EVOLUTION

INTERNATIONAL JOURNAL OF ORGANIC EVOLUTION

PUBLISHED BY

THE SOCIETY FOR THE STUDY OF EVOLUTION

Vol. 57 **October 2003** **No. 10**

Evolution, 57(10), 2003, pp. 2197–2215

PERSPECTIVE: MODELS OF SPECIATION: WHAT HAVE WE LEARNED IN 40 YEARS?

SERGEY GAVRILETS

Department of Ecology and Evolutionary Biology, Department of Mathematics, University of Tennessee, Knoxville, Tennessee 37996

E-mail: sergey@tiem.utk.edu

Abstract.—Theoretical studies of speciation have been dominated by numerical simulations aiming to demonstrate that speciation in a certain scenario may occur. What is needed now is a shift in focus to identifying more general rules and patterns in the dynamics of speciation. The crucial step in achieving this goal is the development of simple and general dynamical models that can be studied not only numerically but analytically as well. I review some of the existing analytical results on speciation. I first show why the classical theories of speciation by peak shifts across adaptive valleys driven by random genetic drift run into trouble (and into what kind of trouble). Then I describe the Bateson-Dobzhansky-Muller (BDM) model of speciation that does not require overcoming selection. I describe exactly how the probability of speciation, the average waiting time to speciation, and the average duration of speciation depend on the mutation and migration rates, population size, and selection for local adaptation. The BDM model postulates a rather specific genetic architecture of reproductive isolation. I then show exactly why the genetic architecture required by the BDM model should be common in general. Next I consider the multilocus generalizations of the BDM model again concentrating on the qualitative characteristics of speciation such as the average waiting time to speciation and the average duration of speciation. Finally, I consider two models of sympatric speciation in which the conditions for sympatric speciation were found analytically. A number of important conclusions have emerged from analytical studies. Unless the population size is small and the adaptive valley is shallow, the waiting time to a stochastic transition between the adaptive peaks is extremely long. However, if transition does happen, it is very quick. Speciation can occur by mutation and random drift alone with no contribution from selection as different populations accumulate incompatible genes. The importance of mutations and drift in speciation is augmented by the general structure of adaptive landscapes. Speciation can be understood as the divergence along nearly neutral networks and holey adaptive landscapes (driven by mutation, drift, and selection for adaptation to a local biotic and/or abiotic environment) accompanied by the accumulation of reproductive isolation as a by-product. The waiting time to speciation driven by mutation and drift is typically very long. Selection for local adaptation (either acting directly on the loci underlying reproductive isolation via their pleiotropic effects or acting indirectly via establishing a genetic barrier to gene flow) can significantly decrease the waiting time to speciation. In the parapatric case the average actual duration of speciation is much shorter than the average waiting time to speciation. Speciation is expected to be triggered by changes in the environment. Once genetic changes underlying speciation start, they go to completion very rapidly. Sympatric speciation is possible if disruptive selection and/or assortativeness in mating are strong enough. Sympatric speciation is promoted if costs of being choosy are small (or absent) and if linkage between the loci experiencing disruptive selection and those controlling assortative mating is strong.

Key words.—Allopatric, mathematical, models, parapatric, speciation, sympatric, theory.

Received December 3, 2002. Accepted April 1, 2003.

Speciation, that is, the origin of species, is one of the most intriguing evolutionary processes. Speciation is the process directly responsible for the diversity of life. Understanding speciation still remains a major challenge faced by evolutionary biology even now after almost 150 years since the publication of Darwin's book, *On the Origin of Species*. Although the argument on the appropriate way(s) to define a ''species'' is still unsettled (Wilson 1999), speciation is ultimately a consequence of genetic divergence. Here, I will view the dynamics of speciation as the dynamics of genetic divergence between different populations (or between parts of the same population) resulting in substantial reproductive isolation. A suite of tools for modeling the dynamics of genetic divergence is provided by theoretical population genetics and ecology. These two theories form a foundation for the quantitative/mathematical study of speciation.

Theoretical population genetics has identified a number of factors controlling evolutionary dynamics such as mutation, random genetic drift, recombination, natural and sexual selection, etc. A straight-forward approach for classifying dif-

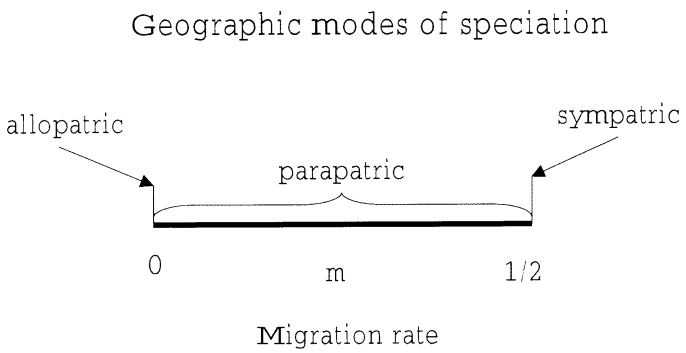


FIG. 1. Geographic modes of speciation.

ferent mechanisms and modes of speciation is according to the type and strength of the factors controlling or driving genetic divergence. In principle, any of the factors listed above can be used at any level of classification. However, traditionally in evolutionary biology the discussions of speciation mechanisms are framed in terms of a classification in which the primary division is according to the level of migration between the diverging (sub)populations (Mayr 1942). In this classification the three basic (geographic) modes of speciation are allopatric, parapatric, and sympatric.

Allopatric speciation (from Greek words *allos* meaning ''other'' and *patra* meaning ''fatherland'' or ''country'') is the origin of new species from geographically isolated (sub)populations. In the allopatric case there is no migration of individuals (and gene flow) between the diverging (sub)populations. At the other extreme of the highest possible migration level lies sympatric speciation (from Greek word *sym* meaning ''the same''). Sympatric speciation is usually defined as the origin of new species from a single local population (Mayr 1942). This definition (as well as similar definitions based on the ''absence of geographic separation,'' occurrence within a ''cruising range,'' occurrence within a ''geographic range'' of ancestor, etc., although intuitively appealing, is not precise enough for modeling purposes. Indeed, to use such a definition one would need to additionally define the exact meaning of ''local population,'' ''geographic separation'' or ''cruising range,'' etc. I will define sympatric speciation as the emergence of new species from a population where mating is random with respect to the place of birth of the mating partners (for a similar approach, see Kondrashov and Mina 1986). This definition is actually implied in most mathematical models of sympatric speciation. The intermediate cases when migration between diverging (sub)populations is neither zero nor maximum possibly fall within the domain of parapatric speciation (from Greek word *para* meaning ''beside'', ''side-by-side,'' ''next to'').

To illustrate this classification of speciation modes, let us consider a system of two demes that exchange a proportion m ($\leq 1/2$) of its members each generation. Assume that mating follows migration and is random within each deme. Then the allopatric case corresponds to $m = 0$, the sympatric case corresponds to $m=1/2$, and parapatric case corresponds to $0<m<1/2$ (see Figure 1). Both this figure and biological intuition suggest that parapatric speciation is the most general (geographic) mode of speciation. Endler (1977) made the same point on the basis of empirical data. I conclude this discussion with a few comments some of which were made repeatedly in the past (e.g., Smith 1955, 1969; Endler 1977). The first is that a consistent use of terminology is very important. Sometimes the notion of sympatric speciation is used as an antonym of allopatric speciation, and thus includes both parapatric speciation and sympatric speciation as defined above. This lays grounds for confusion. Even more misleading is the usage of the term ''sympatric speciation'' as a synonym of ''the origin of species in a well defined geographic area.'' For example, both in the empirical and theoretical literature one can easily find statements about ''sympatric speciation of Lake Victoria cichlids'' which are only a little less misleading than something like ''sympatric speciation of North American birds.'' One should also be absolutely clear that the definitions of the geographic modes of speciation imply nothing about the forces that drive genetic divergence leading to speciation. Any given evolutionary factor can play a role within each of the three geographic modes. For example, sympatric speciation can be driven by mutation and random drift and allopatric speciation can be driven by ecological factors. Models describing both these examples exist. Finally, the traditional stress on the spatial structure of (sub)populations as the primary factor of classification rather than, say, on selection reflects both the fact that it is most easily observed (relative to the difficulties in inferring the type and/or strength of selection acting in natural populations) and the growing realization that spatial structure of populations is very important. Therefore, I do not find compelling recent suggestions to abandon this classification in favor of a classification based on types of selection (Via 2001) or make it subordinate to a ''geography/prezygotic isolating mechanisms'' continuum (Kirkpatrick and Ravigné 2002).

Speciation is a very complex process which is affected by many different factors (genetical, ecological, developmental, environmental, etc.) interacting in nonlinear ways. Both this complexity and the difficulties of experimental approaches coming in particular from the very long time scales that are typically involved imply that mathematical models have to play a very important role in speciation research. The ability of models to provide insights into the speciation process, to train our intuition, to provide a general framework for studying speciation, and to identify key components in its dynamics are invaluable.

For reasons that are unclear, models of speciation were slow to appear. Paradoxically, none of the four greats who are usually viewed as the founders of modern theoretical population genetics—Fisher, Wright, Haldane, Kimura—expressed much interest in developing mathematical models explicitly dealing with speciation. (Among their work most closely related to speciation are Fisher's verbal models of runaway evolution caused by sexual selection [1930], Wright's verbal theory of shifting balance [1931, 1982] and his model of assortative mating [1921], studies of stochastic peak shifts [Haldane 1931; Kimura 1985; Wright 1941], and clines [Fisher 1950; Haldane 1948]). To my knowledge, the first substantial discussions of speciation utilizing formal mathematical models are those by Maynard Smith (1962, 1966) and Bazykin (1965, 1969). Although these papers have been very important for setting the field and one of them has

become a ''citation classic,'' as far as concrete results are concerned, they were mostly about modeling the maintenance of genetic variation rather than about speciation per se. A part of the 1966 paper by Maynard Smith did treat speciation directly, but only in a somewhat awkward numerical way for a single set of parameter values assuming complete reproductive isolation from the start. Systematic studies of speciation utilizing mathematical models started only in the early 1970s with the pioneering papers of Crosby (1970), Dickinson and Antonovics (1973), and Balkau and Feldman (1973). Crosby (1973) was the first to use an individual-based model for describing parapatric speciation. Dickinson and Antonovics (1973) were the first to use numerical iteration of dynamic equations to explore a wide range of parameter values for a specific model of parapatric and sympatric speciation. Balkau and Feldman (1973) were the first to find conditions for parapatric speciation in a certain model analytically. For comparison, according to Rice and Hostert's (1993) review (the title of which I have adapted for this paper), the earliest experimental paper on speciation is that of Koopman published in 1950, and by the early 1970s nineteen more experimental papers had been published.

However now the situation has dramatically changed and the theoretical studies well outnumber the experimental papers. This imbalance only continues to grow. For example, recent reviews of experiments on speciation (Florin and Ödeen 2002; Ödeen and Florin 2000, 2002) list only four new experimental papers published since Rice and Hostert's 1993 review, while during the same time interval dozens of new modeling papers came out. At present there have been at least a hundred papers modeling some aspects of speciation (Kirkpatrick and Ravigné 2002). Unfortunately, theoretical speciation research provides an example where the second law of dialectics has failed, at least so far. (The second law of dialectics is concerned with the transformation of quantity into quality [e.g., Engels 1940, Hegel 1975].) Instead of a sound mathematical theory of speciation, there is what I can only interpret as a (justifiable) frustration among both empiricists and theoreticians with the theoretical speciation research. For example, Via (2001), an empiricist, who is very favorable to the idea of sympatric speciation and appears to be convinced it is very probable from a theoretical point of view still asks theoreticians (who have already published dozens of papers on the subject) to ''generalize, by further theoretical exploration and integration of existing models, the theoretical conditions under which sympatric speciation can occur'' (p. 389). Theoreticians Kirkpatrick and Ravigné (2002) talk about the ''balkanization'' of the theory of speciation and the absence of clear and general results worthy to be included in the textbooks. Kirkpatrick and Ravigné devised a scheme for classifying the published models of speciation in a hope to better understand their results. Turelli et al. (2001) paint a very gloomy picture of the field of theoretical research concluding in particular that because of the complexity of the process of speciation . . . ''progress on major issues . . . is more likely to emerge from empirical than from mathematical analyses'' (p. 330).

I believe the current situation stems primarily from the limitations of the methods used in theoretical speciation research. Most modeling papers on speciation use numerical simulations as the basic tool. The most general feature of simulation models is that their results are very specific. Interpretation of numerical simulations, the interpolation of their results for other parameter values, and making generalizations based on simulations are notoriously difficult. For these reasons the results of simulations are typically wide open to interpretation and narrowing of interpretation is not possible without additional work. However, simulation results are usually impossible to reproduce because many technical details are not described in original publications. At the same time, the authors of the original publications (and this is not limited to theoretical research) almost never return to their work to check the generality of their conclusions under more general or reasonable parameter values or assumptions.

