila, a seemingly unlikely candidate for strong genetic drift. However, although an SB-like process may explain the spread of the infection, it does not produce increased adaptation of the species. Because imperfect maternal transmission of bacteria allows the production of a few uninfected flies that engage in incompatible matings, the average fitness of a population after the spread of the infection is slightly *lower* (by about 2%) than that of an uninfected population.

Summary of the Examples.—With the possible exceptions of chromosome evolution, mimicry in *Heliconius*, and molecular covariation, these examples show that drift can sometimes cause gene-frequency change in the face of selection (phase I of the SBT). Two of the cases (cytoplasmic incompatibility in *Drosophila* and coiling in snails) also appear to show that populations evolved past unstable equilibria to new peaks (phase II). Only one of these cases, however (cytoplasmic incompatibility), shows convincing evidence for phase III, and none show that the entire process has led to higher population fitness. Moreover, with the exception of mimicry in *Heliconius*, the examples involve changes in only one or two genes (*no* genes in the case of *Wolbachia* infection), and not the complex "adaptive systems" envisioned by Wright. We conclude that migration rates, population sizes, and genotypic fitnesses can sometimes permit the operation of individual phases of the SBT, but that there is no evidence for the concatenation of all three phases in a single episode of evolution. Because none of these studies was conducted to demonstrate the operation of the SBT, they should not be criticized on that account. What they do reveal is the extraordinary difficulty in showing the operation of the full process, even in the most favorable material. They also pro-

vide some useful examples of how genetic drift has overcome natural selection.

Evidence from Geographic Races and Hybrid Zones

In phase III of the SBT, selection among different peaks leads to improved adaptation. For this interdeme selection to be a strong force, species must be divided into a very large number of local races, and neighboring patches must have established somewhat incompatible gene combinations. This view gains support from observations of outbreeding depression, in which the fitness of offspring from crosses between populations decreases with increasing distance between them. Parker (1992), for example, presents evidence for the coadaptation of alleles found in two "biotypes" of a selfing annual, *Amphicarpaea bractea*, that are spatially segregated over a 100-m transect. Each biotype performs well in the area where the other is prevalent, suggesting that there is not a great deal of local adaptation on this scale. Nevertheless, in the greenhouse, all F_3 families derived from outcrossed F_2 plants produced seeds with an average biomass below that of the midparent. The selective accumulation of coadapted multilocus genotypes is clearly aided by the high degree of selfing in this population, which is at least 99.5%. (See Waser and Price [1994] for a similar analysis in *Delphinium nelsonii* and a review of the literature.) Experiments of this sort, comparing the performance of parents and hybrids in various environments, can assess whether local populations possess coadapted gene complexes that may be favorable across a broad range of environments. This approach is potentially one of the most powerful tests of the SBT.

There is more abundant evidence about divergence over larger scales. Species are frequently subdivided into broad geographic races separated by narrow hybrid zones. One of the most striking observations is that hybrid zones identified by one character (e.g., karyotype or morphology) usually involve many other genetic differences as well (Barton and Hewitt 1985; Hewitt 1988; Avise 1994). This clustering of clines, however, does not support the view that the SB process explains *adaptive* differences among races or species. First, there is much less opportunity for selection among a few geographic races than among very many local demes. Just as Darwin (1859) saw that selection can be effective in the long term only if variation is generated faster than it is eroded, so the SB can operate rapidly only if geographic variation is generated by local peak shifts faster than it is eroded by agglomeration into a few geographic races. Second, selection among geographic races is based on the combined effects of all the differences between them, and does not separate the effects of differences built up by independent peak shifts. Even if it could be shown that certain geographic races have a propensity to spread at the expense of others, the SB would be able to build up complex adaptations only if there were a wide variety of locally differentiated races, each carrying subtly different combinations of genes. Finally, the distribution of gene combinations often reflects expansion from refugia following climate change (Hewitt 1993; Avise 1994). Again, it is hard to see that such expansion could systematically favor the better-adapted race.

Evidence from Artificial Selection Experiments

Shorthorn Cattle.—Presented in three papers in the Journal of Heredity, Wright's (1923a,b, 1925) analysis of selection in British Shorthorn cattle was both a stimulus for developing the SBT and a lifelong source of evidence for its operation (e.g., 1951, 1978, 1982). The essence of his argument was that the improvement of shorthorns (involving such characters as milk production, body configuration, and vigor) was due to a breeding scheme analogous to the shifting balance process, and that the striking success of selection proved the SB process to be rapid and efficient.

Because shorthorn pedigrees were available back to the late eighteenth century, Wright could identify the contributions of individual animals to the genetic composition of herds. Through repeated matings of a few superior bulls, the breeders maintained a fairly high inbreeding coefficient within herds (F between 0.2 and 0.4). Occasionally, bulls were imported from other herds to infuse "fresh blood" into the group. This process was repeated over many generations, producing, according to Wright, a remarkable transformation of the breed, although he presented no data on the response to selection.

Wright claimed that two aspects of this breeding scheme were crucial to its success. First, herds were fairly inbred, which supposedly allowed the breeders to select among diverse "interaction systems" rather than having to deal with the diversity of recombining interaction systems in outbred populations:

> If Bates had not maintained a close relationship between the animals which he mated, the relatively high degree of inbreeding, and close relationship to one animal (Favourite), his material would probably have been too plastic. The simultaneous variation in all characters would have been more than he could have contended with (Wright 1923a, p. 416).

Second, the movement of superior bulls between herds provided new variation for selection. According to Wright, the entire selection scheme mimicked the SB process: restricted population size and inbreeding allowed fixation of new character complexes (phase I and II), and selective export of desirable sires diffused these characters to other herds (phase III):

> Recognition that the two-level process was much more efficient than mere individual selection led to consideration of whether an analogous two-level process might not occur in nature. The first draft of a paper developing this idea was written immediately after the Shorthorn study in 1925 but was not published until 1931 (Wright 1982, pp. 5–6).

When examined closely, Wright's work on Shorthorns actually provides no evidence for the SBT. First, the breeding scheme was used not because it was known to be the most effective selection regime, but probably because it was the only practical method given the restricted number of cattle available to each breeder. The high amount of inbreeding within herds may not have been a conscious or unconscious strategy of selection, but either a means of achieving uni-

formity of phenotype or an inevitable result of strong selection in fairly small populations (Wray and Thompson 1990; Santiago and Caballero 1995). More important, although selection apparently led to a marked improvement of the breed, Wright gave no evidence that this improvement was more rapid than that attainable by mass selection in a single very large herd. (Granted, such an experiment would have been a practical nightmare.) The absence of a mass-selected control also means that we have no evidence that adaptive valleys were crossed during selection.

The Shorthorn data therefore show only that there was a response to selection within herds when the selection was fueled by the movement of favored sires between herds. This result is consistent with nearly every conceivable model of fitness and gene action, including those based on completely additive alleles and the absence of multiple peaks in the adaptive landscape. It is possible that some of the improved traits in shorthorns were based on recessive alleles, and that inbreeding and subdivision uncovered these alleles, allowing a more rapid response to selection. This is an unresolved issue that would be an alternative to the SB explanation *if* we had evidence that selection in shorthorns was more rapid than that achievable by mass selection. As we show below, however, experiments in other species give little evidence for an unusually high response to selection in subdivided populations.

Other Selection Experiments.—Relatively few experimental studies have attempted to simulate the entire SB process. In this section we review these and other relevant experiments from animal breeding and evolutionary genetics.

Wade and Goodnight (1991) described a selection experiment on the beetle *Tribolium castaneum* that was intended to address the efficacy of the SB process. Their experiment compared the rate of advance of average deme productivity (number of adult offspring produced by 20 breeding adults in 60 d) using two sets of treatments. In the experimental treatments, demes contributed to a migrant pool in proportion to their excess in productivity over the average deme. Each deme then received migrants from this pool in proportion to its deficit in productivity compared to the average deme. Three different selection regimes were used: selection every generation, every second generation, and every third generation. Each regime was unreplicated. In the control treatments, levels of interdeme migration were identical to those of the corresponding experimental lines, but the emigration and immigration rates were independent of group productivity. All three experimental lines showed a substantial response to interdeme selection relative to the controls. Moreover, the greatest response (in terms of realized heritability) occurred in the line selected every other generation.

