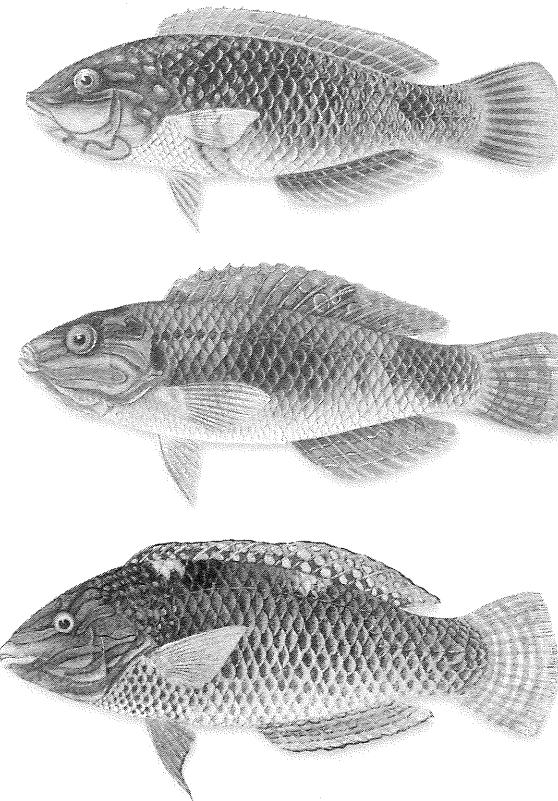


SPECIATION



JERRY A. COYNE

University of Chicago

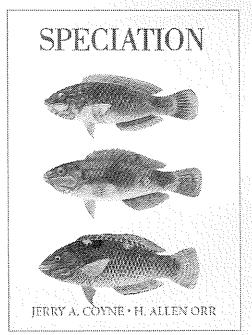
H. ALLEN ORR

University of Rochester

**UNIVERSITY OF NORTHERN
BRITISH COLUMBIA
LIBRARY
Prince George, BC**



Sinauer Associates, Inc. • Publishers
Sunderland, Massachusetts U.S.A.



About the cover

Three congeneric species of Pacific wrasse; from top to bottom, *Halichoeres trimaculatus*, *H. margaritaceus*, and *H. hortulanus*. These paintings, by the Japanese artist Kako Morita, are reproduced from Plates 46 and 47 of *The Fishes of Samoa* by David Starr Jordan and Alvin Seale (1906, Bulletin of the United States Bureau of Fisheries 25:173–456). The illustrations were published with the help of Jordan's friend and fellow naturalist, President Theodore Roosevelt, who interceded when the government's committee on publication deemed the plates too expensive to print. Jordan (1851–1931) was an influential evolutionist, ichthyologist, and a staunch defender of Darwinism at a time

when it was unpopular. A prolific author, his most notable contribution to the study of speciation was his emphasis on the importance of geographic barriers.

SPECIATION

Copyright © 2004 by SINAUER ASSOCIATES, Inc. All rights reserved. This book may not be reprinted in whole or in part without permission from the publisher.

For information or to order, address:

SINAUER ASSOCIATES, Inc.
23 Plumtree Road/PO Box 407
Sunderland, MA 01375 U.S.A.

FAX: 413-549-1118

Email: publish@sinauer.com

www.sinauer.com

Library of Congress Cataloging-in-Publication Data

Coyne, Jerry A.

Speciation / Jerry A. Coyne, H. Allen Orr.

p. cm.

Includes bibliographical references and index.

ISBN 0-87893-091-4 (hardcover) -- ISBN 0-87893-089-2 (paperback)

1. Species. I. Orr, H. Allen. II. Title.

QH380.C68 2004

576.8'6--dc22

2004009505

To Anne and Lynne

Reinforcement

The theory of reinforcement—the enhancement of prezygotic isolation in sympatry by natural selection—has had an extraordinarily tortuous history. Even the origin of the idea is more complex than generally appreciated. Although A. R. Wallace, the codiscoverer of natural selection, is typically given credit for the idea—reinforcement is still sometimes called the “Wallace effect”—Wallace’s hypothesis differs from the modern one. Wallace argued for the enhancement of *postzygotic* isolation by *group* selection (Littlejohn 1981; Howard 1993), a scenario that has little to do with the modern theory of reinforcement. The modern theory instead dates to Dobzhansky (1937b), who described the following scenario: two taxa diverge in allopatry; upon secondary geographic contact, hybridization occurs at some rate, yielding unfit hybrids; because production of hybrids is maladaptive, individuals who mate only with their own taxon enjoy a fitness advantage; natural selection thus favors the evolution of enhanced prezygotic isolation.

Enthusiasm for this new idea quickly outstripped all evidence for it and, by 1940, Dobzhansky went so far as to suggest that even *allopatric* divergence in mate preference may sometimes reflect the evolutionary consequences of rare migrants from another taxon. Dobzhansky also came to believe that reinforcement (a term coined by Blair 1955) represents a nearly obligate last step in speciation, a view that remained popular into the 1970s (e.g., Lewontin 1974). Indeed, Dobzhansky ultimately concluded not only that prezygotic isolation evolves as a direct product of natural selection, but that it is an ad hoc contrivance designed to protect the integrity of species, a remarkably teleological view (e.g., Dobzhansky 1951, p. 208).

In the 1980s, however, the popularity of reinforcement plummeted. As Noor (1999) and Marshall et al. (2002) note, this change had little to do with new data and everything to do with new theoretical objections and verbal arguments.

While we review many of these objections and arguments below, it is worth clarifying one point at the outset. Butlin (1987a) emphasized that cases in which taxa produce completely unfit hybrids must be distinguished from those in which taxa produce partially unfit hybrids. In the first case, no gene flow is possible and the taxa are already good species; in the second case, gene flow is possible and the taxa are not yet good species. Butlin argued that enhancement of prezygotic isolation in the first case has nothing to do with speciation and should be kept distinct from reinforcement proper. He suggested the term “reproductive character displacement” for an increase in isolation between taxa that are already good species. True reinforcement is restricted to cases in which isolation is enhanced between taxa that can still exchange genes; reinforcement might therefore complete speciation. We will abide by Butlin’s distinction: we are interested in speciation, not in changes that occur afterwards. This does not mean, however, that we can learn nothing about reinforcement from reproductive character displacement. On the contrary, we will see that both the experimental and theoretical study of reinforcement can benefit from analysis of cases in which hybrids are completely unfit.