What is missing in the theoretical speciation research are general and transparent analytical results comparable to those in other areas of theoretical population genetics and ecology. It is simple mathematical models allowing for analytical investigation (rather than complex numerical models) that form the basis of most scientific theories, and there are no reasons why evolutionary biology, in general, and speciation research, in particular, should be an exception. The questions that can be answered depend on the tools used in theoretical research. The majority of existing models of speciation (especially models of sympatric speciation) mostly demonstrate that speciation in a certain scenario may occur. What is needed now is a shift in focus to identifying more general rules and patterns. In principle, to learn that ''selection promotes speciation'' or that ''peak shifts by random genetic drift are unlikely'' one does not need mathematical models because this is what biological intuition already tells us. Mathematical models are needed only if one wants to get much more specific conclusions and predictions about the effects of genetic architecture and different types and intensities of selection on the probability of speciation or peak shift and their dynamical characteristics. Such knowledge can be coupled with the estimates of relevant parameters from natural populations to get a much deeper understanding of the process. Mathematical models are also needed if intuition is not working because of the complexity of the process under consideration.

My goal here is to describe and illustrate some existing theoretical results on speciation that are both simple and general. The common wisdom is that a picture is worth a thousand words. In the exact sciences, an equation is worth a thousand pictures. Equations and their interpretations are the most concrete results a theoretician can come up with. Therefore, equations and their interpretations are the focus of this paper. The extensive body of purely numerical work on speciation is deemphasized. The equations to be discussed below are easily accessible to biologists who lack mathematical training. They represent a kind of a ''do-it-yourself'' kit which can be used by biologists to check or train their intuition about speciation by substituting specific numerical values of parameters and interpreting the results.

In the remainder of this paper, I first will show why the classical theories of speciation by peak shifts across adaptive valleys driven by random genetic drift run into troubles (and into what kind of troubles). Then I will describe the Bateson-Dobzhansky-Muller (BDM) model of speciation that does not require overcoming selection. I describe exactly how the

probability of speciation, the average waiting time to speciation, and the average duration of speciation depend on the mutation and migration rates, population size, and selection for local adaptation. The BDM model postulates a rather specific genetic architecture of reproductive isolation. I then show exactly why the genetic architecture required by the BDM model should be common in general. The theoretical results to be described here have led to a major shift in focus from Wright's ''rugged adaptive landscapes'' to nearly neutral networks and holey adaptive landscapes. Next I consider the multilocus generalizations of the BDM model again concentrating on the qualitative characteristics of speciation such as the average waiting time to speciation and the average duration of speciation. Finally, I consider two models of sympatric speciation in which the conditions for sympatric speciation were found analytically. Results of this type, if known more widely, can clear up a lot of controversy that currently surrounds the questions about the plausibility and generality of sympatric speciation. The Appendix contains a glossary of symbols used in the paper.

Dynamics of Stochastic Peak Shifts

Here I describe two simple models of stochastic transitions between adaptive peaks driven by random genetic drift. My main goal is to show in quantitative terms that strong reproductive isolation is a very unlikely outcome of a stochastic peak shift.

One-locus two-allele model with underdominance

Following Lande (1979), let us consider a randomly mating diploid population with discrete and nonoverlapping generations. Assume that there is a single diallelic locus with alleles **A** and **a** controlling fitness (viability). Let the relative fitnesses of genotypes **AA** and **Aa** be $w_{\mathbf{AA}} = 1$, $w_{\mathbf{Aa}} = 1 - s$ and $w_{\mathbf{aa}}$, respectively ($0 \le s \le 1$) . The adaptive landscape corresponding to this model has two ''adaptive peaks'' at genotypes **AA** and **aa**, and an ''adaptive valley'' at genotype **Aa** (see Fig. 2a).

Assume that initially the population is monomorphic for allele **A**. Let us allow for mutation from **A** to **a** with a small probability μ per gene per generation. Selection will tend to eliminate alleles **a**, whereas mutation will continuously supply them. If the population size N is very large, the population will stay at a state where allele **a** is maintained at a very low frequency μ/s. However, if the population size N is small, the frequency of the mutant allele **a** will eventually cross 1/2 as a result of random genetic drift and approach 1 resulting in a ''peak shift.'' Parameter s, measuring the depth of the adaptive valley that separates the peaks, also characterizes the strength of (postmating) reproductive isolation between the population states before and after the peak shift.

How long does one have to wait until a peak shift occurs in this model? If the mutation rate is very small and the product of N and s is at least moderately large (>2) , the average waiting time until a peak shift occurs can be approximated as

$$T \approx \frac{1}{\mu}\frac{\sqrt{\pi}}{2}\frac{e^{Ns}}{\sqrt{Ns}}$$

(Lande 1979). This equation shows that T grows exponentially with the product Ns. (Note that although the term $1/\sqrt{Ns}$ does decrease with increasing Ns, it is dominated by the exponential term). For example, if $\mu = 10^{-5}$ (which is a common estimate of the mutation rate, e.g., Futuyma 1998, Griffiths et al. 1996), $s=0.05$ (which implies a shallow adaptive valley) and $N=400$ (which is a relatively small population size), then $T = 10^{13}$ generations, which is a very long time, in fact, far longer than the probable age of any extant vertebrate species.

Most of this time will be spent waiting for a ''lucky'' mutant allele destined to be fixed to appear in the population as the overwhelming majority of mutant alleles are removed by selection and random genetic drift. One can also estimate the average time τ that it will take for the ''lucky'' mutant allele to take over the population. Time τ characterizes the average actual duration of stochastic transitions between the peaks. Using the diffusion approximation (e.g., Ewens 1979) one finds that this time is approximately

$$\tau \approx \frac{1}{s} + \frac{2}{s}\ln\left(\sqrt{\frac{Ns}{2}}\right).$$

The term $\ln(\sqrt{Ns/2})$ changes very slowly with the product Ns. Therefore, the actual duration of transition is on the order of $1/s$ generations and, thus, is much shorter than the average waiting time to transition T. For example, with the same values of parameters as above, $\tau = 66$ generations. The ratio of τ and T can be viewed as a measure of the likelihood to observe the actual transition. The above results show that observing the transition is very unlikely.

Additive quantitative character

Next, let us consider a diploid population of size N where individuals differ with respect to a single additive quantitative trait z. Assume that the genotypic variance of the trait is somehow maintained at a constant value G. Let the population be under viability selection defined by a fitness function

$$w(z) = \exp[-s(1 - z^2)^2],$$

where $s>0$. This fitness function has two local peaks at $z = -1$ and $z = 1$ separated by a valley at $z=0$ (see Fig. 2b). The height of the peaks is one, and the fitness at the bottom of the valley is $\exp(-s)$, which is approximately $1 - s$ for small s. That is, parameter s, as before, characterizes the depth of the adaptive valley separating the adaptive peaks.

Assume that initially the average trait value is in a neighborhood of one of the peaks. Random genetic drift will eventually result in the population crossing the adaptive valley and approaching the other peak (Barton and Charlesworth 1984; Lande 1985). In this model the average waiting time to the peak shift is approximately

$$T \approx \frac{\sqrt{2}\pi}{4Gs}e^{2Ns}.$$

Thus, as in the previous model, the waiting time to transition grows exponentially with Ns. Most of time T will be spent in a neighborhood of the initial adaptive peak waiting for the ''right'' sequence of stochastic events for the peak shift to

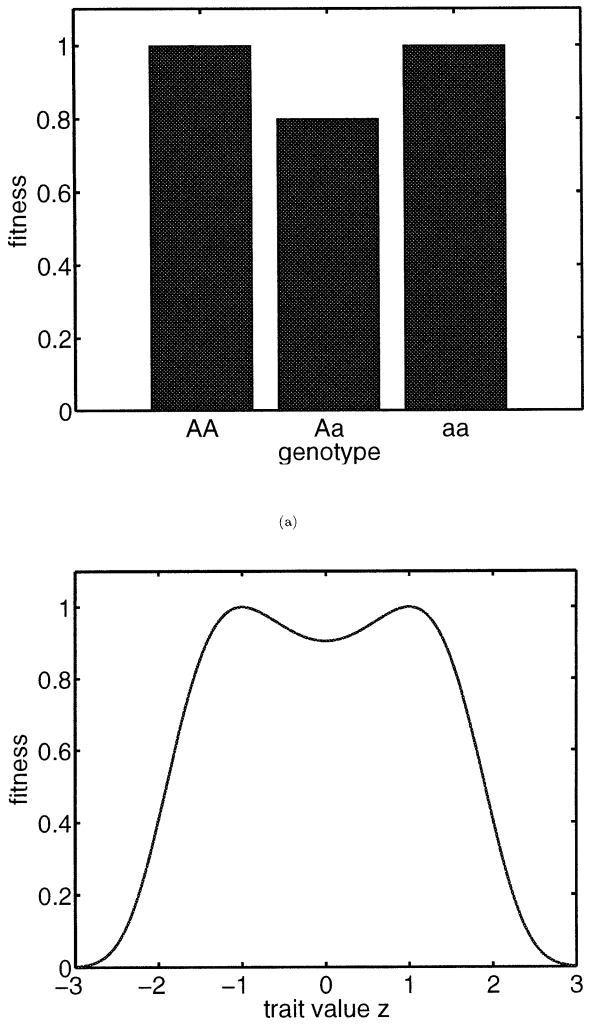


FIG. 2. Simple adaptive landscapes with two equal peaks. (a) One-locus two-allele diploid model with underdominance. (b) Disruptive selection on an additive quantitative character.

occur. The average actual duration of transition τ is approximately

$$\tau \approx \frac{1}{4Gs} \ln[8\sqrt{2}\sqrt{sN}],$$

that is, it has order $1/(Gs)$. For example, if the genetic variance $G = 0.04$ (which implies that the bottom of the valley is at five standard deviations from the peak), $s=0.05$ and $N=400$, then $T = 1.31 \times 10^{20}$ generations and $\tau = 490$ generations.

There are two important conclusions emerging from these simple analyses. First, unless the population size is small and the adaptive valley is shallow, the waiting time to a stochastic transition between the adaptive peaks is extremely long. This implies that it is very unlikely that a single peak shift will result in strong reproductive isolation. Second, if transition does happen, it is very quick. This implies that observing it ''in action'' is practically impossible. These conclusions are very general and apply to a broad class of stochastic transitions between adaptive peaks (e.g., Barton and Charlesworth 1984; Barton and Rouhani 1987). Strong reproductive isolation may follow after a chain of stochastic peak shifts each of which is across a shallow valley and, thus, results in a very small degree of reproductive isolation (Walsh 1982). If there are many adaptive peaks separated by shallow valleys, then the pattern of the evolutionary dynamics will be that of a sequence of rapid transitions between adaptive peaks separated by very long periods of no apparent change spent at each of the ''intermediate'' adaptive peaks. The extended stays at ''intermediate'' peaks implies that in this model of allopatric divergence the pattern of evolutionary dynamics is quite different from that of ''punctuated equilibrium'' (Eldredge 1971; Eldredge and Gould 1972; Gould 2002) in which intermediate stages are (almost) never observed. Below we will see that the pattern of ''punctuated equilibrium'' arises in models of parapatric speciation.

The Bateson-Dobzhansky-Muller Model

By the Bateson-Dobzhansky-Muller (BDM) model, I mean a very specific assumption about the genetic architecture of reproductive isolation, namely that there are two loci with two alleles at different loci being ''incompatible'' (Bateson 1909; Dobzhansky 1937; Muller 1942). The adaptive landscapes implied by this model are illustrated in Figure 3. In this Figure it is alleles **a** and **B** that are assumed to be incompatible in the sense that individuals carrying both of them have zero viability (Fig. 3a) or that females carrying allele **a** do not mate with males carrying allele **B** (Fig. 3b). Speciation occurs if two populations that have initially the same genetic composition end up at the opposite sides of the ridge of high fitness values as illustrated in Figure 3. In the BDM model, the populations are not required to cross any adaptive valleys as they simply follow a ridge of high fitness values. For example, hybrid inviability in the fish *Xiphophorus* behaves as if controlled by a simple BDM model (Orr and Presgraves 2000).

Assume that initially the population has genotype **AABB** fixed. Let us allow for mutation from **A** to **a** and from **B** to **b** at an equal rate μ per generation, and let us neglect the possibility of backward mutations. In this model the population will definitely reach the state with genotype **aabb** fixed (see Fig. 3b). This means that speciation is certain. How long does one have to wait until speciation and what is the actual duration of the speciation process? Can speciation happen by mutation and random genetic drift alone? What are the quantitative effects of selection for local adaptation on the dynamics of speciation? These are the questions I intend to answer in this section.