Although Wade and Goodnight (1991, p. 1015) and Crow (1991) interpret these data as experimental confirmation of the SBT, there is some ambiguity in the results. Wade and Goodnight's data show that if demes contribute to the next generation in proportion to their productivity, mean productivity rises. Interdeme selection, however, is effective under nearly any mode of gene action, including the simple case of selection for an additive trait governed by a single locus. The scheme is a variant of the family-selection method used to select for quantitative traits of low heritability. In that method, only the families with the highest mean contribute to the next generation, whereas in the Wade and Goodnight design, families (demes) contribute in proportion to their phenotypic value. Such a design could successfully alter nearly any character, including bristle number in *Drosophila*, a classical subject of mass-selection experiments.

Wright's central claim for the SBT was that the interaction between epistasis, population subdivision, and interdemic selection will often produce a response greater than that achievable by mass selection. This crucial question, however, was not addressed by Wade and Goodnight's study: because they did not practice mass selection (the control lines were unselected), they could not compare the responses to mass and interdeme selection. It might be possible to compare these selection regimes by letting each deme in the control regime contribute to a single random-mating pool in proportion its productivity, and then arbitrarily dividing the mated adults into groups to assay productivity in the next generation. In the absence of such a control, however, the interpretation of the experimental results is necessarily ambiguous, and in this sense is similar to Wright's claim that selection response in shorthorn cattle, which also lacked a mass-selected control, supports the SBT.

A second claim by Wade and Goodnight (1991) is that the larger response in their selection-every-other-generation treatment indicates a nonadditive basis for the trait, so that the gains they report could not have resulted from additive genetic differences accessible to ordinary mass selection. Unfortunately, their experiment used only one line in each selection regime, so it is unclear whether the variation in response is characteristic of these regimes or is due, say, to a sampling artifact in the single line selected every other generation. Even if their result were seen in multiple lines, however, it would not prove the existence of epistasis: it is also the case that in models of group selection *acting on a single locus*, the largest response can occur with intermittent selection (Wilson and Colwell 1981).

It is often claimed that nonadditive variance for fitness is common, but much of this evidence comes from inbred populations or homozygous lines. Experimental studies of outbred populations, however, show little evidence for significant epistatic variation in components of fitness (Tachida and Cockerham 1988; Charlesworth and Hughes, in press).

There have been other experiments testing whether the response to selection is enhanced in subdivided populations. Barker (1979), Cohan and Hoffmann (1989), Cohan et al. (1989), Hill (1989), and Hill and Caballero (1992) reviewed seven such studies in *Drosophila* and *Tribolium*, concluding that only the work of Katz and Young (1975) on body size in *D. melanogaster* supported the SBT. As Hill (1989) noted, one possible explanation is that this character generally exhibits more nonadditive variance than do the other traits, which may explain its enhanced response to selection with subdivision. However, no explanation is really needed for Katz and Young's result, which provides little evidence that subdivided populations respond more rapidly to selection. As in the study of Wade and Goodnight (1991), each of Katz and Young's three selection regimes (mass selection, selection every generation in a subdivided population, and selection every second generation in a subdivided population) was

unreplicated. Moreover, although the two "subdivided" lines achieved the highest means by generation 18, the variation among all three lines was quite small, well within the amount of divergence routinely seen among replicate responses to a single mode of selection. Hence, there is no evidence for the "treatment effects" needed to establish that selection is more efficient in subdivided populations.

Hill (1989, 1996) and Hill and Caballero (1992) also reviewed relevant work derived from animal breeding. Overall, these provide no strong support for the idea that selection response is enhanced by breeding schemes designed to exploit epistatic variance via inbreeding and population subdivision. Hill (1989, p. 194) concluded that had Wright known of this work, "I rather doubt he would have arrived at and championed the shifting-balance theory." It would clearly be desirable, however, to conduct additional selection experiments in subdivided and mass populations, making sure that each selection regime is replicated so that any treatment effects can be discerned.

Cohan et al. (1989) studied the importance of epistasis in producing genetic divergence between independently selected populations of *Drosophila*. Their assay involved comparing the means of F_1 and F_2 populations to those of the selected parents. They found no significant evidence for epistasis, in contrast to some similar experiments on other traits (reviewed in Cohan et al. 1989). The explanation of this discrepancy may again be that genetic variance in the selected character (the ability of flies to resist knockdown by ethanol fumes) is predominantly additive. Given our profound ignorance of the genetics of adaptation (Orr and Coyne 1992), we have no idea whether the evolution of "typical" adaptations in nature involves epistasis of the type postulated by the SBT.

Conclusions

Our reviews of both the theoretical issues and of the empirical literature do not provide compelling, or even moderate, support for the SBT.

Theory shows that the SB can sometimes be an efficient mechanism of adaptation, but only under restrictive conditions. The massive effort expended on estimating population structure from genetic data (Lewontin et al. 1981; Slatkin 1987) has shown that in most cases, random drift would be strong enough to cause shifts only between peaks separated by shallow valleys. Such tests are not decisive, however, because drift may be strong enough at some times and places to cause more drastic peak shifts, and because there are many other mechanisms of divergence that could drive populations to different adaptive peaks. The key obstacles to seeing the SB as a common adaptive process are that most adaptations may be driven by natural selection alone without requiring that populations cross adaptive valleys; adaptations whose fixation requires drift are susceptible to being trapped by barriers to dispersal, unlike adaptations fixed by directional selection; and adaptive changes that occur independently are hard to bring together by recombination. We require a better empirical understanding of whether subpopulations actually occupy different adaptive peaks that are not determined by local selection pressures, how new adaptive gene combinations might spread (despite fitness valleys) from the marginal populations where they may be most likely to arise, and whether better adapted genetic combinations spread preferentially.

The empirical evidence is equally nebulous. Although Wright (1977, 1978) interpreted most empirical observations in light of his theory, nearly all of these observations are, like that of selection response in Shorthorn cattle, equally well explained by Fisherian mass selection. Compelling empirical support for the SBT can come only from cases in which Fisherian selection is incapable of explaining the evolution of an adaptation. We presented six cases that indicate the action of at least one component of the SBT. However, none of these examples demonstrates all three phases of the SBT, at least as Wright conceived it. The case of *Wolbachia* infection in *D. simulans* is perhaps the closest analog to the entire process, for it demonstrates a "peak shift" that is spreading through a species (probably by a "wave of advance" rather than Wrightian interdeme selection). However, this example shows the spread not of advantageous alleles, but of a maternally transmitted infection. The inability of any case to completely exemplify all phases of the SBT does not necessarily rule out the SB process in nature, but does show the extreme difficulty of garnering strong evidence. With such a situation, it will be impossible to support Wright's claim that the SBT explains a large fraction of evolutionary change in nature.

Our examples do offer weaker support for the SBT by showing that genetic drift has affected the evolution of some adaptations, and can at least occasionally overcome selection, forcing populations off of adaptive peaks. Regardless of whether all three phases have operated together, such observations at least allow us to expand our view of evolution beyond that of Fisherian selection in large populations, incorporating factors such as genetic drift, population structure, and epistasis. It must be remembered, however, that much of the interest in such phenomena was based on their importance as constituents of Wright's theory. Are there any adaptations in nature whose evolution involved intermediate steps of lowered fitness? This question is hard to answer because adaptations are usually fixed within a species, burying their evolutionary path in the unrecoverable past. Finding support for the SBT then depends on being able to either reconstruct the sequence of genetic and adaptive change involved in a given adaptation—almost always an insuperable task—or having the rare opportunity to witness peak shifts occurring in nature (as in *Wolbachia* infection or the breakdown of tristyly).

Evolutionary cytogenetics may be one fruitful area for observing the complete process, for in some cases of chromosome evolution it may be possible to reconstruct the pathway of evolution and determine whether it involves crossing an adaptive valley. Although interdeme selection is not required for underdominant chromosome arrangements to spread (Barton 1979), the SB process would be supported if one can demonstrate the fixation of a fitness-enhancing chromosome rearrangement subject to strong underdominance. Similarly, studies of outbreeding depression, when combined with control experiments to determine whether the phenomenon is caused only by local adaptation, can reveal whether populations in nature are on separate fitness peaks. Such ex-

periments cannot, however, demonstrate whether the transitions between those peaks occurred by drift.

The SBT is sometimes used, as in the case of Müllerian mimicry in *Heliconius*, to explain the evolution of traits whose intermediate steps seemed to involve a loss of fitness. Usually, however, there are alternative explanations involving selection alone. For example, the intermediate stages in the evolution of a Batesian mimic were once thought to be maladaptive, for edible "half-mimics" are conspicuous to the predator but not sufficiently similar to avoid being eaten. This potential adaptive valley, however, was crossed not by drift but by the fixation of major advantageous mutations that leap the valley by causing strong resemblance to the model (Turner 1977; Orr and Coyne 1992). As Whitlock et al. (1995) emphasize, supporting the SBT requires that one rule out the possibility that a supposed "peak shift" actually involved unseen ridges that could be scaled by directional selection.