By about 1990, reinforcement once again became popular, reflecting the emergence of new data revealing that *something* interesting was happening in sympatry and that, at the least, this something resembled reinforcement. On the heels of these data, many theorists found that reinforcement was possible after all. Indeed, several found that it was easy.

Although we have played some part in the revival of enthusiasm for reinforcement, we are not uncritical champions of its importance. We believe the present data and theory show that reinforcement is possible—and must be taken seriously—but they do not show that reinforcement is common, much less ubiquitous (Marshall et al. 2002; Servedio and Noor 2003). Here, we critically review these data and theory.

Our approach is somewhat unusual. Inverting our usual order, we consider the data before the theory. There are several reasons for this. First, this order better reflects the history of research on reinforcement. Second, it is important to establish if patterns *consistent* with reinforcement are common in nature before worrying about whether these patterns are best explained by reinforcement or by many alternative hypotheses. Our presentation thus proceeds in several steps: (1) we ask if there is evidence consistent with an increase of prezygotic isolation in sympatry; (2) finding that there is, we ask if reinforcement provides a plausible explanation; (3) we review alternative hypotheses that might also explain the data; and (4) finally, we suggest a test to empirically distinguish reinforcement from these alternatives.

The Data

Data testing reinforcement come in two forms: selection experiments and observations from nature (Rice and Hostert 1993).

Selection experiments

Before considering the frequency of reinforcement in nature, it is worth knowing if reinforcement can happen under *any* circumstances. The most direct way to find the answer is clear: one must attempt to reproduce the phenomenon in the laboratory. One should be able to show that enhanced prezygotic isolation evolves in response to *artificial* selection against hybrids in the laboratory, just as it allegedly does to natural selection in the wild. As Rice and Hostert (1993) emphasize in their review of such experiments, these efforts fall into two classes. The first involves “destroy-the-hybrids” designs, and the other involves disruptive selection on arbitrary traits. For the moment, we include experiments in which hybrids are completely unfit (reproductive character displacement, not reinforcement). If no response to selection occurs in this extreme case, we need not concern ourselves with less extreme, but more biologically relevant, cases.

The first destroy-the-hybrid experiment was that of Koopman (1950), who simulated complete hybrid inviability between two species of *Drosophila*, asking if production of unfit hybrids leads to the evolution of enhanced prezygotic isolation. Koopman studied replicate population cages containing both *D. pseudoobscura* and *D. persimilis*, each of which carried a different recessive eye color mutation. This allowed him to distinguish the progeny of conspecific versus heterospecific matings by eye color. Each generation, Koopman discarded all offspring of interspecific couplings, breeding only from the progeny of conspecific couplings. Enhanced prezygotic isolation evolved remarkably quickly: in the first generation between 22% and 50% of all progeny were hybrid, but within six generations the figure dropped to less than 5%.

Most destroy-the-hybrid experiments show at least some enhancement of prezygotic isolation, albeit often weaker and more slowly evolving than in Koopman’s study. A good example is that of Knight et al. (1956) who successfully selected for sexual isolation within *D. melanogaster* over nearly 40 generations. Similarly, Crossly (1974) successfully selected for increased sexual isolation between two mutant strains of *D. melanogaster*. Such experiments are not limited to *Drosophila*. Indeed, one of the most impressive involved selection for prezygotic isolation between two strains of maize, *Zea mays* (Paterniani 1969). (Plants provide excellent material for such experiments. Unlike many animals where a female mates with one or a few males, plants may receive pollen from many individuals. By scoring the frequency of hybrid kernels in maize, one obtains a quantitative measure of prezygotic isolation per individual.) After sowing non-hybrid kernels from individuals showing low rates of intercrossing, Paterniani saw a remarkable 30%–40% decrease in the frequency of intercrossing in six generations. Rice and Hostert (1993) review similar experiments.

Despite this record, destroy-the-hybrids experiments have been criticized on several grounds. First, there is a suspicion that their success rate reflects a publication bias. Responses to selection for greater isolation are exciting and are published, while failed responses are uninteresting and may go unpub-

lished. Although there are some exceptions to this pattern (see Robertson 1966, and Fukatami and Moriwaki 1970 for negative, or at least non-significant, results), it does seem likely that failed responses are underreported. (Indeed one of our own early experiments involved a failed—and unpublished—attempt to replicate Koopman's results in the wasp *Nasonia*.) Second, as noted, these experiments test reproductive character displacement, not reinforcement. One cannot, then, necessarily extrapolate from these findings to true reinforcement.

The obvious experimental modification—a destroy-the-hybrids design in which a few hybrids are allowed to breed—has been attempted, so far as we know, only once. Harper and Lambert (1983) selected against hybridization between two lines of *D. melanogaster*, allowing variable degrees of gene flow. Though no response to selection was seen, this may have reflected a lack of genetic variation in their marker stocks, which are usually highly inbred.

The closest analogue to modified destroy-the-hybrid experiments involves allopatric populations that are first adapted to different environments (e.g., food with EDTA versus food without EDTA) and then brought into contact (e.g., Robertson 1966a; Wallace 1982; Ehrman et al. 1991). If population hybrids suffer intermediate phenotypes and so are poorly adapted to available environments (extrinsic postzygotic isolation), increased prezygotic isolation should be favored. Although such experiments allow gene flow, they probably allow far too much of it: it seems unlikely that, after a brief period of adaptation in the laboratory, most hybrids would die. In any case, these experiments are more complex than needed to address the above concern. What we want to know is if response to selection disappears when going from no gene flow to low gene flow in Koopman-like experiments.

In the absence of such work, attention shifted to a second class of selection experiment—that in which an arbitrary character experiences disruptive selection in a single population. Although such experiments are usually considered tests of sympatric speciation, they bear on reinforcement: they ask if assortative mating evolves when intermediate “hybrid” phenotypes are selected against. The first and most celebrated of these experiments was that of Thoday and Gibson (1962). Because we described this experiment in Chapter 4, we only sketch it here. Thoday and Gibson performed disruptive selection on bristle number in *D. melanogaster*. Virgin flies with the highest and lowest bristle number were selected from a base population; they were then placed together and allowed to mate freely; high and low bristle females were separated and allowed to produce progeny. This was continued for 12 generations. Although this design allows random mating among flies, Thoday and Gibson saw a rapid response to selection: by the end of the study, high bristle females produced almost all high bristle progeny and low bristle females almost all low bristle progeny. Natural selection appeared to promote the evolution of assortative mating.