Let us define the average waiting time to speciation T as the average time that it takes to get from the ancestral state (i.e., with genotype **AABB** fixed) to the state of reproductive isolation (i.e., with genotype **aabb** fixed). Let us define the average actual duration of speciation τ as the average time that it takes to get from the ancestral state to the state of reproductive isolation *without returning to the ancestral state.* Time τ can also be thought of as the average duration of the intermediate stages in the actual transition to a state of reproductive isolation (speciation). In a sense, the ratio of τ and T characterizes the probability of observing incipient speciation. The dynamics of speciation can be modeled as a

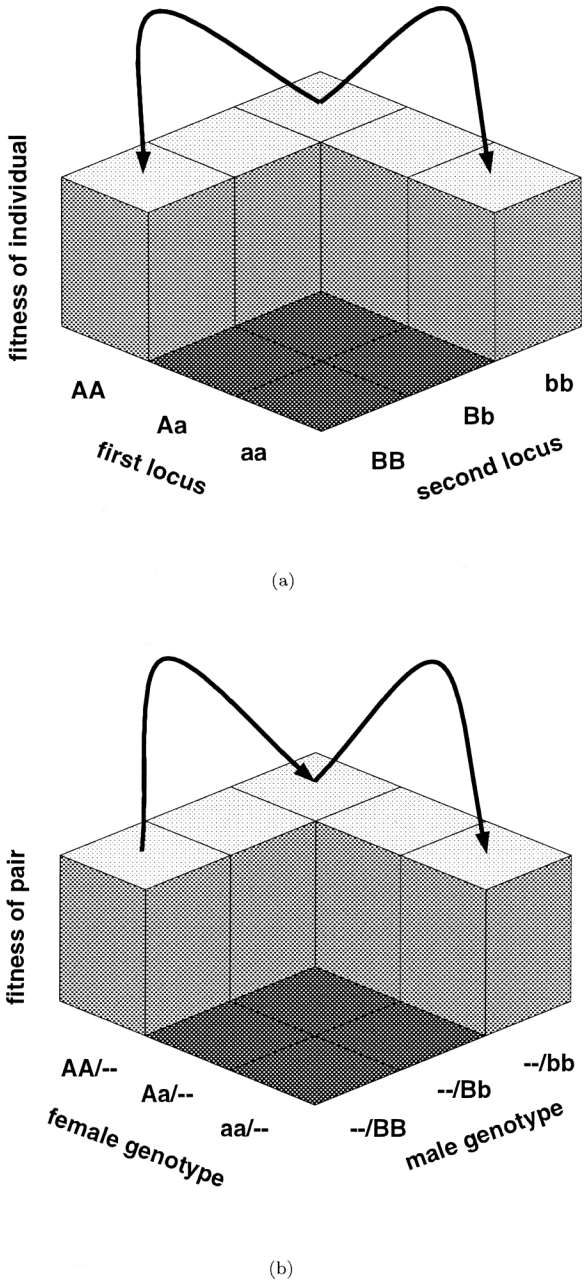


FIG. 3. Adaptive landscapes in the diploid BDM model. The height of a bar gives the fitness of the corresponding combination of genes. (a) Fitness of an individual. The presence of alleles **a** and **B** results in zero viability. (b) Fitness of a mating pair. Females carrying allele **a** do not mate with males carrying allele **B**. It is assumed that alleles at locus B are not expressed in females whereas alleles at locus A are not expressed in males. The arrows show two possible routes to the evolution of reproductive isolation.

random walk performed by the most common genotype in the population along the ridge of high fitness values (Gavrilets 2000a). I will describe the corresponding approximations for T and τ separately for allopatric and parapatric cases.

Allopatric speciation

First assume that the two loci under consideration have no other pleiotropic effects on fitness not related to reproductive isolation. In this case, speciation will be driven by mutation and random genetic drift which represents a general null model of speciation. The average waiting time to speciation and the average duration of speciation are equal to

$$T = \tau = \frac{2}{\mu}.$$

This last expression is very easy to understand. The average waiting time until fixation of a neutral allele is approximately the reciprocal of the mutation rate (Nei 1976). To evolve to a state with haplotype **ab** fixed, the population has to fix two neutral alleles—the fact reflected in the expression for T above. Note that because in the neutral case the rate of substitutions does not depend on the population size, the population size N does not enter the above equations for T and τ. Because in this model genetic divergence is irreversible, the average duration of speciation is equal to the average waiting time to speciation.

Next we allow for selection for local adaptation. I will assume that mutant alleles **a** and **b** have a small selective advantage s_a over the ancestral alleles **A** and **B** in the environment experienced by the population. In this case the average waiting time to speciation and the average duration of speciation are approximately

$$T = \tau \approx \frac{2}{\mu}\frac{1}{S_a},$$

where $S_a = 4Ns_a$ (and it is assumed that S_a is at least moderately large, i.e., >3) . The above expression strongly supports previous arguments (e.g., Schluter 2000) that selection for local adaptation can dramatically accelerate speciation. For example, increasing S_a by one order of magnitude will decrease the waiting time to speciation by one order of magnitude.

Parapatric speciation

In the parapatric case, we assume that each generation a small proportion m of the population is substituted by immigrants coming from a source population. All immigrants have ancestral genotype **AABB**.

First assume that speciation is driven by mutation and random genetic drift (the null model). Then the waiting time to speciation is

$$T = \left(2 + \frac{m}{\mu}\right)\frac{1}{\mu} \approx \frac{m}{\mu^2},$$

while the average duration of speciation is

$$\tau = \frac{1}{\mu}\frac{2 + \frac{m}{\mu}}{1 + \frac{m}{\mu}} \approx \frac{1}{\mu},$$

where the approximate equalities assume that the rate of mutation is much smaller than the rate of migration ($\mu \ll m$). For example, if $m = 0.01$ and $\mu = 10^{-5}$, then $T \approx 10^8$ generations and $\tau \approx 10^5$ generations. Note that the mutation rate has more influence on the dynamics of parapatric speciation than the migration rate.

Next we consider two models of selection for local adaptation. The first model describes *direct* selection on the alleles underlying reproductive isolation. Specifically, let us assume that mutant alleles **a** and **b** have a small selective advantage s_a over the ancestral alleles **A** and **B**. Then the average waiting time to speciation and the average duration of speciation are

$$T \approx \frac{1}{\mu}\left(2 + \frac{m}{\mu}e^{-S_a}\right)\frac{1}{S_a},$$

$$\tau \approx \frac{1}{\mu}\frac{2 + \frac{m}{\mu}e^{-S_a}}{1 + \frac{m}{\mu}e^{-S_a}}\frac{1}{S_a},$$

where it is assumed that the coefficient $S_a = 4Ns_a$ is at least moderately large (>3) . For example, with the same parameters as above and with $S_a=5$, then $T=1.73 \times 10^5$ generations and $\tau = 2.24 \times 10^4$ generations. That is, the average waiting time to speciation is reduced by three orders of magnitude relative to that in the model with no selection, and this brings it near a realistic range of values.

The second model considers effects of a *genetic barrier to neutral gene flow*. We assume that new alleles **a** and **b** have no effects on adaptation to local conditions but that the population has already diverged from the source population in some other loci controlling adaptation to local conditions. Now immigrants will have a reduced fitness, which will affect the fate of the ancestral alleles **A** and **B** that they carry. The effects of selection ''induced'' on neutral alleles via their association with some locally deleterious alleles can be characterized in terms of the *gene flow factor* (Bengtsson 1985; Barton and Bengtsson 1986; Gavrilets 1997a; Gavrilets and Cruzan 1998) defined as the probability that a neutral gene brought by immigrants makes it to the local genetic background. For example, assume that immigrating adults differ from the residents in two genes: a gene reducing the viability of F_1 hybrids to $1-s$ (relative to viability 1 of the residents) and a neutral gene. Assume also the average number of offspring produced by matings between the immigrants and residents, α, and by matings between F_1 hybrids and residents, β, is reduced relative to that of matings among residents (which is normalized to be 1) . This can happen if there is fertility and/or sexual selection against immigrants and/or hybrids. Then the gene flow factor is

$$\gamma = \frac{r(1-s)\alpha\beta}{1-(1-r)(1-s)\beta},$$

where r is the rate of recombination between the selected and neutral loci (Gavrilets and Cruzan 1998). For example, let $s = \alpha = \beta = 0.5$. Then if $r = 0.5$, then $\gamma = 0.07$; if $r = 0.05$, then $\gamma = 0.016$. That is, joint action of viability selection and fertility selection and assortative mating can significantly decrease the effective rate of immigration of neutral alleles even if they are unlinked to the locus under selection. Returning back to modeling speciation, let us assume that the gene flow factor resulting from selection for local adaptation is γ. Then approximately

$$T \approx \gamma\frac{m}{\mu^2}, \qquad \tau \approx \frac{1}{\mu}.$$

These equations show that a strong genetic barrier to the neutral gene flow (i.e., low γ) will significantly decrease the waiting time to speciation but will not affect its average duration. For example, if $\gamma = 0.1$, the waiting time to speciation reduces to one tenth of that in the case of divergence driven by mutation and drift.

There are several conclusions emerging from this analysis. Unless there is selection for local adaptation, the waiting time to speciation in the BDM model is very long (but not as long as under stochastic peak shift). Migration can significantly delay speciation. It is generally agreed upon that spatial heterogeneity in selection is common and that this can have profound effects of the possibility of speciation even in the presence of substantial gene flow (Endler 1977; Ogden and Thorpe 2002; Schneider et al. 1999; Smith et al. 1997). Our model makes this intuition more precise. Selection for local adaptation (either acting directly on the loci underlying reproductive isolation via their pleiotropic effects or acting indirectly via establishing a genetic barrier to gene flow) can significantly decrease the waiting time to speciation. Direct selection is much more effective than indirect selection. In the parapatric case the average actual duration of speciation is much shorter than the average waiting time to speciation.

Nearly Neutral Networks and Holey Adaptive Landscapes

We have seen that evolution of strong reproductive isolation via stochastic transitions between isolated adaptive peaks is unlikely. At the same time, the BDM model shows that strong reproductive isolation can be achieved if the population evolves along a ridge of high fitness values in the genotype space. How likely are such ridges in general? My goal in this section is to answer this question quantitatively. I will follow the general approach of Gavrilets and Gravner (1997) and Gavrilets (1997b, 2003).

To get some intuition about this question let us first consider the set of all possible haploid genotypes that have only two genes each with a very large number of alleles. The alleles at the first locus will be denoted as $\mathbf{A}_i$ and the alleles at the second locus will be denoted as $\mathbf{B}_j$. Assume a stepwise mutation pattern (Nei et al. 1983) allowing allele $\mathbf{A}_i$ mutate only to alleles $\mathbf{A}_{i+1}$ and $\mathbf{A}_{i-1}$ and in a similar way, allowing allele $\mathbf{B}_i$ mutate only to alleles $\mathbf{B}_{i+1}$ and $\mathbf{B}_{i-1}$. The genotype space (i.e., the set of all possible genotypes) in this model can be represented by a two-dimensional lattice of square sites where the x- and y- coordinates specify the alleles at the first and second locus, respectively.

Following Gavrilets and Gravner (1997), assume that genotype fitnesses (viabilities) are generated randomly and independently and are only equal to one (viable genotype) or zero (inviable genotype) with probabilities $\mathcal{P}$ and $1-\mathcal{P}$, respectively. Here, one might think of the set of all possible genotypes playing one round of Russian roulette with $\mathcal{P}$ being the probability to get a blank. In this model, any change in genotype no matter how large or small, results in a fitness value completely independent of the initial value. Note that this assumption is, of course, a great oversimplification which however does not affect our conclusions.