A test of whether the SB as a whole has the *potential* to make a significant contribution to adaptation could come from careful comparisons between artificial populations maintained under mass selection versus conditions conducive to Wright's process. The few such tests to date have given no strong support to the theory (see Hill 1989 and above), but the characters studied may not have been amenable to evolution by the SB because their genetic variance was primarily additive. If such experiments show that population subdivision significantly enhances the rate of evolution, this can be construed as supporting the SBT only if the response is due to shifts between adaptive peaks in the subdivided population and not to the more efficient use of favorable recessive and nonadditive variation under partial inbreeding. These outcomes could be discriminated by seeing whether hybrids between demes have reduced fitness.

We do not doubt that the SB process might sometimes operate in nature. Given the multifarious nature of evolution, almost every conceivable scenario must occasionally occur. Many species are subdivided, genetic drift must sometimes oppose natural selection, and some evolutionary transitions probably involved peak shifts. Nevertheless, we have found no compelling evidence that Wright's SBT accounts for the evolution of a single adaptation, much less a significant proportion of adaptations, in nature. Until such evidence is at hand, we favor the view that adaptations are usually produced by Fisherian mass selection, a process that is not only more parsimonious than the SBT, but has also been shown to occur widely (Endler 1986).

Acknowledgments

We thank the following people for discussion and comments on the manuscript: S. Barrett, J. Bull, B. Charlesworth, D. Charlesworth, P. DeVries, S. Gavrilets, J. H. Gillespie, R. K. Grosberg, W. G. Hill, A. A. Hoffmann, M. Kirkpatrick, C. H. Langley, R. C. Lewontin, J. B. Mallet, M. Noor, L. Nunney, H. A. Orr, T. Prout, M. Slatkin, J. Spofford, W. Stephan, J. B. Walsh, P. Ward, K. Weber, J. Willis, and M. Zwick. We are especially grateful to D. J. Futuyma and D. Schemske for their exhaustive criticism of the manuscript. Needless to say, not all of these reviewers agree with our ideas. This work was supported by National Institutes of Health grant GM 50355 to JAC, National Science Foundation grant DEB 9527808 to MT, and grants from the Darwin Trust of Edinburgh and the Biotechnology and Biological Sciences Research Council (GRJ/76057, GR/H/09928) to NHB.

Literature Cited

Akin, E. 1979. The geometry of population genetics. Lecture notes in biomathematics 31. Springer-Verlag, Berlin.

Avise, J. C. 1994. Molecular markers, natural history and evolution. Chapman and Hall, New York.

Baker, R. J., and J. W. Bickham. 1986. Speciation by monobrachial centric fusions. Proc. Nat. Acad. Sci. USA 83:8245–8248.

Barker, J. S. F. 1979. Inter-locus interactions: A review of experimental evidence. Theor. Pop. Biol. 16:323–346.

Barrett, S. C. H. 1993. The evolutionary biology of tristyly. Oxf. Surv. Evol. Biol. 9:283–326.

Barrett, S. C. H., M. T. Morgan, and B. C. Husband. 1989. The dissolution of a complex genetic polymorphism: The evolution of self-fertilization in tristylous *Eichhornia paniculata* (Pontederiaceae). Evolution 43:1398–1416.

Barrett, S. H. 1992. Evolution and function of heterostyly. Springer-Verlag, Berlin.

Barton, N. H. 1979. The dynamics of hybrid zones. Heredity 43: 341–359.

———. 1980. The hybrid sink effect. Heredity 44:277–278.

———. 1986. The maintenance of polygenic variation through a balance between mutation and stabilising selection. Genet. Res. 47:209–216.

———. 1989a. Founder effect speciation. Pp. 229–256 *in* D. Otte and J. A. Endler, eds. Speciation and its consequences. Sinauer, Sunderland, MA.

———. 1989b. The divergence of a polygenic system under stabilising selection, mutation and drift. Genet. Res. 54:59–77.

———. 1992. On the spread of new gene combinations in the third phase of Wright's shifting balance. Evolution 46:551–557.

Barton, N. H., and B. Charlesworth. 1984. Genetic revolutions, founder effects, and speciation. Ann. Rev. Ecol. Syst. 15:133–164.

Barton, N. H., and A. Clark. 1990. Population structure. Pp. 115–174 *in* K. Wöhrmann and S. Jain, eds. Population genetics and evolution. Springer-Verlag, Berlin.

Barton, N. H., and G. M. Hewitt. 1985. Analysis of hybrid zones. Ann. Rev. Ecol. Syst. 16:113–148.

Barton, N. H., and S. Rouhani. 1987. The frequency of shifts between alternative equilibria. J. Theor. Biol. 125:397–418.

———. 1991. The probability of fixation of a new karyotype in a continuous population. Evolution 45:499–517.

———. 1993. Adaptation and the "shifting balance." Genet. Res. 61:57–74.

Bazykin, A. D. 1969. Hypothetical mechanism of speciation. Evolution 23:685–687.

Bergman, A., D. B. Goldstein, K. Holsinger, and M. W. Feldman. 1995. Population structure, fitness surfaces, and linkage in the shifting balance process. Genet. Res. 66:85–92.

Berry, R. J. 1986. Genetics of insular populations of mammals, with particular reference to differentiation and founder effects in British small mammals. Biol. J. Linn. Soc. 28:205–230.

Bickham, J. W., and R. J. Baker. 1980. Reassessment of the nature of chromosomal evolution in *Mus musculus*. Syst. Zool. 29:159–162.

Boerlijst, M. C., and P. Hogeweg. 1995. Attractor and spatial patterns in hypercycles with negative interactions. J. Theor. Biol. 176:199–210.

Brodie, E. D. 1991. Evolutionary response of predators to dangerous prey: Reduction of toxicity of newts and resistance of garter snakes in island populations. Evolution 45:221–225.

Brown, K. S., P. M. Sheppard, and J. R. G. Turner. 1974. Quaternary refugia in tropical America: Evidence from race formation in *Heliconius* butterflies. Proc. R. Soc. Lond. B Biol. Sci. 187:369–378.

BRYANT, E. H., AND L. M. MEFFERT. 1988. Effect of an experimental bottleneck on morphological integration in the housefly. Evolution 42:698–707.

BRYANT, E. H., S. A. MCCOMMAS, AND L. M. COMBS. 1986. The effect of an experimental bottleneck upon additive genetic variation in the housefly. Genetics 114:1191–1211.

CABALLERO, A., P. D. KEIGHTLEY, AND W.G. HILL. 1991. Strategies for increasing fixation probabilities of recessive mutations. Genet. Res. 58:129–138.

CACCONE, A., E. N. MORIYAMA, J. M. GLEASON, L. NIGRO, AND J. R. POWELL. 1996. A molecular phylogeny for the *Drosophila melanogaster* subgroup. Mol. Biol. Evol. 13:1224–1232.

CARROLL, R. L. 1988. Vertebrate paleontology and evolution. Freeman, New York.

CHARLESWORTH, B., AND K. A. HUGHES. In press. The quantitative genetics of life history traits. *In* R. Singh and C. Krimbas, eds. From molecules to morphology. Oxford Univ. Press, Oxford.

CHARLESWORTH, B., AND S. ROUHANI. 1988. The probability of peak shifts in a founder population. II. An additive polygenic trait. Evolution 42:1129–1145.

CHARLESWORTH, B., AND D. B. SMITH. 1982. A computer model of speciation by founder effects. Genet. Res. 39:227–236.

CHARLESWORTH, D., AND B. CHARLESWORTH. 1979. A model for the evolution of distyly. Am. Nat. 114:467–498.

———. 1980. Sex differences in fitness and selection for centric fusions between sex chromosomes and autosomes. Genet. Res. 35:205–214.

COHAN, F. M., AND A. A. HOFFMANN. 1989. Uniform selection as a diversifying force in evolution: Evidence from *Drosophila*. Am. Nat. 134:613–637.

COHAN, F. M., A. A. HOFFMANN, AND T. W. GAYLEY. 1989. A test of the role of epistasis in divergence under uniform selection. Evolution 43:766–774.

COYNE, J. A. 1984. Correlation between heterozygosity and rate of chromosome evolution in animals. Am. Nat. 123:725–729.