Given their dramatic results, Thoday and Gibson's work spawned a minor industry. Unfortunately, this work arrived at a consistently depressing result:

every attempt to replicate Thoday and Gibson's findings with new stocks failed (reviewed in Thoday and Gibson 1970; Scharloo 1971; for discussion of a possible exception, see Rice and Hostert 1993). In retrospect, the reason seems clear. As Felsenstein (1981) explained, the evolution of prezygotic isolation by disruptive selection requires that a population leap an especially high population-genetic hurdle. Reinforcement requires the establishment and maintenance of strong non-random associations (linkage disequilibria) between the alleles underlying the disruptively selected character and those underlying assortative mating. Although disruptive selection automatically generates such disequilibrium, recombination destroys it and—unless the two kinds of loci are tightly linked—recombination wins. We discuss Felsenstein's insight in more detail below. For now, it suffices to realize that there are good reasons why disruptive selection experiments *should fail*.

Another variety of disruptive selection experiment has been more successful. The key insight was provided by Slatkin (1982), who noted that disruptive selection on a single character might *automatically* yield prezygotic isolation. Slatkin gave the example of flowering time. If selection favors early or late flowering, reproductive isolation between subpopulations arises as an automatic consequence of response to selection, as early breeders will not hybridize with late breeders. Formally, this idea sidesteps Felsenstein's objection by allowing the rate of recombination between the genes underlying the character under selection and those underlying assortative mating to go to zero—they are the *same* genes. Rice and colleagues have been vocal champions of this idea. Most important, they have demonstrated the efficacy of the process in large disruptive selection experiments on habitat preference in *Drosophila* (reviewed in Chapter 4).

While such findings prove the efficacy of the single-character scenario—and have clear implications for sympatric speciation—the connection between disruptive selection and reinforcement has grown tenuous as we moved from two-character (Thoday-and-Gibson-like) to one-character (Rice-like) experiments. In the first case, one asks for the evolution of prezygotic isolation *in order* to prevent production of unfit hybrids. One asks, that is, for reinforcement. In the second case, prezygotic isolation is fortuitous—one disruptively selects on a character that just *happens* to cause prezygotic isolation. This is not, we would argue, reinforcement.

In summary, while selection experiments provide some support for the idea that prezygotic isolation can be enhanced by natural selection, those experiments that work best are only loosely connected to reinforcement, while those that work worst are more closely connected to reinforcement. More encouraging results emerge from the second kind of evidence, that from nature.

Evidence from nature: case studies

Almost all work on reinforcement in nature takes the same form: a contrast between the strength of prezygotic isolation in sympatry versus allopatry. In some cases the taxa compared are species while in others they are populations.

Such data can of course only detect a pattern, not identify a process. Several processes other than reinforcement may produce a pattern of increased prezygotic isolation in sympatry. To be clear, then, we will refer to a *pattern* of stronger prezygotic isolation in sympatry as “enhanced isolation,” implying nothing about its evolutionary cause.

The earliest data bearing on enhanced isolation come from Dobzhansky and Koller’s (1938) study of sexual isolation between *D. miranda* and *D. pseudoobscura*. (Two “races” of *D. pseudoobscura*, A and B, were used; race B was later renamed *D. persimilis*.) Early data suggested that *D. miranda* was geographically concentrated in the northwestern United States, near Puget Sound, while *D. pseudoobscura* had a broader distribution in the western United States, overlapping with that of *D. miranda*. Dobzhansky and Koller found that, as one moved from populations of *D. pseudoobscura* (race A or B) far from *D. miranda*’s range, to those closer to (and ultimately sympatric with) *D. miranda*, sexual isolation between the species increased. This discovery had a profound effect on Dobzhansky’s view of speciation. Indeed, Howard (1993) dates Dobzhansky’s conversion to reinforcement to this finding. Not surprisingly, the *D. pseudoobscura*–*D. miranda* story played a prominent part in Dobzhansky’s (1937b) discussion of reinforcement in the first edition of *Genetics and the Origin of Species*, where many evolutionists first learned of reinforcement. (As the dates indicate, Dobzhansky and Koller’s 1938 paper was not yet published at the time of Dobzhansky’s book, but much of the data were in hand.)

A closer reading of the 1938 paper, however, yields a more complex story. The complication springs from the fact that—after the publication of the first edition of *Genetics and the Origin of Species*—Dobzhansky and Koller found that *D. miranda* was not restricted to Puget Sound. Instead, they report the discovery in 1937 of populations of *D. miranda* in the Sierra Nevada Mountains of California. As they note, this finding casts doubt on Dobzhansky’s earlier interpretation. Dobzhansky and Koller tested the strength of sexual isolation between Sierra Nevada populations of *D. miranda* and *D. pseudoobscura*. Surprisingly, these sympatric flies crossed fairly freely—as freely as any allopatric combination—a result, they conclude, that is “the reverse of what one might have expected by analogy with the behavior of the different strains of *D. pseudoobscura* toward the [Puget Sound] race of *D. miranda*.” Curiously, this finding seems to have had little effect on Dobzhansky’s enthusiasm for reinforcement. While the second (1941) edition of his book includes a description of the troubling Sierra Nevada findings (p. 266), Dobzhansky still concludes that selection plays a likely role in the origin of isolation. Remarkably, the Sierra Nevada finding disappears from the third (1951) edition, with Dobzhansky merely reporting that *D. pseudoobscura* flies from regions close to *D. miranda* “show, in general, greater sexual isolation than do strains from distant regions” (p. 210). It is hard to see how this claim can be reconciled with the data.

Over the years, other reports of enhanced isolation appeared sporadically in the *Drosophila* literature. Ehrman (1965), for example, compared the strength of sexual isolation between sympatric and allopatric populations of several

semispecies of *D. paulistorum*, finding greater isolation in sympatry in most cross combinations. Similarly, Wasserman and Koepfer (1977) found enhanced isolation between *D. mojavensis* and *D. arizonensis* (later renamed *D. arizonae*), two species that live in the American Southwest and that have partly overlapping distributions. One of the most recent reports of enhanced isolation involves *D. pseudoobscura* and *D. persimilis*. *D. pseudoobscura* is widely distributed throughout western North America while *D. persimilis* is found only in Pacific coastal states; the latter species is thus completely embedded within the distribution of the former. Because hybrids sometimes form in nature (albeit rarely) and hybrid males are sterile, selection surely punishes hybridization. One would thus predict that *D. pseudoobscura* females from regions sympatric with *D. persimilis* would discriminate against *D. persimilis* males more than do *D. pseudoobscura* females from regions of allopatry. This was precisely the pattern seen by Noor (1995): in five of six contrasts, sympatric females are choosier than allopatric females.