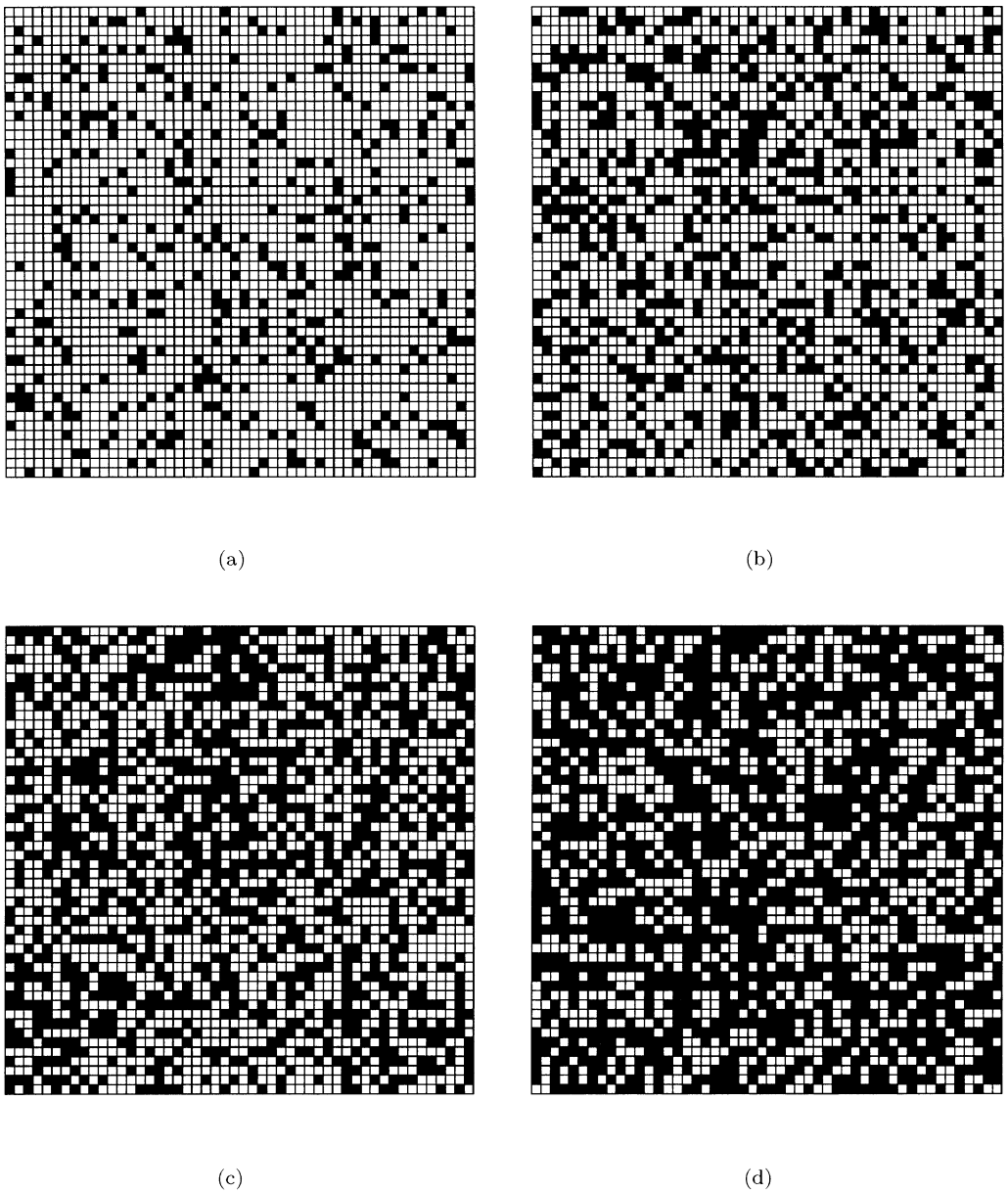


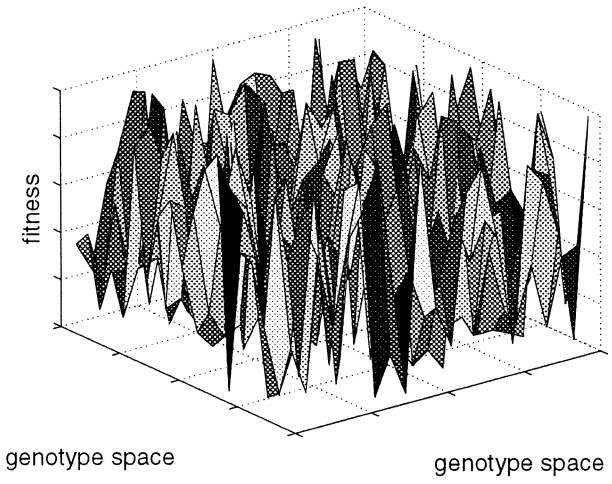
FIG. 4. Formation of clusters of viable genotypes (painted in black) in two dimensions for different values of the probability of being viable p. Inviable genotypes are painted in white. (a) $P = 0.15$, (b) $P = 0.30$, (c) $P = 0.45$, (d) $P = 0.60$.

Let us paint viable genotypes (sites) in black and inviable genotypes (sites) in white (see Fig. 4). For each site (genotype), its four adjacent sites (directly above, below, on the left, and on the right) represent other genotypes that can be obtained by a single mutation. In this model, viable genotypes tend to form connected networks (or ridges) that populations can evolve along by fixing single mutations without the need to go across any inviable states. The number and the structure of networks depend on the probability $\mathcal{P}$. For small values of $\mathcal{P}$ there are many networks of small size (see Fig. 4a). As $\mathcal{P}$ increases, the size of the largest network increases (see Fig. 4a,b,c). As $\mathcal{P}$ exceeds a certain threshold $\mathcal{P}_c$, known as the ''percolation threshold,'' there emerges the largest network (known as the ''giant component'') that extends (''percolates'') through the whole system and includes a significant proportion of all viable genotypes (see Fig. 4d). In this model, describing a so-called ''site percolation'' on an infinite two-dimensional lattice, the percolation threshold is $\mathcal{P}_c = 0.593$ (e.g., Grimmett 1989).

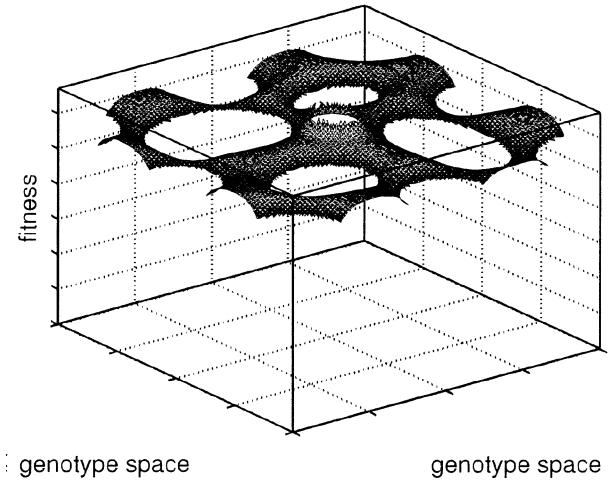
The percolation threshold drops dramatically with more loci and alleles. For example, if there are $\mathcal{L}$ loci each with $\mathcal{A}$ alleles, then the percolation threshold is

$$\mathcal{P}_c = \frac{1}{\mathcal{L}(\mathcal{A} - 1)}.$$

For example, if $\mathcal{L} = 10{,}000$ and $\mathcal{A} = 10$, then values of $\mathcal{P}$ larger than 10^{-5} will typically result in the existence of an extensive network of high-fitness ridges expanding throughout the genotype space. Here, the genotypes that belong to this network have the same fitness. In this sense, the network



(a)



(b)

FIG. 5. Adaptive landscapes. (a) A rugged adaptive landscape formed by genotypes with fitnesses assigned randomly from a uniform distribution. (b) A holey adaptive landscape formed by genotypes with fitnesses within a narrow fitness band. Evolution along a holey landscape is nearly neutral.

is *neutral*. In more general situations where a range of variation of fitness values is allowed, the above results can be generalized to demonstrate the existence of connected *nearly neutral networks* in which genotypes have approximately the same fitness (Gavrilets 1997b, 2003; Gavrilets and Gravner 1997).

Starting with Wright (1932) typical adaptive landscapes are usually imagined as very rough surfaces with many different peaks and valleys (see Fig. 5a). Continuous evolution on such ''rugged adaptive landscapes'' requires crossing adaptive valleys (as modeled above). The results on neutral and nearly neutral networks show that Wright's picture can be very misleading if adaptive landscapes have a very high dimensionality because such landscapes are characterized by the existence of extensive nearly neutral networks. Among different nearly neutral networks, those with sufficiently high fitnesses are of particular evolutionary importance because they allow for continuous evolutionary innovations without any significant loss in fitness. An important notion describing such networks is that of ''holey adaptive landscape.'' A *holey adaptive landscape* is defined as an adaptive landscape where relatively infrequent high-fitness genotypes form a contiguous set that expands throughout the genotype space. An appropriate three-dimensional image of such an adaptive landscape that focuses exclusively on the percolating network is a nearly flat surface with many holes representing genotypes that do not belong to the network (see Fig. 5b). The smoothness of the surface in this figure reflects close similarity between the fitnesses of the genotypes forming the corresponding nearly-neutral network. The ''holes'' include both lower fitness genotypes (''valleys'' and ''slopes'') and very high fitness genotypes (the ''tips'' of the adaptive peaks). The BDM model considered above provides one of the simplest examples of holey adaptive landscapes. Many more examples are known (Gavrilets 1997b, 2003; Gavrilets and Gravner 1997).

This section has dealt exclusively with the *structure* of adaptive landscapes. A number of mechanisms, such as random drift, pleiotropic selection, or random fluctuations in fitness can drive evolutionary divergence on the landscapes. The next section considers the first two mechanisms.

Multilocus Generalizations of the BDM Model

The BDM model was formulated in terms of only two loci. However, existing data on the genetics of reproductive isolation show that typically there are many different loci underlying reproductive isolation even at very early stages of divergence (Naveira and Masida 1998; Wu 2001; Wu and Palopoli 1994). If genetic architecture of reproductive isolation is known, the dynamics of speciation can be modeled, at least in principle. First, however, the genetic architecture of reproductive isolation is never known completely and, second, the mathematical treatment becomes extremely complicated as the number of genes involved increases. Therefore, rather than studying the dynamics of speciation *given a specific genetic architecture*, it becomes much more fruitful to look at the dynamics of speciation expected ''on average.'' (We have already used a similar approach in our discussion of the properties of typical adaptive landscapes in the previous section.) One such approach (Orr 1995; Orr and Orr 1996; Orr and Turelli 2001) reduces the complexity of adaptive landscapes underlying reproductive isolation to simpler effects of genetic ''incompatibilities'' of certain types that arise with certain probabilities. In what follows, two (or more) genes are called incompatible if their joint presence in an individual's genotype or in the genotypes of a (potential) mating pair results in reduction of a fitness component. The goal of this section is to find the average waiting time to speciation and the average duration of speciation.

Accumulation of genetic incompatibilities

Let us assume that there is a large number of diallelic loci potentially affecting reproductive isolation. Let any combination of k alleles at different loci be incompatible with probability q. (Note that in the BDM model $k=2$). Consider two populations that have diverged in d loci. Variable d is a measure of genetic distance separating the populations. As shown by Orr (1995), the expected number of incompatibilities between them is

$$J = q\binom{d}{k}, \tag{1}$$

where $\binom{d}{k} = d!/[k!(d-k)!]$ is the binomial coefficient (i.e., the number of combinations of k objects chosen from a set of d objects). The value of J increases very rapidly (''snowballs'') with genetic distance d (approximately as the k-th order of d) . Notice that the snowball effect is much more pronounced with larger values of k.

How do the incompatibilities translate into reproductive isolation between the populations? Let us assume that complete reproductive isolation occurs when C incompatibilities separate the populations. (Note that in the BDM model, $C=1$). Then the probability that two genotypes at distance d are not reproductively isolated is approximately

$$w(d) = \frac{\Gamma(C, J)}{\Gamma(C)} \tag{2}$$

(SG, unpubl. data). Here $\Gamma(\cdot, \cdot)$ and $\Gamma(\cdot)$ are the incomplete gamma function and gamma function, respectively (Gradshteyn and Ryzhik 1994) and J is given by equation (1). That the probability w decreases with d is compatible with a general empirical pattern that reproductive compatibility decreases with genetic distance between parents (Edmands 2002). The probability that two genotypes at distance d are reproductively isolated is $1 - w(d)$. One can also approximate the average K and the coefficient of variation CV_K of the number of substitutions required for speciation (i.e., complete reproductive isolation):

$$K \approx \left(\frac{k!}{q}C\right)^{1/k}, \tag{3a}$$

$$CV_K \approx \frac{1}{k\sqrt{C}} \tag{3b}$$

(S. Gavrilets, unpubl. data). As expected, K increases with the ratio C/q. The dependence of K on k is nonmonotonic with K minimized at some intermediate values of k. Figure 6 shows that as genetic divergence exceeds value K, the probability that no complete reproductive isolation occurs undergoes a rapid transition from 1 to 0. This ''threshold effect'' is especially strong when many complex incompatibilities are required for complete reproductive isolation. The latter feature is also apparent from the fact that the coefficient of variation CV_K quickly goes to zero as both C and k become large.

It should be intuitively clear that the adaptive landscapes implied by this model are ''holey.'' However, in contrast to the Russian Roulette model described above, here the adaptive landscape is ''correlated'' in the sense that fitnesses of similar genotypes are similar which is more realistic.

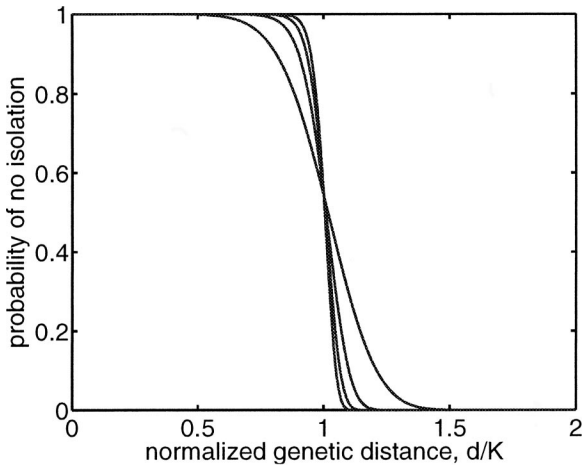


FIG. 6. The probability of no reproductive isolation with $C = 10$. Different lines correspond to $k = 2$ (the shallowest), 4, 6, 8 (the steepest).