COYNE, J. A., S. AULARD, AND A. BERRY. 1991. Lack of underdominance in a naturally occurring pericentric inversion in *Drosophila melanogaster* and its implications for chromosome evolution. Genetics 129:791–802.

COYNE, J. A., W. MEYERS, A. P. CRITTENDEN, AND P. SNIEGOWSKI. 1993. The fertility effects of pericentric inversions in *Drosophila melanogaster*. Genetics 134:487–496.

CROMPTON, A. W. 1963. On the lower jaw of *Diarthrognathus* and the origin of the mammalian lower jaw. Proc. Zool. Soc. Lond. 140:697–753.

CROW, J. F. 1991. Was Wright right? Science 253:973.

CROW, J. F., W. R. ENGELS, AND C. DENNISTON. 1990. Phase three of Wright's shifting-balance theory. Evolution 44:233–247.

CRUMPACKER, D. W. 1967. Genetic loads in maize (*Zea mays* L.) and other cross-fertilized plants and animals. Evol. Biol. 1:306–324.

DARWIN, C. 1859. On the origin of species by means of natural selection. John Murray. London.

DOBZHANSKY, T. 1937. Genetics and the origin of species. Columbia Univ. Press, New York.

EBERHARD, W. G. 1985. Sexual selection and animal genitalia. Harvard Univ. Press, Cambridge, MA.

EIGEN, M. 1986. The physics of molecular evolution. Chemica Scripta 26:13–26.

ELDREDGE, N. 1995. Reinventing Darwin. Wiley, New York.

ENDLER, J. A. 1986. Natural selection in the wild. Princeton Univ. Press, Princeton, NJ.

EPLING, C., AND T. DOBZHANSKY. 1942. Genetics of natural populations. VI. Microgeographical races in *Linanthus parryae*. Genetics 27:317–332.

EPLING, C., H. LEWIS, AND F. M. BALL. 1960. The breeding group and seed storage: A study in population dynamics. Evolution 14:238–255.

EWENS, W. J. 1979. Mathematical population genetics. Springer-Verlag, Berlin.

FALCONER, D.S. 1989. Introduction to quantitative genetics. 3d ed. Longman, London.

FELSENSTEIN, J. 1971. On the biological significance of the cost of gene substitution. Am. Nat. 105:1–11.

———. 1979. Excursions along the interface between disruptive and stabilizing selection. Genetics 93:773–795.

FISHER, R. A. 1930. The genetical theory of natural selection. Oxford Univ. Press, Oxford.

FONTANA, W., D. A. M. KONINGS, P. F. STADLER, AND P. SCHUSTER. 1988. Physical aspects of evolutionary optimization and adaptation. Phys. Rev. A 40:3301–3321.

FONTANA, W., T. GREISMACHER, W. SCHNABL, P. F. STADLER, AND P. SCHUSTER. 1991. Statistics of landscapes based on free energies, replication and degradation rate constants of RNA secondary structures. Monatshefte Chemie 122:795–819.

FONTANA, W., W. SCHNABL, E. G. BORNBERG-BAUER, T. GREISMACHER, I. L. HOFACKER, M. TACKER, P. TARAZONA, E. D. WEINBERGER, AND P. SCHUSTER. 1993a. RNA folding and combinatory landscapes. Phys. Rev. E 47:2083–2099.

FONTANA, W., W. SCHNABL, AND P. SCHUSTER. 1993b. Statistics of RNA secondary structures. Biopolymers 33:1389–1404.

FORREST, S. 1993. Genetic algorithms—principles of natural selection applied to computation. Science 261:872–878.

GAVRILETS, S. 1996. On phase three of the shifting-balance theory. Evolution 50:1034–1041.

GAVRILETS, S., AND A. HASTINGS. 1996. Founder effect speciation: A theoretical reassessment. Evolution 147:466–491.

GIBBS, H. L., AND P. R. GRANT. 1987. Oscillating selection in Darwin's finches. Nature 327:511–513.

GILLESPIE, J. H. 1983. Some properties of finite populations experiencing strong selection and weak mutation. Am. Nat. 121: 691–708.

———. 1984. Molecular evolution over the mutational landscape. Evolution 38:1116–1129.

———. 1991. The causes of molecular evolution. Oxford Univ. Press, Oxford.

GINGERICH, P. D. 1983. Rates of evolution: Effects of time and temporal scaling. Science 222:159–161.

GITTERBERGER, E. 1988. Sympatric speciation in snails: A largely neglected model. Evolution 42:826–828.

GOLDSTEIN, D. B., AND K. E. HOLSINGER. 1992. Maintenance of polygenic variation in spatially structured populations: Roles for local mating and genetic redundancy. Evolution 46:412–429.

GOODNIGHT, C. J. 1988. Epistasis and the effect of founder events on the additive genetic variance. Evolution 42:441–454.

———. 1995. Epistasis and the increase in additive genetic variance: Implications for Phase I of Wright's shifting balance process. Evolution 49:502–511.

GOULD, S. J. 1980. Is a new and general theory of evolution emerging? Paleobiology 6:119–130.

GOULD, S. J., N. D. YOUNG, AND B. KASSON. 1985. The consequences of being different: Sinistral coiling in *Cerion*. Evolution 39:364–1379.

GRANT, P. R. 1986. The evolution and ecology of Darwin's finches. Princeton Univ. Press, Princeton, NJ.

HALDANE, J. B. S. 1932. The causes of evolution. Longmans, Green and Co., New York.

———. 1959. Natural selection. Pp. 101–149 *in* P. R. Bell, ed. Darwin's biological work: Some aspects reconsidered. Wiley, New York.

HASTINGS, A. 1989. Deterministic multilocus population genetics: An overview. Lect. Math. Life Sci. 27:318–336.

HEDRICK, P. W. 1981. The establishment of chromosomal variants. Evolution 35:322–332.

HEDRICK, P. W., AND D. A. LEVIN. 1984. Kin-founding and the fixation of chromosomal variants. Am. Nat. 124:789–797.

HEUCH, I., AND R. T. LIE. 1985. Genotype frequencies associated with incompatibility systems in tristylous plants. Theor. Popl. Biol. 27:318–336.

HEWITT, G. M. 1988. Hybrid zones: Natural laboratories for evolutionary studies. Trends Ecol. Evol. 3:158–167.

———. 1993. After the ice: *Parallelus* meets *erythropus* in the Pyrenees. Pp. 40–164 *in* R. G. Harrison, ed. Hybrid zones and the evolutionary process. Oxford Univ. Press, Oxford.