Reports of enhanced isolation have not been limited to *Drosophila*. In the damselfly *Calopteryx*, for instance, Waage (1975, 1979) showed that males who derive from regions of geographic overlap between two species distinguish females of the two species better than do males from non-overlapping regions; interestingly, females of these species differ in wing coloration in sympatry but not in allopatry. Similarly, songs of Hawaiian crickets of the genus *Laupala* differ more among sympatric populations than expected by chance (i.e., than if songs were randomly assigned without regard to geography) (Otte 1989). Also, the walking stick *Timema cristinae* shows greater prezygotic isolation between populations that exchange migrants at intermediate rates. Nosil et al. (2003) argue that this pattern reflects reinforcement, which should occur neither under negligible migration rates (as there is insufficient selection against hybridization), nor under high migration rates (as response to selection is swamped by gene flow; see also Cooley et al. 2001 on the periodical cicadas *Magicicada*.)

Beyond insects, some of the best-known cases of enhanced isolation involve gametic isolation in marine organisms. We have already described these cases in Chapter 6—they include the abalone *Haliotis* (Lee et al. 1995), the sea urchin *Arbacia* (Metz et al. 1998b), the sea urchin *Echinometra* (Geyer and Palumbi 2003), and the mussel *Mytilus* (Springer and Crespi 2004; see also Wullschleger et al. 2002 for results from snails).

Frogs have provided particularly good material for studies of enhanced isolation as the primary mechanism of prezygotic isolation is both known and easily studied: male calls attract females and, in a number of frog families, females favor conspecific over heterospecific calls (Blair 1974). A pattern of enhanced isolation is fairly common. In *Gastrophryne olivacea* and *G. carolinensis*, for example, calls of males from allopatric populations are similar, while those from sympatric populations differ in both duration and frequency. One of the best-studied cases involves the Australian tree frog *Hyla*. Littlejohn (1965) showed that calls of distant allopatric populations of *H. ewingi* and *H. verreauxi* are similar, while those of sympatric populations are different. Subsequent

work revealed that, when sympatric females of each species are given a choice of sympatric males from both species, they invariably "choose" conspecific males (Littlejohn and Loftus-Hills 1968). Two other species of tree frog, *Hyla cinerea* and *H. gratiosa*, also show a pattern of enhanced isolation: female preference for conspecific male calls in *H. cinerea* is greater in sympatric than allopatric populations (Hobel and Gerhardt 2003; see also Hillis 1981 and Chapter 5 for a possible case of reinforcement in the leopard frog *Rana* involving temporal isolation; and see Pfennig 2003 for a possible case of reinforcement in the spadefoot toad *Spea*). In a review of the older literature, Blair (1974) emphasized that no sympatric species are known that have the same call. He concluded that reinforcement plays a probable role in the origin of frog species.

Similar cases occur in fish. Rundle and Schlüter (1998), for example, found enhanced isolation in the threespine stickleback *Gasterosteus aculeatus*. Benthic and limnetic morphs co-occur in several coastal lakes in British Columbia; in other lakes, only single forms are found. Rundle and Schlüter found that benthic females from lakes that include limnetics discriminate against limnetic males, while females from lakes that do not include limnetics, do not. This result is particularly interesting as the sympatric morphs cross at a low rate in nature and mtDNA data reveal past gene flow as well. Moreover, the resulting hybrids are unfit under natural conditions (Chapter 7).

In birds, males from two species of Darwin's finches (*Geospiza fuliginosa* and *G. difficilis*) that live in sympatry on the Galápagos archipelago prefer conspecific female models, whereas males of the same species from regions of allopatry show little preference (Ratcliffe and Grant 1983b). Unfortunately, though, this difference may not reflect response to selection: if it is genetically based, the findings are obviously consistent with reinforcement, but there is no hard evidence for such heritability. Instead, as Ratcliffe and Grant note, the difference might reflect learning, as sympatric (but not allopatric) males could learn to distinguish between conspecific and heterospecific females. (Indeed, such learning is known to occur in guppies of the genus *Poecilia*, yielding a false appearance of reinforcement; Magurran and Ramnarine 2004.) This problem afflicts all studies of reinforcement in organisms that can learn; unambiguous tests of enhanced isolation in such cases must involve naïve individuals.

In another bird study, Saetre et al. (1997) found evidence of enhanced isolation between two species of *Ficedula*, the collared and pied flycatchers. While females of both species are brown, males of the collared flycatcher are black and white, and males of the pied flycatcher are either black and white, or brown. Interestingly, brown males occur only where the species overlap. Saetre et al. argue that brown color evolved to prevent hybridization, which yields mostly sterile hybrids (Alatalo et al. 1990; Saetre et al. 1997). Saetre et al. show that, when given a choice of sympatric males, females of both species choose conspecific males. When given a choice of allopatric males, many females choose heterospecific males. While suggestive, this pattern might reflect a kind of mimicry: collared males are dominant to pied males in competition for nesting sites. But collared males tend to ignore brown pied males, possibly allowing

ing pied males to acquire better nesting sites without interference (Alatalo et al. 1994). Brown color, then, might represent female mimicry that acts to reduce interspecific competition.

Finally, there are several reports of enhanced isolation in plants and fungi. Grant (1966c), for instance, studied nine species of leafy-stemmed *Gilia*, all of which produce sterile hybrids. Five species occur in the foothills and valleys of California and are sympatric; the remaining four species are maritime and allopatric to each other. Grant showed that the sympatric foothill-and-valley species are strongly isolated from each other while the allopatric maritime species are not. While this isolation probably involves a mixture of pre- and postzygotic barriers, at least part reflects prezygotic incompatibilities between pollen and pistils. Similarly, two species of tropical herbs, *Costus allenii* and *C. laevis*, occur sympatrically, flower at the same time, and share the same bee pollinator; despite considerable interspecific pollen flow in the wild, artificial hybridization is difficult, reflecting failed pollen-pistil interactions (Schemske 1981; Kay and Schemske 2003). Another species, *C. guanaiensis*, which is ecologically segregated from both *C. allenii* and *C. laevis*—and so shows little interspecific pollen flow with either species in nature—easily produces hybrids in the greenhouse with both *C. allenii* and *C. laevis* (Schemske 1981). Thus, the species that experience high pollen exchange in nature show strong prezygotic isolation, while those that experience low pollen exchange in nature do not. For other possible plant examples, see Levin's (1985) work in *Phlox*, and McNeilly and Antonovics's (1968) work in *Agrostis* and *Anthoxanthum* (Chapters 3 and 4). Recent work has also revealed a pattern of enhanced isolation in fungi. In particular, crosses between species belonging to the genus *Neurospora* show greater reproductive isolation when populations are sympatric than allopatric, although it is unclear how much of this enhanced isolation involves prezygotic barriers (see Dettman et al. 2003 and references therein).