Allopatric speciation

The simplest approach to model allopatric speciation is to assume that the populations accumulate substitutions at a constant rate, say ω substitutions per generation (Orr 1995; Orr and Orr 1996; Gavrilets 2000a; Orr and Turelli 2001). If ''on average'' K substitutions are required for complete reproductive isolation between two populations, the average waiting time to speciation is approximately

$$T = \frac{K}{2\omega} \tag{4}$$

generations, where the coefficient 2 is because there are two populations accumulating substitutions. Thus, the average waiting time to allopatric speciation can be predicted on the basis of K alone without the need to specify the fitness function $w(d)$ precisely. For example, if reproductive isolation results from C pairwise incompatibilities ($k=2$) , then equation (3) predicts that $K = \sqrt{2C/q}$, leading to the average waiting time to speciation

$$T = \frac{1}{\omega}\sqrt{\frac{C}{2q}}.$$

Orr (1995) and Orr and Turelli (2001) arrived to a similar expression using a different approach. How can one approximate the rate of substitutions ω? One way is to assume that genes underlying reproductive isolation have no other pleiotropic effects on fitness. In this case, genetic divergence will be driven exclusively by mutation and random drift (the null model of speciation). Let ν be the rate of neutral mutations per gamete per generation. If the within-population genetic variation is absent (or very low), each new mutation that is compatible with the current genetic state can be treated as selectively neutral. Then the rate of accumulation of substitutions can be written as

$$\omega = \nu$$

This approach predicts that the average waiting time to speciation does not depend on the population size either. However, in general the process of accumulation of incompatibilities cannot be treated as neutral even if the relevant genes do not control any other fitness components. This is because the alleles producing reproductive isolation are weakly selected against when rare (Nei et al. 1983; Gavrilets et al. 1998; Gavrilets 1999). A more careful analysis of the incompatibility model used here shows that large populations will actually evolve slower than small populations (Nei et al. 1983; Gavrilets et al. 1998, 2000; Gavrilets 1999). If alleles underlying reproductive isolation also increase fitness in the local conditions through some pleiotropic effects by a small amount s_a, then

$$\omega \approx \nu S_a,$$

where as before $S_a = 4Ns_a$. Strong selection for local adaptation will speed up both the genetic divergence and speciation.

One can also find the variances of these waiting times. If substitutions accumulate at a constant rate, then the coefficient of variation CV_T of the time to speciation is the same as the coefficient of variation of the number of substitutions necessary for complete reproductive isolation (see eq. 3b). That is,

$$CV_T \approx \frac{1}{k\sqrt{C}}. \tag{5}$$

Orr and Turelli (2001) arrived to a similar expression for $k = 2$ using a different approach. Interestingly, equations (3a, 4, 5) show that speciation times of the organisms characterized by large C are expected to have both larger averages and narrower relative ranges (Orr and Turelli 2001).

Parapatric speciation

To predict the dynamics of parapatric speciation one needs to specify function $w(d)$ precisely. A simple choice is the threshold function of reproductive compatibility

$$w(d) = \begin{cases} 1 & \text{for } d \le K, \\ 0 & \text{for } d > K \end{cases}$$

(Gavrilets et al. 1998; Gavrilets 1999). This function should be viewed as a limiting case of function $w(d)$ given by equation (2) when complete reproductive isolation requires many complex incompatibilities (that is, when k or C are large; see the brief discussion of the ''threshold effect'' below eq. 3). The neutral case (no reproductive isolation) corresponds to K equal to the number of loci. Some other choices of $w(d)$ were considered in Gavrilets (1999, 2000b).

Suppose there is a large number of loci and let ν be the probability of mutation per gamete per generation. As before, we assume that the population is subject to immigration at a constant rate m and that all immigrants have an ''ancestral'' genotype. Now speciation can be modeled as a random walk (Gavrilets 2000a) on the integers 0, 1, 2, . . . performed by the average pairwise distance d between the immigrants and the genotype most common in the population. Each new mutation fixed in the population will increase d by one. The fixation of an ancestral allele brought by the immigrants in a locus that has already diverged will decrease d by one. Speciation occurs when genetic distance d reaches the value of $K + 1$.

Neglecting within-population genetic variation, the average waiting time to speciation is approximately

$$T \approx \frac{1}{\nu}\left(\frac{m}{\nu}\right)^K K! \tag{6}$$

(Gavrilets 2000a). The average duration of speciation is approximately

$$\tau \approx \frac{1}{\nu}\left[1 + \frac{\Psi(K+1) + 0.577}{m/\nu}\right],$$

where the number 0.577 is Euler's constant and $\Psi(\cdot)$ is the psi (digamma) function (Gradshteyn and Ryzhik 1994). [Function $\Psi(K + 1)$ slowly increases with K and is equal to 0.42 at $K = 1$, to 2.35 at $K = 10$, and to 4.61 at $K = 100$.] For example, if $m = 0.01$, $\nu = 0.001$ and $K = 5$, then the waiting time to speciation is very long: $T = 1.35 \times 10^{10}$ generations, but if speciation does happen, its duration is relatively short: $\tau = 1236$ generations. Figure 7 illustrates the dependence of T and τ on model parameters in more detail. Notice that τ is of the order $1/\nu$ across a wide range of parameter values.

Direct selection for local adaptation.—Assume that each ''new'' allele improves adaptation to the local conditions. Let s_a be the average selective advantage of a new allele over the corresponding ancestral allele and $S_a = 4Ns_a$. The waiting time to speciation is approximately

$$T_s \approx T\frac{1}{S_a \exp(KS_a)},$$

where T is given by equation (6). The average duration of speciation is approximately

$$\tau_s \approx \frac{1}{\nu}\left[1 + \frac{\Psi(K+1) + 0.577}{(m/\nu)}\exp(S_a)\right]\frac{1}{S_a}.$$

For example, with the same values of parameters as above and $S_a = 2$, $T_s = 2.74 \times 10^4$ generations and $\tau_s = 2170$ generations. Thus, selection for local adaptation dramatically decreases T (in the numerical example, by the factor $\approx$ 50,000). Selection for local adaptation also somewhat increases τ relative to the case of speciation driven by mutation and genetic drift.

Effects of a genetic barrier.—Assume that alleles underlying reproductive isolation have no other pleiotropic effects on fitness but that the population has already diverged from the source population in some other loci underlying adaptation to local conditions. Let γ be the corresponding gene flow factor resulting from selection for local adaptation. In this case, the migration rate m has to be replaced by the effective migration rate $m_e = \gamma m$. The average waiting time to speciation can now be approximated as

$$T_\gamma \approx T\gamma^K,$$

where T is given by equation (6). Even a weak genetic barrier can significantly decrease the waiting time to speciation if there are many loci underlying reproductive isolation. For example, if $\gamma = 0.1$ and $K = 5$, then T_γ reduces by five orders

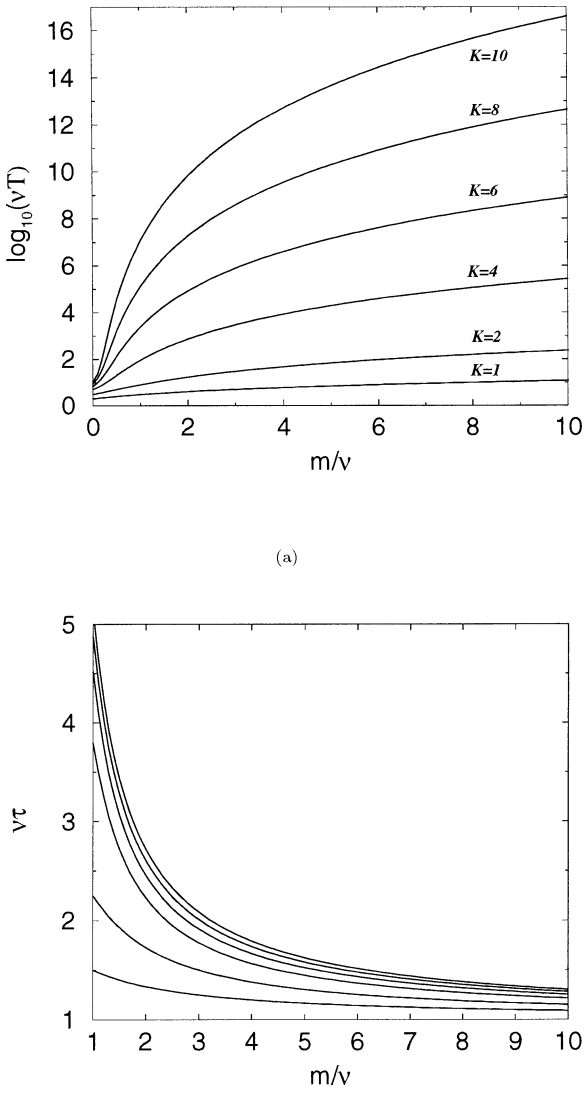


FIG. 7. The average waiting time to speciation, T, and the average duration of speciation, τ, in the case of speciation driven by mutation and random drift. The horizontal axes give the ratio of migration and mutation rates. In (a), the vertical axis gives the product of T and the mutation rate ν on the logarithmic scale. In (b), the vertical axis gives the product of τ and the mutation rate ν on the linear scale. In (b), the lines correspond to $K = 1,2,4,6,8$, and 10 (from bottom to top). [Using composite variables for the vertical axes allows one to represent relevant dependencies in two dimensions.]

of magnitude relative to the neutral case. The average duration of speciation is approximately

$$\tau_\gamma = \frac{1}{\nu}\left[1 + \frac{\Psi(K+1) + 0.577}{m/\nu}\frac{1}{\gamma}\right],$$

and is increased by the presence of the barrier.

It would be very important to know the variances of the waiting time to speciation and of the duration of speciation. Unfortunately, no analytical results comparable to expression (5) are known for the parapatric case.

Most species consist of geographically structured populations, some of which experience little genetic contact for long periods of time (Avise 2000). Different mutations are expected to appear first and increase in frequency in different populations necessarily resulting in some geographic differentiation even without any variation in local selection regimes. This genetic differentiation by mutation and drift alone can result in allopatric and parapatric speciation. However, in the parapatric case the waiting time to speciation is relatively short only if a very small number of genetic changes is sufficient for complete reproductive isolation. If more significant genetic change is necessary, then parapatric speciation without selection for local adaptation is hardly possible. The results presented above show that even relatively weak selection (acting directly on the loci underlying reproductive isolation or indirectly via genetic barrier) reduces the waiting time to speciation by orders of magnitude.

The waiting time to speciation, T, is very sensitive to parameters: changing a parameter by a small factor can increase or decrease T by several orders of magnitude. Most of the parameters of the model (such as the migration rate, intensity of selection for local adaptation, the population size, and, probably, the mutation rate) directly depend on the state of the environment (biotic and abiotic) the population experiences. This suggests that speciation is triggered by changes in the environment (Eldredge 2003). If it is a significant environmental change that initiates speciation, the populations of many different species inhabiting the same geographic area should all be affected in a similar way. In this case, one expects more or less synchronized bursts of speciation in a geographic area, that is, a ''turnover pulse'' (Vrba 1985).

The results about the duration of speciation τ lead to two important generalizations. The first is that the average duration of parapatric speciation, τ, is much smaller than the average waiting time to speciation, T. This feature of the models studied here is compatible with the patterns observed in the fossil record which form the empirical basis of the theory of punctuated equilibrium (Eldredge 1971; Eldredge and Gould 1972; Gould 2002). The second generalization concerns the absolute value of τ which is on the order of one over the mutation rate for a subset of the loci affecting reproductive isolation for a wide range of migration rates, population sizes, intensities of selection for local adaptation, and the number of genetic changes required for reproductive isolation. Given a ''typical'' mutation rate on the order of $10^{-5} - 10^{-6}$ per locus per generation (Griffiths et al. 1996; Futuyma 1998) and assuming that there are at least on the order of 10–100 genes involved in the initial stages of the evolution of reproductive isolation (Singh 1990; Wu and Palopoli 1994; Coyne and Orr 1998; Naveira and Masida 1998; Wu 2001), the duration of speciation is predicted to range between 10^3 and 10^5 generations.

Sympatric Speciation

The greatest share of modeling work has been on sympatric speciation. In spite of this and for the reasons discussed in the introduction, no clear and simple quantitative results are generally known. Here I consider two simple models of sympatric speciation for which conditions for sympatric speciation can be found analytically.

The first model, developed by Udovic (1980) describes a diploid population with alleles **A** and **a** and the first locus

and alleles **B** and **b** at the second locus. There are two common indices measuring the progress toward sympatric speciation in models of this kind. One is the *heterozygote deficiency index*

$$I = 1 - \frac{y}{2p(1-p)}$$

where y is the frequency of heterozygotes at a locus under consideration, p and $1 - p$ are the frequencies of the two alleles at the locus. The range of possible values of I is between 0 and 1. If the population is in Hardy-Weinberg proportions (which is expected if mating is random), $y = 2pq$ and $I = 0$. If the population has split into two homozygous groups that do not mate and hybrids (i.e., heterozygotes) are completely absent, then $y = 0$ and $I = 1$. Another index is the *normalized linkage disequilibrium, D′*, between the two loci which is defined as

$$D' = \frac{D}{\sqrt{p_A p_a p_B p_b}}.$$

Here D is the gametic linkage disequilibrium (defined as $D = x_{AB}x_{aB} - x_{Ab}x_{aB}$, where x_{AB}, x_{ab}, x_{Ab}, and x_{aB} are the frequencies of the corresponding gametes) and p_A, p_a, p_B and p_b are the corresponding allele frequencies. The range of possible values of D' is from -1 to $+1$. If alleles are distributed randomly between different gametes, then $D' = 0$. If the population has split into two genetic clusters so that the only gametes present are **AB** and **ab**, then $D' = 1$. If the only gametes present in the population are **Ab** and **aB**, then $D' = -1$.