HILL, W. G. 1982. Predictions of response to artificial selection from new mutations. Genet. Res. 40:255–278.
———. 1989. Sewall Wright and quantitative genetics. Genome 31:190–195.
———. 1996. Sewall Wright's "Systems of Mating." Genetics 143:1499–1506.
HILL, W. G., AND A. CABALLERO. 1992. Artificial selection experiments. Ann. Rev. Ecol. Syst. 23:287–310.
HOFFMANN, A. A., M. TURELLI, AND G. M. SIMMONS. 1986. Unidirectional incompatibility between populations of *Drosophila simulans*. Evolution 40:692–701.
HOLLAND, J. 1975. Adaptation in natural and artificial systems. Univ. of Michigan Press, Ann Arbor.
HUSBAND, B. C., AND S. C. H. BARRETT. 1992a. Genetic drift and the maintenance of the style length polymorphism in tristylous populations of *Eichhornia paniculata* (Pontederiaceae). Heredity 69:440–449.
———. 1992b. Effective population size and genetic drift in tristylous *Eichhornia paniculata* (Pontederiaceae). Evolution 46: 1875–1890.
JOHN, B. 1981. Chromosome change and evolutionary change: A critique. Pp. 23–51 *in* W. R. Atchley and D. S. Woodruff, eds. Evolution and speciation. Essays in honor of M. J. D. White. Cambridge Univ. Press, Cambridge.
JOHN, B., AND L. G. MIKLOS. 1988. The eukaryotic genome in development and evolution. Allen and Unwin, London.
JOHNSON, M. S. 1982. Polymorphism for direction of coil in *Partula suturalis*: Behavioural isolation and positive frequency dependent selection. Heredity 49:145–151.
JOHNSON, P. A., R. E. LENSKI, AND F. C. HOPPENSTEADT. 1995. Theoretical analysis of divergence in mean fitness between initially identical populations. Proc. R. Soc. Lond. B Biol. Sci. 259: 131–136.
KATZ, A. J., AND S. S. Y. YOUNG. 1975. Selection for high adult body size in *Drosophila* populations with different structures. Genetics 81:163–175.
KAUFFMAN, S., AND S. A. LEVIN. 1987. Towards a general theory of adaptive walks on a rugged landscape. J. Theor. Biol. 128: 11–46.
KEY, K. H. L. 1968. The concept of stasipatric speciation. Syst. Zool. 17:14–22.
KIMURA, M. 1983. The neutral theory of molecular evolution. Cambridge Univ. Press, Cambridge.
———. 1985a. The role of compensatory neutral mutations in molecular evolution. J. Genet. 64:7–19.
———. 1985b. Diffusion models in population genetics with special reference to fixation time of molecular mutants under mutational pressure. Pp. 19–39 *in* T. Ohta and K. Aoki, eds. Population genetics and molecular evolution. Japan Scientific Societies Press, Tokyo.
KING, M. 1993. Species evolution. Cambridge Univ. Press, Cambridge.
KIRBY, D. A., S. V. MUSE, AND W. STEPHAN. 1995. Maintenance of pre-mRNA secondary structure by epistatic selection. Proc. Nat. Acad. Sci. USA 92:9047–9051.
KIRKPATRICK, M. 1982. Quantum evolution and punctuated equilibrium in continuous genetic characters. Am. Nat. 119:833–848.
KIRKPATRICK, S., C. D. GELATT, AND M. P. VECCHI. 1983. Optimization by simulated annealing. Science 220:671–680.
KOHN, J. R., AND S. C. H. BARRETT. 1994. Pollen discounting and the spreading of a selfing variant in tristylous *Eichhornia paniculata*: Evidence from experimental populations. Evolution 48: 1576–1594.
KONDRASHOV, A. S. 1992. The third phase of Wright's shifting balance theory: A simple analysis of the extreme case. Evolution 48:1972–1975.
LANDE, R. 1976. Natural selection and random genetic drift in phenotypic evolution. Evolution 30:314–334.
———. 1979. Effective deme sizes during long-term evolution estimated from rates of chromosomal rearrangement. Evolution 33:234–251.
———. 1984. The expected fixation rate of chromosomal inversions. Evolution 38:743–752.
———. 1985a. Expected time for random genetic drift of a population between stable phenotypic states. Proc. Nat. Acad. Sci. USA 82:7641–7645.
———. 1985b. The fixation of chromosomal rearrangements in a subdivided population with local extinction and recolonisation. Heredity 54:323–332.
———. 1989. Fisherian and Wrightian theories of speciation. Genome 31:221–227.
LATTER, B. D. H. 1960. Natural selection for an intermediate optimum. Aust. J. Biol. Sci. 13:30–35.
LEIGH, E. G. 1983. When does the good of the group override the advantage of the individual? Proc. Nat. Acad. Sci. USA 74:4542–4546.
LEWONTIN, R. C. 1974. The genetic basis of evolutionary change. Columbia Univ. Press, New York.
LEWONTIN, R. C., J. A. MOORE, W. B. PROVINE, AND B. WALLACE, EDS. 1981. Dobzhansky's genetics of natural populations I–XLIII. Columbia Univ. Press, New York.
LLOYD, D. G., AND C. J. WEBB. 1992. The selection of heterostyly. Pp. 179–207 *in* S. H. Barrett, ed. Evolution and function of heterostyly. Springer-Verlag, Berlin.
LUDWIG, D. 1975. Persistence of dynamical systems under random perturbation. SIAM Review 17:605–640.
LUSH, J. L. 1937. Animal breeding plans. Iowa State Univ. Press, Ames.
LYNCH, M., AND R. LANDE. 1993. Evolution and extinction in response to environmental change. Pp. 234–250 *in* P. M. Kareiva, J. G. Kingsolver, and R. B. Huey, eds. Biotic interactions and global change. Sinauer, Sunderland, MA.
MACKAY, T.F.C., J. D. FRY, R.F. LYMAN, AND S. V. NUZHDIN. 1994. Polygenic mutation in *Drosophila melanogaster*: Estimates from response to selection of inbred strains. Genetics 136:937–951.
MALLET, J. L. B. 1986. Hybrid zones of *Heliconius* butterflies in Panama, and the stability and movement of warning colour clines. Heredity 56:191–202.
———. 1993. Speciation, raciation, and color pattern evolution in *Heliconius* butterflies: Evidence from hybrid zones. Pp. 226–260 *in* R. G. Harrison, ed., Hybrid zones and the evolutionary process. Oxford Univ. Press, New York.
MALLET, J., AND N. H. BARTON. 1989. Strong natural selection in a warning color hybrid zone. Evolution 43:421–421.
MALLET, J., N. H. BARTON, G. LAMAS, J. SANTISTEBAN, M. MUEDAS, AND H. EELY. 1990. Estimates of selection and gene flow from measures of cline width and linkage disequilibrium in *Heliconius* hybrid zones. Genetics 124:921–936.
MANGEL, M. 1994. Barrier transitions driven by fluctuations, with applications to ecology and evolution. Theor. Popl. Biol. 45:16–40.
MANI, G. S., AND B. C. CLARKE. 1990. Mutational order: A major stochastic process in evolution. Proc. R. Soc. Lond. B Biol. Sci. 240:29–37.
MAYNARD SMITH, J. 1966. Sympatric speciation. Am. Nat. 100: 637–650.
———. 1970. Natural selection and the concept of a protein space. Nature 225:563–564.
———. 1976. Group selection. Q. Rev. Biol. 51:277–283.
———. 1989. Evolutionary genetics. Oxford Univ. Press, Oxford.
MAYR, E. 1942. Systematics and the origin of species. Columbia Univ. Press, New York.
———. 1959. Where are we? Cold Spring Harbor Symp. Quant. Biol. 24:1–14.
———. 1963. Animal species and evolution. Harvard Univ. Press, Cambridge, MA.
MCCARTHY, E. M., M. A. ASMUSSEN, AND W. W. ANDERSON. 1995. A theoretical assessment of recombinational speciation. Heredity 74:502–509.
MCCAULEY, D., AND W. F. EANES. 1987. Hierarchical population structure analysis of the milkweed beetle, *Tetraopes tetraophthalmus* (Forster). Heredity 58:193–202.
MCKUSICK, V. A., ED. 1978. Medical genetic studies of the Amish: Selected papers. Johns Hopkins Univ. Press, Baltimore, MD.
MENKEN, S. B. J. 1987. Is the extremely low heterozygosity level

in *Yponomeuta rorellus* caused by bottlenecks? Evolution 41: 630–637.

MERREL, D. J. 1994. The adaptive seascape: The mechanism of evolution. Univ. of Minnesota Press, Minneapolis.

MICHALAKIS, Y., AND M. SLATKIN. 1996. Interaction of selection and recombination in the fixation of negative-epistatic genes. Genet. Res. 67:257–270.

MOORE, F. B. G., AND F. J. TONSOR. 1994. A simulation of Wright's shifting-balance process: Migration and the three phases. Evolution 48:69–80.

MULLER, H. J. 1940. Bearing of the *Drosophila* work on systematics. Pp. 185–268 *in* J. S. Huxley, ed. The new systematics. Clarendon Press, Oxford.

NACHMAN, M. W., AND J. B. SEARLE. 1995. Why is the mouse karyotype so variable? Trends Ecol. Evol. 10:397–402.

NAGYLAKI, T. 1976. Selection in one- and two-locus systems. Springer Verlag, New York.

ORR, H. A. 1991. Is single-gene speciation possible? Evolution 45: 764–769.

———. 1995. The population genetics of hybrid incompatibilities. Genetics 139:1805–1813.

ORR, H. A., AND J. A. COYNE. 1992. The genetics of adaptation revisited. Am. Nat. 140:725–742.

PARKER, M. A. 1992. Outbreeding depression in a selfing annual. Evolution 46:837–841.

PATON, R. 1994. Computing with biological metaphors. Chapman and Hall, London.

PETRY, D. 1982. The pattern of phyletic speciation. Paleobiology 8:56–66.

PHILLIPS, P. C. 1993. Peak shifts and polymorphism during phase three of Wright's shifting balance process. Evolution 47:1733–1743.

PIALEK, J., AND N. H. BARTON. In press. Recombination between geographic races: Two-locus models. Genetics.

PRICE, T., M. TURELLI, AND M. SLATKIN. 1993. Peak shifts produced by correlated response to selection. Evolution 47:280–290.

PROUT, T. 1980. Some relationships between density independent selection and density dependent population growth. Evol. Biol. 13:1–68.