While the above reports tend to support reinforcement, others do not. The literature includes a fair number of studies in which enhanced isolation was looked for but not found (Marshall et al. 2002). Patterson and Stone (1952, pp. 357–358 and 549–550) listed several cases in *Drosophila* (e.g., *D. microspina* × *D. limpiensis*) in which prezygotic isolation increases with geographic distance between two species, contrary to the pattern emphasized by Dobzhansky. Coyne et al. (2002) recently found that *Drosophila santomea* and *D. yakuba*, which have partially overlapping ranges on the volcanic island of São Tomé, show no enhanced isolation in sympatry. Similarly, Walker (1974) found no evidence of song displacement among acoustic insects (although he studied fewer cases than often reported [Howard 1993]). Similarly, Loftus-Hills (1975) argued that evidence for enhanced isolation between the toads *Bufo americanus* and *B. woodhousii* is weaker than previously thought, while Sanderson et al. (1992) found no evidence for enhanced isolation in a hybrid zone between the toads *Bombina bombina* and *B. variegata*. In addition, Doherty and Howard (1996) found that, although male song in the cricket *Allonemobius fasciata* appears displaced in some populations that are sympatric with its sister species *A. socius*, females

are mostly oblivious to these differences. Butlin and colleagues (Butlin 1989; Ritchie et al. 1989; Butlin and Ritchie 1991) further argued that a hybrid zone between two subspecies of the grasshopper *Chorthippus parallelus* provides little evidence of enhanced isolation. But most important of all, Butlin (1987a,b, 1989) argued that several alleged—and classic—cases of reinforcement are flawed in several ways, most seriously in that they often involve taxa that produce *completely* sterile or inviable hybrids. Such cases, he argued, represent reproductive character displacement, not reinforcement.

The case study data, then, do not present a particularly simple picture. Part of the problem is inherent in the approach itself. A case or two of enhanced isolation is seen in one small group, but not in another, quite different, group. Given that all else is surely not equal across these groups, it is hard to know if such inconsistencies reflect the rarity of reinforcement or are consequences of looking at taxa of different evolutionary ages, or taxa that have been in contact for different periods of time, or taxa that show different degrees of postzygotic isolation, and so on.

Evidence from nature: comparative studies

The solution to this problem is clear: we require larger, more systematic studies, ones that allow comparison of the strength of prezygotic isolation in sympatry versus allopatry over large groups, *controlling* for the age of taxa and the strength of postzygotic isolation between them, etc.

This was one of the motivations for our survey of patterns of speciation in *Drosophila*, in which we collected data on geography (allopatry versus sympatry), genetic distance (a molecular proxy for evolutionary age), and the strength of pre- and postzygotic isolation between many pairs of species. Our original survey included 119 species pairs (Coyne and Orr 1989a), while our subsequent update included 171 (Coyne and Orr 1997; see Chapter 3). Some of our results are shown in Figure 10.1. Prezygotic isolation is obviously much stronger between sympatric than allopatric species. Indeed, young sympatric taxa (Nei's genetic distance less than 0.5) show an average prezygotic isolation index of 0.83 (where 0 means no prezygotic isolation, and 1 means complete isolation), while young allopatric taxa show an average of 0.29. This difference cannot be explained by the age of species (i.e., sympatric taxa are not simply older than allopatric). Most important, *postzygotic* isolation does not differ between sympatry and allopatry. Just as expected under the theory of reinforcement, then, pre- but not postzygotic isolation is greater among sympatric taxa.

These data had some effect on the status of reinforcement among evolutionary biologists, an effect that had much to do with the theoretical climate in which these results appeared. As we will see, theorists had all but abandoned reinforcement by the late 1980s. Although the pattern shown in Figure 10.1 does not prove reinforcement—we will consider a long list of possible alternatives—it was, we suggested, sufficiently striking to require taking reinforcement seriously.

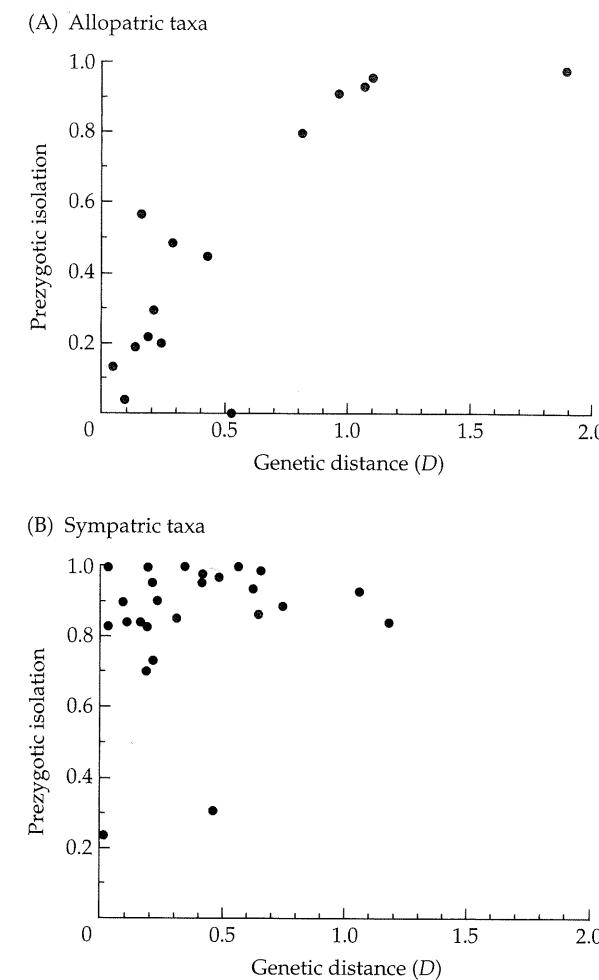


Figure 10.1 Enhanced prezygotic isolation in *Drosophila*. The top plot shows prezygotic isolation between allopatric pairs of species, while the bottom plot shows prezygotic isolation between sympatric pairs of species. (From Coyne and Orr 1997.)