In the second model, which deals with sexual conflict, individuals are haploid, and there are multiple alleles. The progress toward speciation in this model will be apparent from the formation of discrete genotypic clusters and from the degree of reproductive isolation between members of different clusters.

The Udovic model

Udovic's paper published in 1980 was the only one in almost a 40-year-long time interval that got very close to finding conditions for sympatric speciation analytically in a nontrivial model. However, both his approach and results were mainly ignored, being overshadowed by other papers describing numerical simulations. In this section, I describe and extend Udovic's ground-breaking results.

Consider a very large population with discrete nonoverlapping generations. There are two diallelic possibly linked loci with r being the recombination rate. The first locus with alleles **A** and **a** will be called the assortative mating locus or the AM locus. The AM locus controls assortative mating according to the symmetric version of the O'Donald model (O'Donald 1960). That is, each individual mates with probability α with another individual that has the same genotype at the AM-locus. With probability $1 - \alpha$ the individual is engaged in random mating. If $\alpha = 0$, the population is randomly mating. If $\alpha = 1$, the population is split into three reproductively isolated genotypic clusters corresponding to the three genotypes **AA**, **Aa**, and **aa**, of which the cluster of heterozygotes will gradually disappear. In the O'Donald model, if there are no other loci, the allele frequencies at the AM locus do not change whereas the genotype frequencies approach an equilibrium at which the index of heterozygote deficiency is

$$I_A = \frac{\alpha}{2 - \alpha} \tag{7}$$

(O'Donald 1960) The second locus with alleles **B** and **b** will be called the disruptive selection locus or the DS locus. The DS locus is subject to symmetric frequency-dependent disruptive selection favoring rare genotypes. This implies that if mating is random (that is, if $\alpha = 0$), the population evolves to a stable polymorphic equilibrium where the allele frequencies are $p_B = p_b = 1/2$ and the heterozygotes experience a relative fitness loss which we will denote as S ($0 \le S \le 1$). If $S = 1$, heterozygotes are inviable and (postmating) reproductive isolation between the two homozygotes is complete.

In this model there always exists a line of equilibria where the allele frequency at the AM locus can be arbitrary, the deficiency of heterozygotes at the AM locus is given by equation (7), the allele frequencies at the DS locus are at 1/2, the genotype frequencies at the DS locus are in Hardy-Weinberg proportions ($I_B = 0$) , and the linkage disequilibrium between the loci is absent ($D' = 0$). In words, on this line of equilibria the loci behave as completely independent. Generically, this line of equilibria is locally stable if disruptive selection and assortative mating are weak and linkage is loose.

Under certain conditions the line of equilibria becomes unstable. If this happens, the population evolves to one of the two alternative polymorphic equilibria at which the frequencies of all four alleles are equal to 1/2, heterozygotes are in deficiency and there is a statistical association between the alleles in different loci. The two equilibria differ in the sign of linkage disequilibrium D' and which one is approached by the population depends on initial conditions. At these equilibria, the population splits into two genetic clusters which arises within the population sympatrically and the loci mutually reinforce their effects leading to stronger reproductive isolation between the clusters. Evolution toward such an equilibrium represents (a step toward) sympatric speciation.

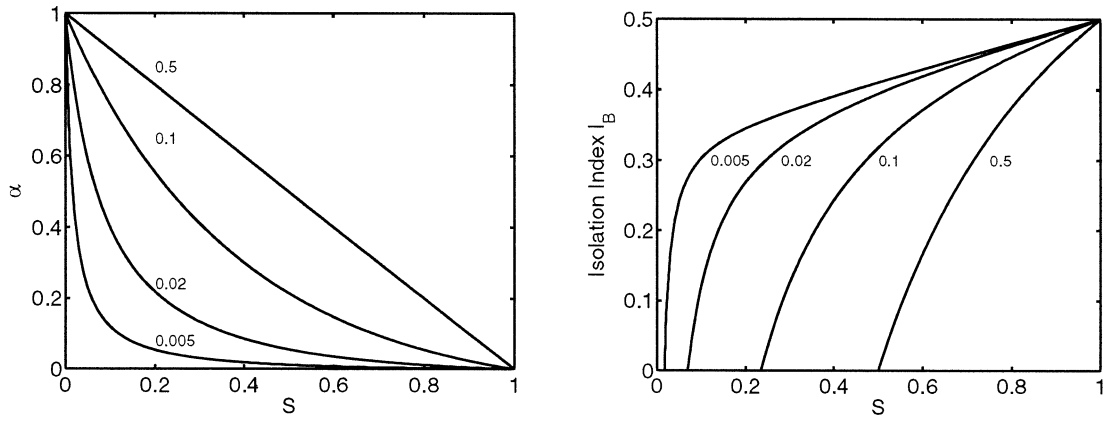
No linkage.—If the loci are unlinked ($r = 1/2$), sympatric speciation occurs if

$$\alpha + S > 1. \tag{8}$$

This condition has a very simple biological interpretation. Considering each locus in isolation, conditions $S = 1$ and $\alpha = 1$ imply complete postmating isolation and complete premating isolation, respectively. Thus, equation (8) tells us that sympatric speciation occurs only if the cumulative strength of selection against hybrids and the strength of assortative mating, $S + \alpha$, is larger than the threshold strength for each of them to cause complete reproductive isolation when acting in isolation.

If condition (8) is satisfied, then at equilibrium the index of heterozygotes deficiency in the DS locus is

$$I_B = \frac{S + \alpha - 1}{S}.$$



(a) (b)

FIG. 8. The Udovic model. (a) Conditions for sympatric speciation. Speciation occurs for parameter values above the corresponding curve. (b) Isolation index I_B for $\alpha = 0.50$. The recombination rates are given next to the corresponding curves.

Notice that the condition for sympatric speciation (8) coincides with the condition for I_B to be positive.

The index of heterozygote deficiency at the AM locus is

$$I_A = \frac{(S + 2\alpha)\alpha^2}{\alpha^2 S + \alpha^2 + \alpha S - S + 1}.$$

The normalized linkage disequilibrium is

$$D' = \mp \frac{\sqrt{I_A I_B}}{\alpha}.$$

Note that the maximum possible value for each of these three measures is observed at $S = 1$ and is equal to α.

Linkage.—With linkage (i.e., if $r < 1/2$), sympatric speciation occurs if

$$\alpha > \frac{2r(1 - S)(2 - S)}{S + 4r(1 - S)}. \quad (9)$$

The equilibrium values of I_A, I_B and D' are too cumbersome to be given here. Figure 8 illustrates the effects of parameters on the possibility of sympatric speciation and the degree of resulting genetic differentiation. Close linkage (i.e., small r) and strong selection are known to both make sympatric speciation easier and result in stronger genetic differentiation (Dickinson and Antonovics 1973; Udovic 1980; Felsenstein 1981). The inequality (9) and Figure 8 provide a quantitative description of these effects. I note that the equation for I_A was given in the appendix of Udovic's original paper. The fact that positivity of I_B is required for the stability of the equilibrium is also apparent from the numerical simulations reported by him.

As an example, let us consider the case of unlinked loci ($r = 1/2$) assuming linear frequency-dependent selection with fitnesses

$$W_{BB} = 1 + sp_b, \qquad W_{Bb} = 1, \qquad W_{bb} = 1 + sp_B,$$

where $s > 0$. This is one of the simplest cases of frequency-dependent selection (e.g., Gavrilets and Hastings 1995, 1998) which implies that rare alleles have a fitness advantage over common alleles. For example, genotypes using an underexplored resource can get a competitive advantage over other genotypes. In this model the equilibrium with allele frequencies $p_B = p_b = 1/2$ exists and is stable for any $s > 0$. At this equilibrium the relative fitness loss of heterozygotes is $S = s/(2 + s)$ and condition (8) for sympatric speciation is

$$\alpha > \frac{2}{2 + s}.$$

For example, if $s = 1$ (that is, selection is weak), then α has to be larger than 0.67 (that is, strong assortativeness in mating is required) whereas if $s = 8$ (that is, selection is strong), then α has to be larger than 0.20 (that is, weak assortativeness in mating is sufficient). Alternatively, if $\alpha = 0.2$ (that is, weak assortativeness in mating), then s has to be larger than 8 (that is, strong selection is required) and if $\alpha = 0.67$ (strong assortativeness in mating), then s has to be larger than 1 (i.e., weak selection is sufficient). Figure 9 illustrates the distributions of genotype frequencies for three different combinations of parameters obtained by numerical iterations of dynamic equations.

A somewhat inconspicuous feature of the Udovic model inherited from the O'Donald model of assortative mating is that organisms pay no costs for being ''choosy'' no matter how rare their preferred mates are. Under this condition the Udovic model shows that sympatric speciation is possible if disruptive selection and/or assortativeness in mating are strong enough. Closer linkage between the AM locus and the DS locus promote sympatric speciation.

Sexual conflict

In the model considered in the previous section sympatric speciation resulted as an outcome of disruptive natural selection emerging from ecological interactions between individuals. In this section I discuss a model where sympatric speciation is a consequence of sexual selection.

Consider a large sexual haploid population with distinct

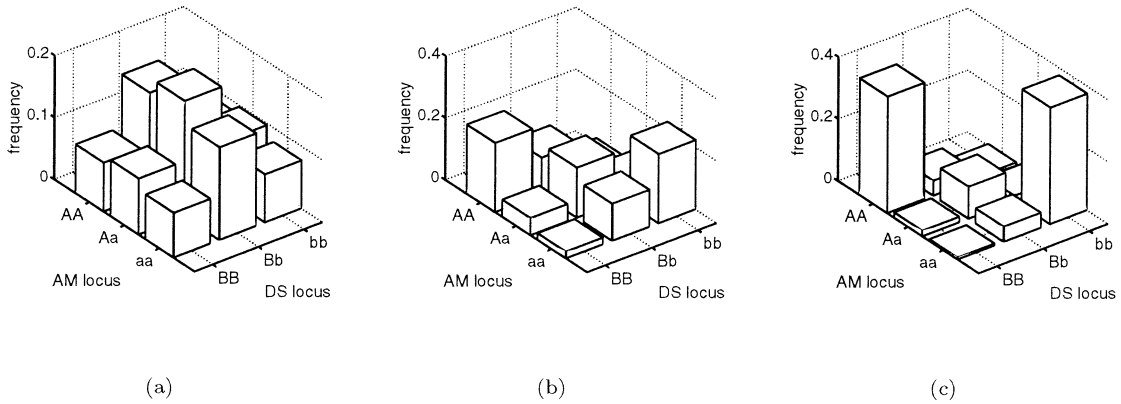


FIG. 9. The distributions of genotype frequencies after 200 generations starting with a random distribution. The loci are unlinked ($r = 1/2$) and $s = 2$, $S = 0.5$. The frequencies of double heterozygotes **AB**/**ab** and **Ab**/**aB** are pulled together. (a) $\alpha = 0.4$ so that condition (8) is not satisfied and no association of **A** and **B** evolves ($I_A = 0.25$, $I_B = 0$, $D' = 0$). (b) $\alpha = 0.6$ so that condition (8) is satisfied but the association of **A** and **B** is weak ($I_A = 0.46$, $I_B = 0.20$, $D' = 0.50$). (c) $\alpha = 0.8$ so that condition (8) is satisfied and the association of **A** and **B** is strong ($I_A = 0.72$, $I_B = 0.60$, $D' = 0.82$).

nonoverlapping generations. We concentrate on two possibly linked multiallelic loci assuming step-wise mutation. The alleles $\mathbf{A}_i$ at the first locus are only expressed in females (or eggs), and the alleles $\mathbf{B}_j$ at the second locus are only expressed in males (or sperm). The genotype space in this model is identical to that in the Russian Roulette model considered above.

I will say that two individuals are ''compatible'' if mating, fertilization, and offspring development are not prevented by isolating mechanisms. Assume that the probability Ψ_{ij} that a female (or egg) carrying an $\mathbf{A}_i$ allele is compatible with a male (or sperm) carrying a $\mathbf{B}_j$ allele is

$$\Psi_{ij} = \exp\left[-\frac{(i-j)^2}{2\sigma^2}\right].$$

This choice of function Ψ_{ij} formalizes a general observation that successful reproduction requires certain ''matching'' of male and female genes (Tregenza and Wedell 2000). Specifically, function Ψ_{ij} is maximized for all pairs of female and male genes with $i = j$. The coefficient σ controls the tolerance of matching: small σ implies that matching has to be pretty good for successful mating, and large σ implies that even when the value of $|i - j|$ is large, mating can happen easily.