PROVINE, W. 1986. Sewall Wright and evolutionary biology. Univ. of Chicago Press, Chicago.

———. 1989. Founder effects and genetic revolutions in microevolution and speciation: An historical perspectives. Pp. 43–76 *in* L. V. Giddings, K. Y. Kaneshiro, and W. W. Anderson, eds. Genetics, speciation, and the founder principle. Oxford Univ. Press, New York.

REED, K. M., I. F. GREENBAUM, AND J. W. SITES JR. 1995. Cytogenetic analysis of chromosomal intermediates from a hybrid zone between two chromosome races of the *Sceloporus grammicus* complex (Sauria, Phrynosomatidae). Evolution 49:37–47.

RICHARDS, O. W., AND G. C. ROBSON. 1926. The species problem and evolution. Nature 117:345–347,382–384

RIDLEY, M. 1993. Evolution. Blackwell Scientific Publications, Oxford.

RISCH, R., D. DE LEON, L. OELIUS, P. KRAMER, L. ALMASY, B. SINGER, S. FAHN, X. BREAKFIELD, AND S. BRESSMAN. 1995. Genetic analysis of idiopathic torsion dystonia in Ashkenazi Jews and their recent descent from a small founder population. Nat. Genet. 9:152–159.

ROBSON, G. C. 1928. The species problem. Oliver and Boyd, London.

ROBSON, G. C., AND O. W. RICHARDS. 1936. The variation of animals in nature. Longmans, Green and Co., London.

ROUGHGARDEN, J. 1979. Theory of population genetics and evolutionary ecology: An introduction. MacMillan, New York.

ROUHANI, S., AND N. H. BARTON. 1987a. The probability of peak shifts in a founder population. J. Theor. Biol. 126:51–62.

———. 1987b. Speciation and the "shifting balance" in a continuous population. Theor. Pop. Biol. 31:465–492.

———. 1993. Group selection and the "shifting balance." Genet. Res. 61:127–136.

RUIZ, A., AND J. ALVEROLA. 1983. Lack of evidence of embryonic mortality in the progeny of *Drosophila buzzatii* females heterozygous for included inversions. Drosophila Inf. Serv. 59:1065–1067.

SANTIAGO, E., AND A. CABALLERO. 1995. Effective size of populations under selection. Genetics 139:1013–1030.

SCHLEFER, E., M. ROMANO, S. I. GUTTMAN, AND S. RUTH. 1986. Effects of twenty years of hybridization in a disturbed habitat on *Hyla cinerea* and *Hyla gratiosa*. J. Herp. 20:210–221.

SCHLUTER, D. 1996. Ecological causes of adaptive radiation. Am. Nat. Suppl. 148:S40–S64.

SIMMONS, M. J., AND J. F. CROW. 1977. Mutations affecting fitness in *Drosophila* populations. Ann. Rev. Genetics 11:49–78.

SIMPSON, G. G. 1944. Tempo and mode in evolution. Columbia Univ. Press, New York.

———. 1953. The major features of evolution. Columbia Univ. Press, New York.

SITES, J. W., JR., AND C. MORITZ. 1987. Chromosomal evolution and speciation revisited. Syst. Zool. 36:153–174.

SLATKIN, M. 1975. Gene flow and selection in a two-locus system. Genetics 81:209–222.

———. 1981. Estimating levels of gene flow in natural populations. Genetics 99:323–335.

———. 1987. Gene flow and the geographic structure of natural populations. Science 236:787–792.

———. 1989. Population structure and evolutionary progress. Genome 31:196–202.

SMITH, T. B. 1993. Disruptive selection and the genetic basis of bill size polymorphism in the African finch *Pyrenestes*. Nature 363:618–620.

SPIRITO, F. 1992. The exact values of the probability of fixation of underdominant chromosomal rearrangements. Theor. Pop. Biol. 41:111–120.

STEPHAN, W. S. 1996. The rate of compensatory evolution. Genetics 143:419–426.

STEPHAN, W. S., AND D. A. KIRBY. 1993. RNA folding in *Drosophila* shows a distance effect for compensatory mutations. Genetics 135:97–103.

SUMIDA, B. H., AND W. D. HAMILTON. 1994. Both Wrightian and "parasite" peak shifts enhance genetic algorithm performance in the travelling salesman problem. Pp. 254–279 *in* R. Paton, ed. Computing with biological metaphors. Chapman Hall, London.

SUMIDA, B. H., A. I. HOUSTON, J. M. MCNAMARA, AND W. D. HAMILTON. 1989. Genetic algorithms and evolution. J. Theor. Biol. 147:59–84.

TACHIDA, H., AND C. C. COCKERHAM. 1988. Variance components of fitness under stabilizing selection. Genet. Res. 51:47–53.

TEMPLETON, A. R. 1980. The theory of speciation via the founder principle. Genetics 94:1011–1038.

THEWISSEN, J. G. M., S. T. HUSSAIN, AND M. ARIF. 1994. Fossil evidence for the origin of aquatic locomotion in Archaeocete whales. Science 263:210–212.

TURELLI, M. 1984. Heritable genetic variation via mutation-selection balance: Lerch's zeta meets the abdominal bristle. Theor. Pop. Biol. 25:138–193.

TURELLI, M., AND N. H. BARTON. 1994. Genetic and statistical analyses of strong selection on polygenic traits: What, me normal? Genetics 138:913–941.

TURELLI, M., AND A. A. HOFFMANN. 1991. Rapid spread of an inherited incompatibility factor in California *Drosophila*. Nature 353:440–442.

———. 1995. Cytoplasmic incompatibility in *Drosophila simulans*: Dynamics and parameter estimates from natural populations. Genetics 140:1319–1338.

TURNER, J. R. G. 1971. Müllerian mimicry in burnet moths and Heliconiid butterflies. Pp. 224–260 *in* E. R. Creed, ed. Ecological genetics and evolution. Blackwells, Oxford.

TURNER, J. R. G. 1977. Butterfly mimicry: The genetical evolution of an adaptation. Evol. Biol. 10:163–206.

TURNER, J. R. G., AND J. L. B. MALLET. 1996. Did forest islands drive the diversity of warningly coloured butterflies? Biotic drift and the shifting balance. Phil. Trans. R. Soc. Lond. B Biol. Sci. 351:835–845.