Several other comparative tests of reinforcement have since appeared. The most exhaustive is that of Howard (1993), who searched the literature for cases that, by meeting certain criteria, allow one to test for enhanced isolation in sympatry with some confidence. (Howard required, for instance, that the character studied be known to at least sometimes play a role in reproductive isolation.) Howard identified 48 cases that satisfied his criteria. Remarkably, 33 of these cases (69%) showed a pattern of enhanced isolation. Because these data derive from a wide variety of organisms—*insects, birds, mammals, amphibians, reptiles, fish, and plants*—enhanced isolation is clearly not hard to find. One could argue, however, that Howard's findings reflect a publication bias: cases in which preliminary studies suggest greater isolation in sympatry might be more likely to be subjected to fuller analysis, resulting in publication. To guard against this possibility, Howard further scoured the

literature for cases in which patterns of assortative mating had been assessed in hybrid zones: random or assortative mating may occur here and the results would presumably be published regardless. Howard found that, in a sample of 37 hybrid zones in which patterns of assortative mating had been characterized by genetic, morphological, or behavioral means, 19 showed positive assortative mating and 16 showed random mating, findings that are at least consistent with reinforcement.

Oddly, we know of only one comparative survey of isolation in sympatry vs. allopatry that focuses on plants. Although this study (Moyle et al. 2004) uncovered no evidence of enhanced isolation, only two genera were surveyed (*Silene* and *Glycine*) and aspects of prezygotic isolation that act before pollination (like pollinator behavior and flowering time) were not studied. It is hard, therefore, to know how representative these findings are. The relative neglect of plants by reinforcement workers is unfortunate. Indeed a number of novel predictions could be tested here (e.g., because wind-pollinated plants cannot achieve reinforcement by changing pollen vectors or flower color, reinforcement should often take the form of changed flowering time and/or “aerodynamic isolation” in which heterospecific pollen is less likely than homospecific pollen to land on the stigma; Chapter 6).

Although the above surveys reveal that sympatry can enhance prezygotic isolation, they do not tell us *how often* sympatry matters. The problem is that biases causing either over- or underestimation of the frequency of enhanced isolation probably exist. In Coyne and Orr’s (1989a, 1997) work, for instance, single species were sometimes used in multiple comparisons with other taxa. A single bout of reinforcement between two species could, therefore, give rise to a pattern of enhanced isolation between several species: a species that experienced selection for greater prezygotic isolation might *generally* show greater choosiness. On the other hand, case studies might underestimate the frequency of enhanced isolation: if increased isolation evolves in sympatry by reinforcement but then spreads throughout a species range, contrasts between sympatric and allopatric populations will reveal no reinforcement although it occurred (Walker 1974).

Noting these problems, Noor (1997b) introduced a method that allows estimation of the frequency with which sympatry yields enhanced isolation. The idea is simple. Consider three species: an outgroup species A, and two ingroup species, B and C (Figure 10.2). Species A and B are sympatric while species A and C are allopatric. If reinforcement occurs, A-B will show more prezygotic isolation than A-C. In searching for such contrasts, Noor imposed two restrictions. First, sympatric species can be used in one comparison only and, second, the allopatric species cannot be sympatric with any other close relative (as such sympatry might fortuitously heighten isolation between species A and C). Surveying data from *Drosophila*, Noor concluded that sympatry yields enhanced isolation in 21% of phylogenies like those in Figure 10.2. Because this value is a lower bound on a confidence interval, and because his data incorporated only one component of prezygotic isolation (sexual isolation), Noor argues that it probably underestimates the true frequency of enhanced isolation.

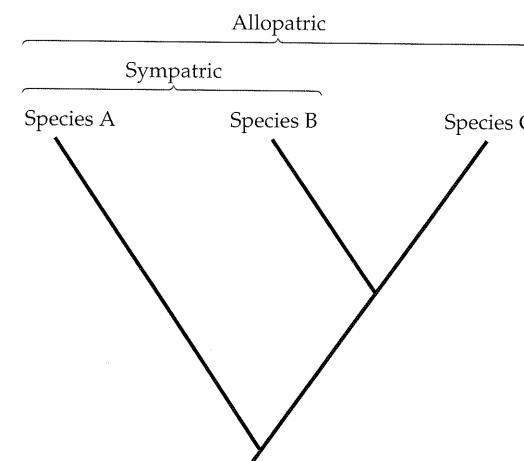


Figure 10.2 Noor’s (1997b) test of the frequency of enhanced isolation. Species A and B are sympatric, while species A and C are allopatric.

Overall, then, the comparative data are consistent with reinforcement. While we have not yet considered possible alternative explanations, work over the last two decades strongly suggests that a *pattern* of increased prezygotic isolation in sympatry is reasonably common in nature. Before turning to theoretical attempts to explain this pattern, we briefly consider whether postzygotic isolation shows a similar pattern.

Reinforcement of postzygotic isolation

While the logic behind selection for increased prezygotic isolation is clear, the same cannot be said for postzygotic isolation. But Grant (1966c, p. 104) and Coyne (1974) argued that reinforcement of postzygotic isolation is possible in species having substantial parental investment. If hybrids are sterile or inviable (where lethality occurs late in development) and if parents sacrifice further reproduction during a period of pregnancy, child rearing, or seed/fruit production, it may pay to abort hybrid development as early as possible (i.e., to evolve hybrid lethality). (See Johnson and Wade 1995 for theoretical treatment of this idea.) Such reinforcement seems most likely in organisms like mammals or plants, where maternal investment is large.

Coyne described several possible examples. One involves the “corky” syndrome in cotton. Hybrids between two species of *Gossypium* suffer a syndrome of morphological and physiological problems (Stephens 1946, 1950). The alleles causing these problems are mostly restricted to regions of geographic overlap between the species. In another case, discussed earlier, Grant (1966c) showed that sympatric species of *Gilia* suffer greater barriers to hybridization than allopatric species. While some of these barriers are prezygotic, others appear to be postzygotic.

There are therefore hints that postzygotic isolation might sometimes be reinforced. Once again, however, the problem is that the approach taken so far has been entirely anecdotal. While we can point to isolated examples of height-

ened hybrid inviability in sympatry, we have no idea how common these cases are relative to those in which postzygotic isolation is greater in *allopatry*, cases that may often go unpublished. Again, the obvious solution is a more systematic approach. (Systematic data have been collected in *Drosophila* [Coyne and Orr 1989a, 1997], but flies do not provide promising material given that they do not necessarily sacrifice future reproduction by producing unfit progeny.) Fortunately, the Grant-Coyne idea makes two predictions that could be readily tested in plants. First, crosses between sympatric taxa should produce fewer viable seeds than crosses between allopatric taxa. Second, no such difference should appear in the strength of hybrid sterility, which cannot be reinforced (at least not in a straightforward way). These predictions could be systematically tested across plants having large fruit or seed investment.