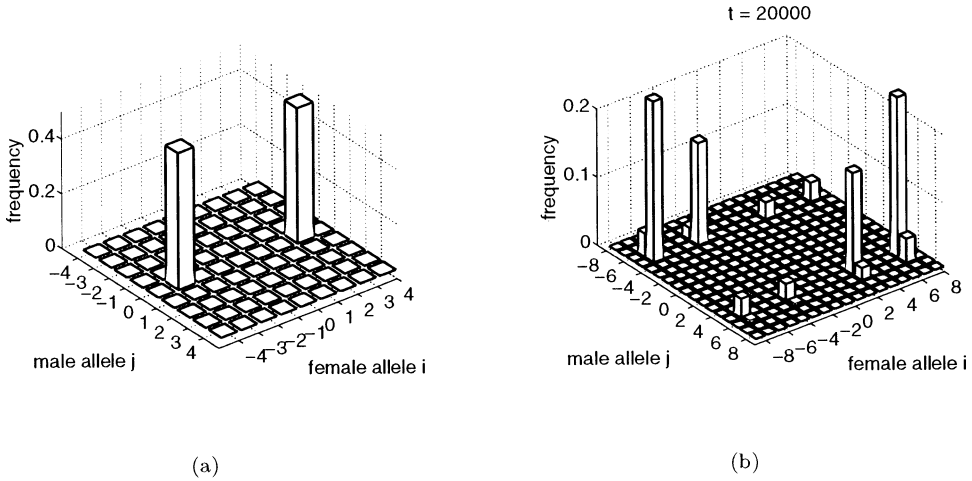
Let P_i be the proportion of the males in the population that are compatible with females carrying allele $\mathbf{A}_i$. For each female genotype the value of P_i can be found if one knows the frequencies of male genotypes in the population. P_i can also be thought of as a proxy of a female mating rate. We assume that males can be involved in multiple mating and that they compete for fertilization opportunities. Although males generally benefit from high mating rates, in many situations multiple matings reduce female viability and total fitness, that is, there is a conflict between the sexes with regard to the optimum mating rate (Arnqvist and Rowe 1995; Chapman and Partridge 1996; Rice 1996; Holland and Rice 1998; Parker and Partridge 1998). To formalize this fact we assume that the overall probability that an $\mathbf{A}_i$ female leaves offspring is a unimodal function of P_i that reaches a maximum at a certain value $P_{opt} < 1$ (Gavrilets 2000b; Gavrilets et al. 2001; Gavrilets and Waxman 2002). For example, in sea urchins, egg fitness is maximized at a level of sperm density which is much smaller than levels common under natural conditions (Franke et al. 2002). In our model, female mating rate is directly proportional to the proportion of compatible males and the assumption $P_{opt} < 1$ formalizes the idea of sexual conflict over mating rate because for the males, it is optimal to have $P_{opt} = 1$ since then all females are susceptible to fertilization by any male. To clarify the implications of the above assumptions, assume that the population is monomorphic for male allele $\mathbf{B}_j$. Then $P_i = \Psi_{ij}$ and the females that have the optimum mating rate and the highest overall fitness are those for which $\Psi_{ij} = P_{opt}$. Using the definition of Ψ it is easy to see that there are two such female alleles: $\mathbf{A}_{j+\delta}$ and $\mathbf{A}_{j-\delta}$ both at distance δ from the male allele, where

$$\delta = \sigma\sqrt{\ln(P_{opt}^{-2})}.$$

This simple model exhibits three general dynamic regimes (Gavrilets and Waxman 2002). The first regime is an *endless coevolutionary chase* between the sexes in which females continuously evolve to decrease the mating rate while males continuously evolve to increase it (Holland and Rice 1998; Gavrilets 2000b; Gavrilets et al. 2001; Gavrilets and Waxman 2002). In this regime, there is a dynamic compromise between the sexes, and the proportion of compatible pairs is intermediate between P_{opt} and 1. The coevolutionary chase is generically observed if the level of genetic variation is not too large.

Here we are concerned with two other regimes observed when the population size or mutation rates are sufficiently large. In the *Buridan's Ass regime* (Gavrilets and Waxman 2002) there is very low variation in male alleles maintained by mutation whereas female alleles split into two clusters both at the optimum distance δ from the male allele (see Fig. 10a). In this regime, males get trapped between the two female subclusters and have relatively low mating success.

In the *sympatric speciation regime* (Gavrilets and Waxman 2002) males answer the diversification of females by diversifying themselves and splitting into two clusters that start



(a) (b)

FIG. 10. Population genetic states in the sexual conflict model. Parameters: mutation rate $\mu = 10^{-5}$, recombination rate $r = 0.5$, $\sigma^2 = 10$, and female overall fitness $w_f = \exp(-(P - P_{opt})^2)$. (a) The Buridan's Ass regime ($P_{opt} = 0.6$; the average of value of P_i is 0.64; there is a single male allele $\mathbf{B}_0$ and two female alleles $\mathbf{A}_3$ and $\mathbf{A}_{-3}$); 5000 generations of selection. (b) The sympatric speciation regime ($P_{opt} = 0.4$; the average of values of P_i is 0.42); 8000 generations of selection.

evolving toward the corresponding female clusters. As a result, the initial population splits into different genetic clusters (species) that are reproductively isolated and which have emerged sympatrically (see Fig. 10b). The regime of coevolutionary chase within-species ends after increasing genetic variation in female alleles leads to the splitting of female alleles into two subclusters within each species. By contrast, genetic variation in male alleles remains very low within each species. At equilibrium, female P_i values are close to P_{opt} whereas males get trapped between two female subclusters and have low mating success.

The probability Ψ_{ij} can be written as a function $\Psi(d)$ of the ''genetic distance'' $d = i - j$ between the female and male alleles. In the limit of very low mutation rates, sympatric speciation occurs if

$$\Psi(\delta - 1) + \Psi(\delta + 1) > 2\Psi(\delta). \tag{10a}$$

If Ψ is a Gaussian function with zero mean and variance σ^2, this inequality can be rewritten as

$$\sigma < \delta \tag{10b}$$

(Gavrilets and Waxman 2002). If the above conditions are not satisfied, the population stays in the Buridan's Ass regime. Sympatric speciation requires small values of P_{opt} implying that sexual conflict over mating rates must be strong. For example, if $P_{opt} = 0.67$, then $\delta = 3$ and condition (10a) is not satisfied (see Fig. 10a). If $P_{opt} = 0.45$, then $\delta = 4$ and condition (10a) is satisfied (see Fig. 10b). Sufficiently small values of P_{opt} can result in more than two species emerging sympatrically (Gavrilets and Waxman 2002). The above results are not affected by the recombination rate between the loci.

In this model sympatric speciation requires sufficiently strong selection (i.e., small P_{opt}) and sufficiently strong assortativeness in mating (i.e., small σ). In contrast to the Udovic model, costs of being choosy are explicitly included in the sexual conflict scenario. That sympatric speciation still occurs is explained by the fact that the loci underlying reproductive isolation also experience direct selection for diversification induced by sexual conflict. In this model selection does not have to overcome the homogenizing effect of recombination that otherwise can prevent sympatric speciation (Udovic 1980; Felsenstein 1981; Rice 1984).

CONCLUSIONS

The theory of speciation needs many more simple and general analytical results allowing for transparent biological interpretation. In spite of the complexity of the processes leading to speciation, analytical approaches can be successful as demonstrated above. The major shift in the focus of theoretical studies needs to be from the demonstration that speciation is possible in a specific scenario (which has so far been the major goal of speciation models) to answering much more detailed questions about the probability of speciation, the waiting time to speciation, the duration of speciation, the degree of genetic and phenotypic divergence between sister species, the way different resources (including space) are partitioned between the sister species, etc. Several important generalizations about speciation have already emerged from analytical models.

(1) Unless the population size is small and the adaptive valley is shallow, the waiting time to a stochastic transition between the adaptive peaks is extremely long. This implies that it is very unlikely that a single peak shift will result in strong reproductive isolation. If transition does happen, it is very quick. This implies that observing it ''in action'' is practically impossible.

(2) Most species consist of geographically structured populations. Different mutations are expected to appear first and increase in frequency in different populations necessarily resulting in some geographic differentiation even without any variation in local selection regimes. Speciation can occur by mutation and random drift alone with no contribution from selection as different populations accumulate incompatible genes. This claim is based on models describing all three geographic modes (allopatric, parapatric, and sympatric). In the

same way as divergence by random drift and mutation represents a null model of molecular evolution, speciation by random drift and mutation represents a null model of speciation. (Note that Wright himself believed that ''the principal evolutionary mechanism in the origin of species must . . . be an essentially nonadaptive one'' [Wright 1932, p. 364]).

(3) The importance of mutations and drift in speciation is augmented by the general structure of adaptive landscapes. Organisms have thousands of genes and millions of nucleotides. Although divergence in a small number of genes leading to strong (or complete) reproductive isolation is possible, in most cases divergence will include multiple genes. Theoretical considerations of multidimensional genotype spaces have led to the realization that their properties are quite different from those of spaces of low dimensionality. In particular, the multidimensional genotype spaces are characterized by the existence of nearly neutral networks and holey adaptive landscapes. Speciation can be understood as the divergence along these networks (driven by mutation, drift, and selection for adaptation to a local biotic and/or abiotic environment) accompanied by the accumulation of reproductive isolation as a by-product.

(4) The waiting time to speciation driven by mutation and drift is typically very long. Migration can significantly delay speciation. However, selection for local adaptation (either acting directly on the loci underlying reproductive isolation via their pleiotropic effects or acting indirectly via establishing a genetic barrier to gene flow) can significantly decrease the waiting time to speciation. Direct selection is much more effective than indirect selection. In the parapatric case the average actual duration of speciation is much shorter than the average waiting time to speciation.

(5) Theoretical studies predict extreme sensitivity of the probability of speciation and the waiting time to speciation on model parameters which in turn strongly depend on the environmental conditions. This suggests that in general speciation is triggered by changes in the environment. Theoretical studies also show that once genetic changes underlying speciation start, they go to completion very rapidly. This is so both for changes driven by strong selection and for changes driven by weak stochastic factors. Thus, the quantitative theory predicts the short duration of intermediate stages in speciation and the difficulties of observing their traces in the fossil record.

(6) Sympatric speciation is possible if disruptive selection and/or assortativeness in mating are strong enough. Sympatric speciation is promoted if costs of being choosy are small (or absent) and if linkage between the loci controlling assortative mating and those experiencing disruptive selection is close.

These are but a few initial steps on a long road. A major future challenge for theoretical speciation research is to develop a comprehensive dynamical theory of speciation and to link microevolutionary processes with macroevolutionary patterns observed in the fossil record.

Acknowledgments

I am really grateful to N. H. Barton, C. R. B. Boake, M. Saum, and an anonymous reviewer for extremely helpful comments of the manuscript. Supported by National Institutes of Health grant GM56693 and by National Science Foundation grant DEB-0111613.

Literature Cited

Arnqvist, G., and L. Rowe. 1995. Sexual conflict and arms races between the sexes: a morphological adaptation for control of mating in a female insect. Proc. R. Soc. Lond. B 261:123–127.

Avise, J. C. 2000. Phylogeography. Harvard Univ. Press, Cambridge, MA.

Balkau, B. J., and M. W. Feldman. 1973. Selection for migration modification. Genetics 74:171–174.

Barton, N. H., and B. O. Bengtsson. 1986. The barrier to genetic exchange between hybridizing populations. Heredity 56: 357–376.

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

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

Bateson, W. 1909. Heredity and variation in modern lights. Pp. 85–101 *in* A. C. Seward, ed. Darwin and modern science. Cambridge Univ. Press, Cambridge, U.K.

Bazykin, A. D. 1965. On the possibility of sympatric species formation (in Russian). Byulleten' Moskovskogo Obshchestva Ispytateley Prirody. Otdel Biologicheskiy 70:161–165.

———. 1969. Hypothetical mechanism of speciation. Evolution 23: 685–687.

Bengtsson, B. O. 1985. The flow of genes through a genetic barrier. Pp. 31–42 *in* J. J. Greenwood, P. H. Harvey, and M. Slatkin, eds. Evolution Essays in Honor of John Maynard Smith. Cambridge Univ. Press, Cambridge, U.K.

Chapman, T., and L. Partridge. 1996. Female fitness in *Drosophila melanogaster*: An interaction between the effect of nutrition and of encounter rate with males. Proc. R. Soc. Lond. B 263: 755–759.

Coyne, J. A., and H. A. Orr. 1998. The evolutionary genetics of speciation. Philos. Trans. R. Soc. Lond. B 353:287–305.

Crosby, J. L. 1970. The evolution of genetic discontinuity: computer models of the selection of barriers to interbreeding between subspecies. Heredity 25:253–297.

Dickinson, H., and J. Antonovics. 1973. Theoretical considerations of sympatric divergence. Am. Nat. 107:256–274.