VAN BATENBURG, F. H. D., AND E. GITTENBERGER. 1996. Ease of

fixation of a change in coiling: Computer experiments on chirality in snails. Heredity 76:278–286.
WADE, M. J., AND C. J. GOODNIGHT. 1991. Wright's shifting balance theory: An experimental study. Science 253:1015–1018.
WAGNER, A., G. P. WAGNER, AND P. SIMILION. 1994. Epistasis can facilitate the evolution of reproductive isolation by peak shifts: A two-locus two-allele model. Genetics 138:533–545.
WASER, N. M., AND M. V. PRICE. 1994. Crossing-distance effects in *Delphinium nelsonii*: Outbreeding and inbreeding depression in progeny fitness. Evolution 48:842–852.
WEBER, K. E. 1996. Large genetic change at small fitness cost in large populations of *Drosophila melanogaster* selected for wind tunnel flight: Rethinking fitness surfaces. Genetics 144:205–213.
WEBER, K. E., AND L. T. DIGGINS. 1990. Increased selection response in larger populations. II. Selection for ethanol vapor resistance in *Drosophila melanogaster* at two population sizes. Genetics 125:585–597.
WHITE, M. J. D. 1973. Animal cytology and evolution. 3d ed. Cambridge Univ. Press, Cambridge.
———. 1978. Modes of speciation. Freeman, San Francisco, CA.
WHITLOCK, M. C. 1995. Variance-induced peak shifts. Evolution 49:252–259.
———. In press. Founder effects and peak shifts without genetic drift: Adaptive peak shifts occur easily when environments fluctuate slightly. Am. Nat.
WHITLOCK, M. C., P. C. PHILLIPS, F. B.-G. MOORE, AND S. J. TONSOR. 1995. Multiple fitness peaks and epistasis. Ann. Rev. Ecol. Syst. 26:601–629.
WILLIAMS, S. M., AND S. SARKAR. 1994. Assortative mating and the adaptive landscape. Evolution 48:868–875.
WILLIS, J. H., AND H. A. ORR. 1993. Increased heritable variation following population bottlenecks: The role of dominance. Evolution 47:949–956.
WILSON, D. S., AND R. K. COLWELL. 1981. Evolution of sex ratio in structured demes. Evolution 35:882–897.
WILSON, D. S., AND M. TURELLI. 1986. Stable underdominance and the evolutionary invasion of empty niches. Am. Nat. 127:835–850.
WILSON, W. G., W. G. LAIDLAW, AND K. VASUDEVAN. 1994. Residual statics estimation using the genetic algorithm. Geophysics 59:766–774.
WRAY, N. R., AND R. THOMPSON. 1990. Predictions of rates of inbreeding in selected populations. Genet. Res. 55:41–54.
WRIGHT, S. 1922. The effects of inbreeding and crossbreeding on guinea pigs. III. Crosses between highly inbred families. Bulletin No. 1121. US Department of Agriculture, Washington, DC.
———. 1923a. Mendelian analysis of the pure breeds of livestock. I. The measurement of inbreeding and relationship. J. Hered. 14:339–348.
———. 1923b. Mendelian analysis of the pure breeds of livestock. II. The Duchess family of Shorthorns as bred by Thomas Bates. J. Hered. 14:405–422.
———. 1925. Mendelian analysis of the pure breeds of livestock. III. The shorthorns. J. Hered. 16:205–215.
———. 1931. Evolution in Mendelian populations. Genetics 16: 97–159.
———. 1932. The roles of mutation, inbreeding, crossbreeding and selection in evolution. Proc. 6th Int. Cong. Genet. 1:356–366.
———. 1935a. The analysis of variance and the correlation between relatives with respect to deviations from an optimum. J. Genet. 30:243–256.
———. 1935b. Evolution in populations in approximate equilibrium. J. Genet. 30:257–266.
———. 1937a. The distribution of gene frequencies in populations. Science 85:504.
———. 1937b. The distribution of gene frequencies in populations. Proc. Nat. Acad. Sci. USA 23:307–320.
———. 1939. Statistical genetics in relation to evolution. Pp. 5–64 *in* Actualités scientifiques et industrielles. No. 802. Eposés de Biométrie et de la statistique biologique XIII. Hermann et Cie., Paris.
———. 1940. The statistical consequences of Mendelian heredity in relation to speciation. Pp. 161–183 *in* J. S. Huxley, ed. The new systematics. Oxford Univ. Press, Oxford.
———. 1941. On the probability of fixation of reciprocal translocations. Am. Nat. 75:513–522.
———. 1943. Isolation by distance. Genetics 28:114–138.
———. 1948. On the roles of directed and random changes in gene frequency in the genetics of populations. Evolution 2:279–294.
———. 1951. The genetical structure of populations. Ann. Eugen. 15:323–354.
———. 1964. Stochastic processes in evolution. Pp. 199–241 *in* J. Garland, ed. Stochastic models in medicine and biology. Univ. of Wisconsin Press, Madison.
———. 1967. "Surfaces" of selective value. Proc. Nat. Acad. Sci. USA 58:165–172.
———. 1977. Evolution and the genetics of populations. Vol. 3. Experimental results and evolutionary deductions. Univ. of Chicago Press, Chicago.
———. 1978. Evolution and the genetics of populations. Vol. 4. Variability within and among natural populations. Univ. of Chicago Press, Chicago.
———. 1982. The shifting balance theory and macroevolution. Ann. Rev. Genet. 16:1–19.
———. 1988. Surfaces of selective value revisited. Am. Nat. 131: 115–123.

Corresponding Editor: T. Markow

APPENDIX: A QUANTITATIVE-GENETIC EXAMPLE OF THE SBT

We consider stabilizing selection on a quantitative trait, following the model introduced by Wright (1935a,b). The trait is assumed to be strictly additive, but the nonlinear relation between trait and fitness leads to epistasis for fitness. In this context, Wright's adaptive landscape is a surface giving the mean fitness of a population as a function of allele frequencies at many loci; just two of these frequencies are shown in Figure 2A. In Wright's model, many genotypes produce a phenotype close to the optimum. Selection acts to fix one of these combinations because segregating mixtures of optimal genotypes will produce recombinant progeny that deviate from the optimum. Thus, there are many adaptive peaks. With 100 loci, each segregating for two alleles of equal and additive effect, any of the 10^{29} combinations of 50 "+" and 50 "−" alleles would give the same average phenotype, and any of these could be fixed by selection for an intermediate optimum. Wright's analysis of his polygenic model ignored several features that are essential to the SBT, such as population subdivision, a mechanism for maintaining variation, and asymmetries that would give different mean fitnesses to different adaptive peaks. All of these factors are considered in our model. Our analysis shows that the SB process can increase fitness in the way Wright suggested, but it also indicates some of the limitations of the process.

In this example, we consider selection on a phenotypic trait determined additively by allelic contributions from $n = 100$ loci, plus an environmental contribution with mean = 0 and variance $V_e = 5$. We assume diploidy and random mating and ignore linkage disequilibrium. We assume that phenotypic fitness, $W(z)$, decreases quadratically with deviation from an intermediate optimum (arbitrarily scaled to be 0): $W(z) = 1 - z^2/2w^2$, where z is the individual phenotype and $w^2 = 20$ measures the strength of stabilizing selection. As noted by Wright (1935a), this nonlinear fitness function produces epistatic fitness interactions. At each of the 100 loci, there are two alleles with effects $\alpha = \pm 0.5$ on the trait. Variation is maintained by symmetric mutation ($\mu = \nu = 10^{-4}$). An infinite panmictic population would reach an equilibrium with m loci almost fixed for the "−" allele, and M for the "+" allele. Any equilibrium with $\{m, M\} = \{50, 50\}$ would give genetic variance near $V_g = 4n\mu(w^2 + V_e) = 1$, and mean fitness near $1 - 2n\mu = 0.98$ (Latter 1960; Turelli 1984). Other equilibria are stable, and correspond to lower adaptive peaks (Barton 1986). For example, the class with $\{m, M\} = \{48, 52\}$ gives $V_g = 1.51$, and mean fitness 0.969. However, these asymmetric equilibria differ relatively little in fitness.

As Wright argued, direct pleiotropic effects of the alleles on fitness are likely to cause substantial differences in the heights of the various adaptive peaks. Suppose that at 50 loci, each "+" allele increases fitness by $S/2$ and each "−" allele decreases fitness by $-S/2$, while at