The Theory

The data reviewed above pose a considerable challenge to theory. A pattern of stronger prezygotic isolation in sympatry seems reasonably common. Does this reflect reinforcement? Is reinforcement theoretically possible and, if so, under what conditions?

The history of theoretical work on reinforcement falls into three phases. In the first phase, attention focused on the case in which hybrids are completely unfit and “reinforcement” was found plausible. In the second phase, starting in the mid-1980s, attention focused on the case in which hybrids have non-zero fitness and reinforcement was found unlikely. In the last phase, beginning in the 1990s, attention focused on ever more realistic (or at least more complex) models, and reinforcement was generally found possible.

Early enthusiasm

Though he presented no formal theory, Fisher (1930) offered the first population-genetic model bearing on reinforcement. He considered a scenario closely allied to, if not identical with, reinforcement. In particular, he considered a widely distributed species that adapts over a continuously varying ecological gradient. A certain set of alleles is favored and so increases in frequency at one end of the species’ range, while another set is favored and increases in frequency at the other end of the range. Fisher argued that under such conditions selection will favor the evolution of restricted dispersal (i.e., modifier alleles that reduce mobility will increase in frequency as such alleles reduce maladaptive hybridization; see also Balkau and Feldman 1973). As this process continues, the population might rupture into two distinct species. But, Fisher argued, the evolution of restricted dispersal is not the only imaginable outcome. One can also imagine direct selection against acceptance of a mate of the “wrong” type:

The grossest blunder in sexual isolation, which we can conceive of an animal making, would be to mate with a different species from its own

and with which the hybrids are either infertile or, through the mixture of instincts and other attributes appropriate to different courses of life, at so serious a disadvantage as to leave no descendants (Fisher 1930, p. 130).

Fisher left little doubt about his belief in the efficacy of direct selection for prezygotic isolation.

No further theoretical work on reinforcement appeared until Wilson’s (1965) presentation of Bossert’s unpublished computer simulations. Wilson and Bossert emphasized what would become an important theme: reinforcement is a race between the enhancement of prezygotic isolation and the fusion of populations. One must do more, in other words, than show that there is a selective advantage for alleles that increase mate discrimination. One must show that these alleles reach high frequency *before* populations fuse through hybridization. (No such race occurs if hybrids have zero fitness, since fusion is impossible.) Wilson and Bossert simulated the case in which two populations come into contact and hybridize at an initial “error rate.” Hybrids suffer low fitness due to heterozygote disadvantage at a single locus, while prezygotic barriers involve many genes. Wilson and Bossert found that low hybrid fitness and low error rates often lead to reinforcement, while high hybrid fitness and high error rates often lead to fusion (Figure 10.3). Crosby (1970) performed another early computer simulation of reinforcement. He considered two scenarios: one in which two taxa are sympatric, and another in which they meet in a hybrid zone. Modeling the evolution of flowering time in an annual plant and allowing for partial hybrid fertility, he showed that, in both scenarios, prezygotic isolation could evolve within several hundred generations.

These studies represent some of the earliest computer simulations in population genetics and, not surprisingly, they suffer the shortcomings that characterized such work: unrealistically small population sizes and little replication. More rigorous, and preferably analytic, work was needed.

Such analytic work initially focused on the case in which hybrids are completely unfit. Although this scenario is irrelevant to speciation, it at least enjoys the merit of being mathematically tractable. More important, it is clear that—if enhanced prezygotic isolation does not evolve under these extreme circumstances—it will not evolve at all. One of these early studies was population genetic and the other quantitative genetic.

In the population genetic study, Sawyer and Hartl (1981) analyzed change in allele frequency for a mutation that lowers the probability of mating with a second species, assuming a hybrid fitness of zero and initial random mating. They showed that, for a rare “choosiness” allele having frequency p , $\Delta p \approx p(1-f_1)m_1$ where f_1 is the fraction of the mating pool made up by the population in which the mutation arises, and m_1 is its probability of rejecting a mate from the “wrong” population. Because change in p is always positive, choosiness alleles always invade. Not surprisingly, then, prezygotic isolation readily increases when hybrids are completely sterile or inviable.

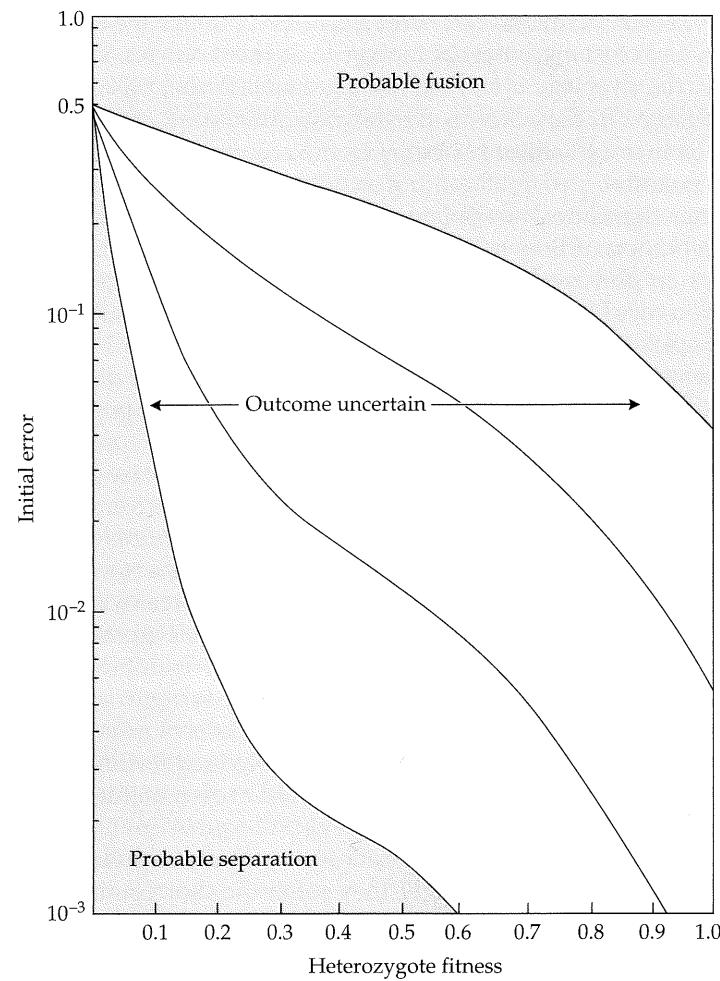


Figure 10.3 The probability of reinforcement versus fusion. (From Wilson 1965, based on computer simulations by Wilson and Bossert.)