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

Edmands, S. 2002. Does parental divergence predict reproductive compatibility? Trends Ecol. Evol. 17:520–527.

Eldredge, N. 1971. The allopatric model and phylogeny in Paleozoic invertebrates. Evolution 25:156–167.

———. 2003. The sloshing bucket: how the physical realm controls evolution. Pp. 3–32 *in* J. Crutchfield and P. Schuster, eds. Towards a comprehensive dynamics of evolution - exploring the interplay of selection, neutrality, accident, and function. Oxford University Press, New York.

Eldredge, N., and S. J. Gould. 1972. Punctuated equilibria: an alternative to phyletic gradualis. Pp. 82–115 *in* T. J. Schopf, ed. Models in paleobiology. Freeman, Cooper, San Francisco.

Endler, J. A. 1977. Geographic variation, speciation, and clines. Princeton Univ. Press, Princeton, NJ.

Engels, F. 1940. Dialectics of nature. International Publishers, New York.

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

Felsenstein, J. 1981. Skepticism towards Santa Rosalia, or why are there so few kinds of animals? Evolution 35:124–138.

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

———. 1950. Gene frequencies in a cline determined by selection and diffusion. Biometrics 6:353–361.

Florin, A.-B., and A. Ödeen. 2002. Laboratory environments are not conducive for allopatric speciation. J. Evol. Biol. 15:10–19.

Franke, E. S., C. A. Styan, and R. Babcock. 2002. Sexual conflict

and polyspermy under sperm-limited conditions: in situ evidence from field simulations with the free-spawning marine echinoid *Evechinus chloroticus*. Am. Nat. 160:485–496.
Futuyma, D. J. 1998. Evolutionary biology. Sinauer, Sunderland, MA.
Gavrilets, S. 1997a. Hybrid zones with epistatic selection of Dobzhansky type. Evolution 51:1027–1035.
———. 1997b. Evolution and speciation on holey adaptive landscapes. Trends Ecol. Evol. 12:307–312.
———. 1999. A dynamical theory of speciation on holey adaptive landscapes. Am. Nat. 154:1–22.
———. 2000a. Waiting time to parapatic speciation. Proc. R. Soc. Lond. B 267:2483–2492.
———. 2000b. Rapid evolution of reproductive isolation driven by sexual conflict. Nature 403:886–889.
———. 2003. Evolution and speciation in a hyperspace: the roles of neutrality, selection, mutation, and random drift. Pp. 135–162 *in* J. Crutchfield and P. Schuster, eds. Towards a comprehensive dynamics of evolution - exploring the interplay of selection, neutrality, accident, and function. Oxford Univ. Press, New York.
Gavrilets, S., and M. B. Cruzan. 1998. Neutral gene flow across single locus clines. Evolution 52:1277–1284.
Gavrilets, S., and J. Gravner. 1997. Percolation on the fitness hypercube and the evolution of reproductive isolation. J. Theor. Biol. 184:51–64.
Gavrilets, S., and A. Hastings. 1995. Intermittency and transient chaos from simple frequency-dependent selection. Proc. R. Soc. Lond. B 261:233–238.
———. 1998. Coevolutionary chase in two-species systems with applications to mimicry. J. Theor. Biol. 191:415–427.
Gavrilets, S., and D. Waxman. 2002. Sympatric speciation by sexual conflict. Proc. Nat. Acad. Sci. USA 99:10533–10538.
Gavrilets, S., H. Li, and M. D. Vose. 1998. Rapid parapatric speciation on holey adaptive landscapes. Proc. R. Soc. Lond. B 265: 1483–1489.
——— 2000. Patterns of parapatric speciation. Evolution 54: 1126–1134.
Gavrilets, S., G. Arnqvist, and U. Friberg. 2001. The evolution of female mate choice by sexual conflict. Proc. R. Soc. Lond. B 268:531–539.
Gould, S. J. 2002. The structure of evolutionary theory. Harvard University Press, Cambridge, MA.
Gradshteyn, I. S., and I. M. Ryzhik. 1994. Tables of integrals, series, and products. Academic Press, San Diego, CA.
Griffiths, A. J. F., J. H. Miller, D. T. Suzuki, R. C. Lewontin, and W. M. Gelbart. 1996. An introduction to genetic analysis. W. H. Freeman, New York.
Grimmett, G. 1989. Percolation. Springer-Verlag, Berlin.
Haldane, J. B. S. 1931. A mathematical theory of natural and artificial selection. V. Metastable populations. Proc. Camb. Philos. Soc. 23:838–844.
———. 1948. The theory of a cline. J. Genet. 48:277–284.
Hegel, G. W. F. 1975. Hegel's logic: being part one of the encyclopaedia of the philosophical sciences (1830). Clarendon Press, Oxford, U. K.
Holland, B., and W. R. Rice. 1998. Chase-away sexual selection: antagonistic seduction versus resistance. Evolution 52:1–7.
Kimura, M. 1985. The role of compensatory neutral mutations in molecular evolution. J. Genet. 64:7–19.
Kirkpatrick, M., and V. Ravigné. 2002. Speciation by natural and sexual selection: models and experiments. Am. Nat. 159: S22–S35.
Kondrashov, A. S., and S. I. Mina. 1986. Sympatric speciation: When is it possible? Biol. J. Linn. Soc. 27:201–223.
Koopman, K. F. 1950. Natural selection for reproductive isolation between *Drosophila pseudoobscura* and *Drosophila persimilis*. Evolution 4:135–148.
Lande, R. 1979. Effective deme size during long-term evolution estimated from rates of chromosomal rearrangement. Evolution 33:234–251.
———. 1985. Expected time for random genetic drift of a population between stable phenotypic states. Proc. Nat. Acad. Sci. USA 82:7641–7645.
Maynard Smith, J. 1962. Disruptive selection, polymorphism and sympatric speciation. Nature 195:60–62.
———. 1966. Sympatric speciation. Am. Nat. 100:637–650.
Mayr, E. 1942. Systematics and the origin of species. Columbia Univ. Press, New York.
Muller, H. J. 1942. Isolating mechanisms, evolution, and temperature. Biol. Symp. 6:71–125.
Naveira, H. F., and Masida, X. R. 1998. The genetic of hybrid male sterility in *Drosophila*. Pp. 330–338 *in* D. J. Howard and S. H. Berrocher, eds. Endless forms: species and speciation. Oxford Univ. Press, New York.
Nei, M. 1976. Mathematical models of speciation and genetic distance. Pp. 723–768 *in* S. Karlin and E. Nevo, eds. Population genetics and ecology. Academic Press, New York.
Nei, M., T. Maruyama, and C.-I. Wu. 1983. Models of evolution of reproductive isolation. Genetics 103:557–579.
Ödeen, A., and A.-B. Florin. 2000. Effective population size may limit the power of laboratory experiments to demonstrate sympatric and parapatric speciation. Proc. R. Soc. Lond. B 267: 601–606.
———. 2002. Sexual selection and peripatric speciation: the Kaneshiro model revisited. J. Evol. Biol. 15:301–306.
O'Donald, P. 1960. Assortative mating in a population in which two alleles are segregating. Heredity 15:389–396.
Ogden, R., and R. S. Thorpe. 2002. Molecular evidence for ecological speciation in tropical habitats. Proc. Nat. Acad. Sci. USA 99:13612–13615.
Orr, H. A. 1995. The population genetics of speciation: the evolution of hybrid incompatibilities. Genetics 139:1803–1813.
Orr, H. A., and L. H. Orr. 1996. Waiting for speciation: the effect of population subdivision on the waiting time to speciation. Evolution 50:1742–1749.
Orr, H. A., and D. C. Presgraves. 2000. Speciation by postzygotic isolation: forces, genes, and molecules. BioEssays 22: 1085–1094.
Orr, H. A., and M. Turelli. 2001. The evolution of postzygotic isolation: accumulating Dobzhansky-Muller incompatibilities. Evolution 55:1085–1094.
Parker, G. A., and L. Partridge. 1998. Sexual conflict and speciation. Philos. Trans. R. Soc. Lond. B 353:261–274.
Rice, W. R. 1984. Disruptive selection on habitat preferences and the evolution of reproductive isolation. Evolution 38: 1251–1260.
———. 1996. Sexually antagonistic male adaptation triggered by experimental arrest of female evolution. Nature 381:232–234.
Rice, W. R., and Hostert, E. E. 1993. Laboratory experiments on speciation: what have we learned in 40 years? Evolution 47: 1637–1653.
Schluter, D. 2000. The ecology of adaptive radiation. Oxford Univ. Press, Oxford, U.K.
Schneider, C. J., T. B. Smith, B. Larison, and C. Moritz. 1999. A test of alternative models of diversification in tropical rainforests: ecological gradients vs. rainforest refugia. Proc. Nat. Acad. Sci. USA 96:13869–13873.
Singh, R. S. 1990. Patterns of species divergence and genetic theories of speciation. Pp. 231–265 *in* K. Wöhrmann and C. K. Jain, eds. Population biology ecological and evolutionary viewpoints. Springer-Verlag, Berlin.
Smith, H. M. 1955. The perspective of species. Turtox News 33: 74–77.
———. 1969. Parapatry: sympatry or allopatry? Systematic Zoology 18:254–255.
Smith, T. B., R. K. Wayne, D. J. Girman, and M. W. Bruford. 1997. A role for ecotones in generating rainforest biodiversity. Science 276:1855–1857.
Tregenza, T., and N. Wedell. 2000. Genetic compatibility, mate choice and patterns of parentage. Mol. Ecol. 9:1013–1027.
Turelli, M., N. H. Barton, and J. A. Coyne. 2001. Theory and speciation. Trends Ecol. Evol. 16:330–343.
Udovic, D. 1980. Frequency-dependent selection, disruptive selec-

tion, and the evolution of reproductive isolation. Am. Nat. 116: 621–641.
Via, S. 2001. Sympatric speciation in animals: the ugly duckling grows up. Trends Ecol. Evol. 16:381–390.
Vrba, E. S. 1985. Environment and evolution: alternative causes of the temporal distribution of evolutionary events. S. Afr. J. Sci. 81:229–236.
Walsh, J. B. 1982. Rate of accumulation of reproductive isolation by chromosome rearrangements. Am. Nat. 120:510–532.
Wilson, R. A. 1999. Species. New interdisciplinary essays. MIT Press, Cambridge, MA.
Wright, S. 1921. Systems of mating. III. Assortative mating based on somatic resemblance. Genetics 6:144–161.
———. 1931. Evolution in Mendelian populations. Genetics 16: 97–159.
———. 1932. The roles of mutation, inbreeding, crossbreeding and selection in evolution. Pp. 356–366 *in* Proceedings of the Sixth International Congress on Genetics. Vol 1. Brooklyn Botanic Garden, Brooklyn, NY.
———. 1941. On the probability of fixation of reciprocal translocations. Am. Nat. 75:513–522.
———. 1982. The shifting balance theory and macroevolution. Annu. Rev. Genet. 16:1–19.
Wu, C.-I. 2001. The genic view of the process of speciation. J. Evol. Biol. 14:851–865.
Wu, C.-I., and M. F. Palopoli. 1994. Genetics of postmating reproductive isolation in animals. Annu. Rev. Genet. 27:283–308.

Corresponding Editor: J. Mitton

Appendix

Notations used herein: The first part of the appendix contains variables used throughout the paper; the following five parts describe variables unique to the five major sections of the paper.

$\mathbf{A}$, $\mathbf{B}$, $\mathbf{a}$, $\mathbf{b}$, $\mathbf{A}_i$, $\mathbf{B}_i$	Alleles at different loci
T, τ	Average waiting time to speciation and average duration of speciation
μ, ν	The rate of mutation per locus and per gamete
r, m, N, w	Recombination rate, migration rate, population size, fitness
s, S, S_a $(=4Ns)$	Different selection coefficients
z, G	Average and variance of an additive quantitative trait
γ	The gene flow factor
a, b	Coefficients characterizing reduction in fertility
$\mathcal{P}$, $\mathcal{P}_c$	Probability of being viable and the threshold value of $\mathcal{P}$
$\mathcal{L}$, $\mathcal{A}$	Number of loci and number of alleles
d	Number of diverged loci
k, q	Number of alleles in an incompatible set and probability of incompatibility
C	Number of incompatibilities necessary for speciation
J	Expected number of incompatibilities
K, CV_{K}	Average and coefficient of variation of the number of substitutions required for speciation
ω	Rate of substitutions
CV_{T}	Coefficient of variation of the time to speciation
I, I_{A}, I_{B}	Heterozygote deficiency indices
p, p_A, p_a, p_b, p_B	Allele frequencies
y	Frequency of heterozygotes
D, D'	Linkage disequilibrium and normalized linkage disequilibrium
α	Probability of assortative mating
Ψ_{ij}	Preference function
P_i	Proportion of compatible males maximizing female fitness $\mathbf{A}_i$
P_{opt}	Optimum proportion of males compatible with a female
δ	Genetic distance resulting in P_{opt}