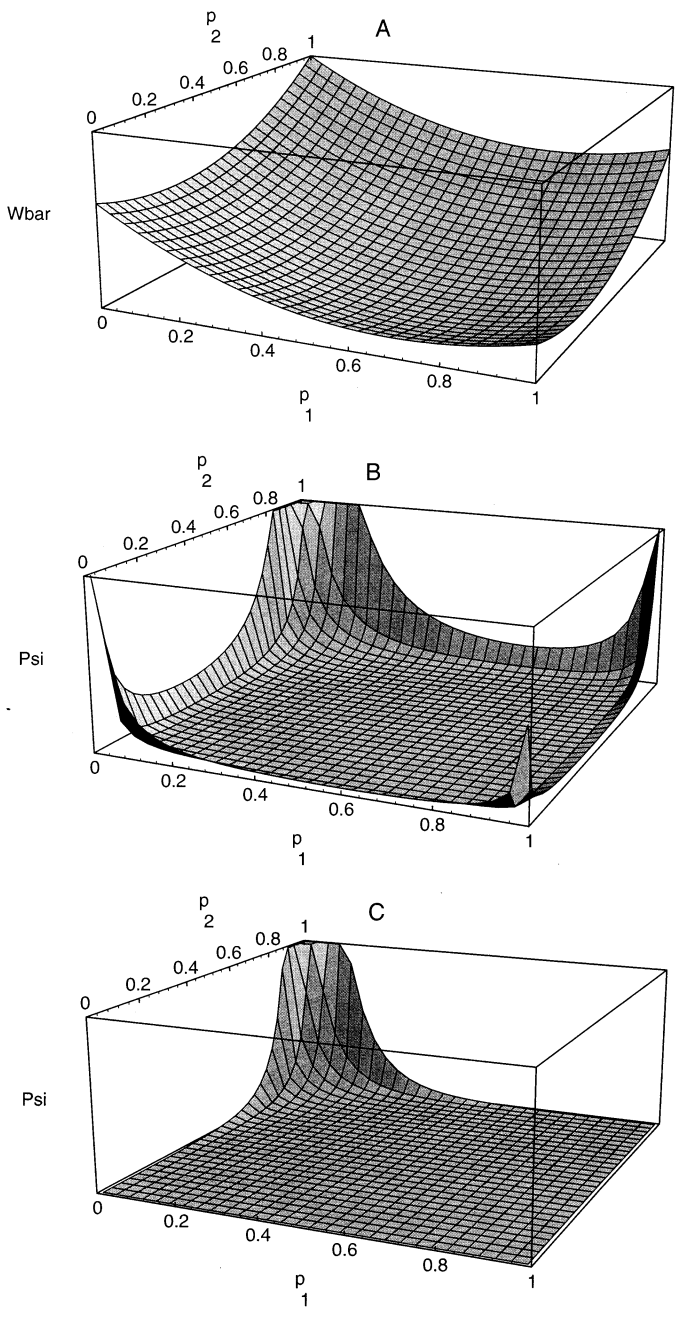

FIG. 2. Stabilizing selection on a polygenic trait (see text for details, including parameter values). (A) The mean fitness, plotted against allele frequencies at two loci. This surface is computed under the assumption that as allele frequencies at these loci change, those at other loci do too, so that the population mean remains at the optimum. Mean fitness is lower for intermediate allele frequencies because they increase the genetic variance; thus selection favors fixation at one of the four adaptive peaks. These peaks differ in height because of the pleiotropic effects on fitness. (B) The equilibrium probability distribution of allele frequencies, computing using equation (1), at two out of 100 loci, with no migration. At one locus the "+" allele has a positive pleiotropic effect on fitness ($S = 0.002$), while at the other, the "−" allele has a positive pleiotropic effect. The phenotypic optimum can be achieved with each locus near fixation for either the "+" or the "−" allele, but because of the pleiotropic effects of the alleles, one of these four states—corresponding to the highest mean fitness—is most likely to be occupied. For these two loci, there is a 99.77% chance that the optimal pair of alleles will be near fixation (highest peak, left rear). However, the probability that all 100 loci are near fixation for the alleles with positive pleiotropic effects is only 6.1×10^{-7}. (C) The equilibrium probability distribution of allele frequencies, as in (B), but with $Nm = 1$ migrants per generation rather than $m = 0$. Now, almost all demes approach the optimal adaptive peak: 89.43% of demes will be close fixation for the fitter alleles at all 100 loci. These distributions are calculated using the two-locus version of equation (A 1.5) of Barton (1989b).

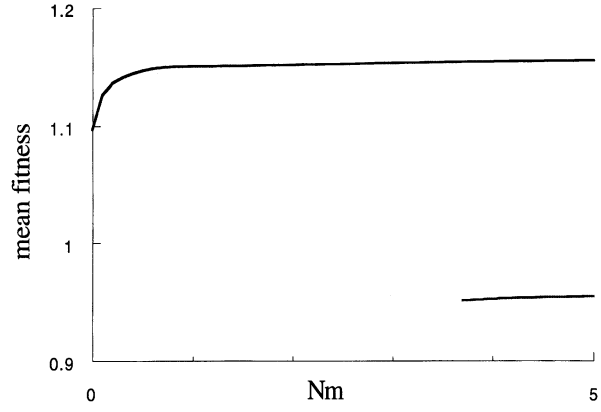

FIG. 3. Mean fitness ($\bar{W}$) as a function of the number of migrants, Nm, for the model of Figure 2. With no migration, the demes drift across the adaptive landscape. Only a small fraction of demes (6.1×10^{-7}) approach the optimal 100-loci gene combination. Mean fitness is $\bar{W} = 1.097$ (left intercept), compared with $W = 1.2$ for the fittest genotype. As migration increases, more demes approach the optimal adaptive peak, and mean fitness rises toward the maximum of 1.180 for a panmictic population (upper curve). (This is reduced below the maximum fitness by the mutation load, $2n\mu = 0.02$.) However, when the number of migrants rises above a critical value ($Nm > 3.6$), the whole metapopulation can also be trapped at an inferior peak. The lower curve shows the mean fitness at a specific nonoptimal stochastic equilibrium at which the common allele has deleterious pleiotropic effects at 50 loci, and favorable effects at the other 50 loci. With Nm above the critical threshold, the population can settle into very many alternative states, corresponding to different combinations of favorable and unfavorable alleles; these can be imagined as a series of curves parallel to those shown. Mean fitness is derived from equations (A1.6–A1.10) of Barton (1989b).

the other 50 loci, the signs of the pleiotropic effects are reversed. We assume that these pleiotropic effects contribute additively to individual fitness, both within loci and across loci, and that these fitness contributions are added to the fitnesses determined by the quadratic stabilizing selection. The intensity of stabilizing selection is quantified by $s = 1/(w^2 + V_e)$. If $S > s\alpha^2/2 = 0.005$, and the population is large, the optimal combination will inevitably be fixed, giving mean fitness $1 - 2n\mu + nS$. However, if $S < s\alpha^2/2$, any of the very many inferior combinations can approach fixation. For example, with $S = 0.002$, the best combination gives mean fitness 1.180, but the worst combination (in which the least favorable allele approaches fixation at all loci) is also stable, giving mean fitness $1 - 2n\mu - nS = 0.780$. In the SBT, random drift allows escape from such inferior equilibria.

Figure 2A shows a cross-section through the adaptive landscape, giving mean fitness as a function of just two of the allele frequencies, allowing the landscape to be depicted in three dimensions. Intermediate values of the allele frequencies lower mean fitness because of increased genetic variation, and pleiotropic effects of the alleles (with $S = 0.002$) tilt the landscape to favor one particular gene combination. The full stochastic model is analyzed by numerically solving for equilibria using equation 1. With no migration, demes of $N = 250$ individuals can

wander across this adaptive landscape, and can approach fixation for many different combinations of alleles (Fig. 2B). In the population as a whole, only a proportion 6.1×10^{-7} of demes approach the optimal gene combination, and mean fitness is $\bar{W} = 1.097$. In contrast, with $Nm = 1$ (Fig. 2C), 89.43% of demes will be near fixation for the fitter alleles across all 100 loci.

Figure 3 shows the effect of migration on mean fitness, calculated by averaging over the whole space of allele frequencies. As the number of migrants (Nm) increases, demes are pulled toward the gene combination that is commonest in the population as a whole. This positive feedback increases the proportion of demes near the highest peak, and raises mean fitness. For instance, with $Nm = 1$, the mean fitness increases to 1.151, and approaches the maximum of 1.180 with panmixis (upper right of Fig. 3). However, as migration rises above a critical value, $Nm_{crit} = 3.6$, the whole set of populations can also become trapped at an arbitrary adaptive peak that may have suboptimal mean fitness. In this case, the gene pool approaches fixation for the "−" allele at some loci and the + allele at others across all demes. The lower curve in Figure 3 shows the mean fitness when the more common allele has favorable pleiotropic effects at half the loci, and unfavorable effects at the others. Mean fitness then rises very slowly from 0.951 at the critical point ($Nm_{crit} = 3.6$) to a maximum of 0.998 with panmixis (only a small part of this rise is shown in Fig. 3). Thus, if we consider sets of populations each starting with different genetic compositions, we expect the highest mean fitness for an intermediate value of Nm, just less than $Nm_{crit} = 3.6$ (numerical analyses suggest that this critical value holds for many of the alternative stationary distributions). This level of migration insures a relatively high mean fitness, but does not permit the population to be trapped at a suboptimal peak.

This example uses the methods introduced by Wright to give a quantitative model of the SBT; for the parameters used here, it supports his intuition that a shifting balance between selection, mutation, migration, and drift can ensure higher mean fitness than with complete panmixis, or complete subdivision. However, even within the constraints of this model, this conclusion holds for only a limited range of parameter values. With stronger pleiotropy ($S > s\alpha^2/2$), mass selection fixes the optimal combination. With weaker pleiotropy, mean fitness can *decrease* as migration increases from zero: this is because although more demes approach the highest peak, migration between demes with different gene combinations introduces genetic variance, which reduces mean fitness (numerical results not shown). With stronger selection ($Ns\alpha^2 \gg 1$), the approach to equilibrium becomes extremely slow ($T \approx \exp[Ns\alpha^2]$), since it relies on successive peak shifts. With weaker selection, on the other hand, demes do not approach definite adaptive peaks; the population then behaves as if panmictic, but with the efficacy of selection reduced by random loss of variation within demes (Barton and Rouhani 1993). Finally, even in the absence of random drift, varying selection will tend to lead the population toward the highest adaptive peak. In Figure 2A, for example, a slight change in the optimum phenotype or the pleiotropic effects will tilt the adaptive landscape enough to remove the adaptive valleys and fix the fittest gene combination.