Sved (1981) considered a quantitative-genetic model in which the characters affecting mate choice are underlain by polygenes. He allowed separate genes for male versus female mating behavior. Assuming that hybrids suffer zero fitness, he calculated the response to selection on female mating behavior due to hybridization. He found that the change in female mating behavior is proportional to the initial difference in male means between populations. As expected intuitively, populations that begin with more divergent mating behavior enjoy faster change. Once again, therefore, we arrive at an unsurprising answer: increase in prezygotic isolation is unproblematic when hybrids suffer zero fitness.

Objections to reinforcement

This optimistic assessment was vigorously challenged in the 1980's. Evolutionists raised many objections to reinforcement, some substantive, and some less so. By the end of the decade, reinforcement was considered unlikely at best, and an exercise in wishful thinking at worst (Noor 1999; Marshall et al. 2002). We discuss the most important objections below.

SELECTION-RECOMBINATION ANTAGONISM. While the above theory shows that prezygotic isolation can be exaggerated if there is no gene flow between populations, the effect of even minimal gene flow is unclear. Although Sved (1981) suggested that his results might be robust to low gene flow, two papers in the early 1980s suggested otherwise.

The first study was by Felsenstein (1981). Although Felsenstein did not explicitly consider reinforcement, his main finding played an important role in all subsequent thinking about the process. Because we have described this work in Chapter 4, we only sketch it here. Felsenstein considered two interbreeding subpopulations. The model featured three loci: one affecting prezygotic isolation, and two affecting postzygotic isolation. At the first gene, individuals carrying the *A* allele tend to mate amongst themselves, while those carrying the *a* allele tend to mate amongst themselves. At the other two loci, *BC* individuals enjoy high fitness in subpopulation 1 while *bc* individuals enjoy high fitness in subpopulation 2. Felsenstein's question was: Will alleles at the prezygotic isolation locus become non-randomly associated with those at the postzygotic isolation loci? That is, will the species split in two, with *ABC* predominating in subpopulation 1 and *abc* predominating in subpopulation 2? Felsenstein showed that, over part of the parameter space, non-random associations (linkage disequilibrium) between the pre- and postzygotic isolation loci do arise. But over much of the parameter space, they do not. More important, he explained the reason for these different outcomes: building an association between the genes causing pre- and postzygotic isolation involves a conflict between two forces, selection and recombination. Selection builds up such associations but recombination tears them down. As Felsenstein further showed, very strong selection is required to overcome recombination.

Felsenstein identified the key barrier to the evolution of prezygotic isolation in sympatry: recombination. This barrier cannot act, of course, when hybrids suffer complete hybrid sterility or inviability. In that case, the only breeding individuals are *BC* and *bc*, and—in each of these now distinct gene pools—alleles favoring assortative mating can increase in frequency. As Felsenstein also emphasized, recombination cannot block the evolution of increased prezygotic isolation if the same allele causes homogametic matings in both subpopulations (the “one allele” model).

Nearly simultaneously, Barton and Hewitt (1981a) noted a related problem. When postzygotic isolation is incomplete and based on many genes, the selection-recombination problem grows more acute: any single modifier of prezy-

gotic isolation can realistically remain in linkage disequilibrium with only a few of the genes causing postzygotic isolation.

In summary, going from zero to non-zero hybrid fitness may fundamentally change the dynamics—and chances—of reinforcement.

THE RACE BETWEEN EXTINCTION AND REINFORCEMENT. While Wilson and Bossert recognized that reinforcement is a race against fusion, they did not discuss another kind of race: reinforcement must occur before one or both of the populations go extinct (Templeton 1981). The problem is that secondary contact between taxa producing unfit hybrids is analogous to heterozygote disadvantage (Paterson 1978, 1982; Harper and Lambert 1983). When heterozygotes at a locus are unfit, there is an unstable equilibrium: any departure from this equilibrium results in loss of one allele. As Paterson emphasized, taxa coming into secondary contact are susceptible to the same unstable equilibrium, with the rarer population prone to extinction. This is because, under random mating, the rarer population meets the “wrong” type more often and so produces unfit hybrids more frequently than the common population. This causes the rarer population to become rarer yet, triggering a positive feedback loop that ultimately dooms the rare population to extinction. Thus, if reinforcement is to evolve, it must do so quickly.

This argument assumes that the two populations are ecologically identical. But as Sved (1981) noted, if this is true, extinction has little or nothing to do with a race with reinforcement: extinction of one taxon is likely by competitive exclusion alone whether or not the populations evolve prezygotic isolation.

There can, however, still be a race between reinforcement and extinction. For it remains possible that extinction could occur before reinforcement *even when populations differ ecologically*. Indeed, extinction might still occur even when populations use completely separate resources and have independent density regulation: populations will still produce fewer progeny when females waste their gametes on males of the wrong type. Thus, populations might be sufficiently distinct ecologically to coexist in the *absence* of hybridization but not in the *presence* of hybridization. In this case, populations can coexist only if prezygotic isolation evolves quickly and we can sensibly speak of a race between reinforcement and extinction.

This point was highlighted in one of the most influential theoretical studies of reinforcement, that of Spencer et al. (1986). Focusing on the case in which hybrids have zero fitness, Spencer et al. performed computer simulations to determine if increased prezygotic isolation can evolve before hybridizing populations go extinct. They considered the case in which populations have independent logistic growth and thus suffer little chance of extinction in the absence of hybridization. With hybridization, however, each population’s growth was reduced by the fraction of females who mated heterotypically. Male and female mating characters were polygenic.

Spencer et al.’s key result was simple. Even with independent population regulation, extinction often occurs before reinforcement. Indeed, reinforcement

was likely only when populations already differed significantly in mating behavior and so already showed considerable prezygotic isolation. Spencer et al. thus concluded that coexistence without hybridization provides no guarantee against extinction *with* hybridization. After gene flow, the threat of extinction poses the most serious challenge to reinforcement.

THE SWAMPING EFFECT. As Moore (1957, p. 335) pointed out, the alleles causing reinforcement have a selective advantage only within the zone of contact between two populations. If these alleles were advantageous outside this zone they would presumably increase in frequency everywhere—and it would be absurd to contend that they were selected to avoid hybridization. The alleles underlying reinforcement are thus probably either neutral or deleterious outside the zone of contact. This poses two problems.

First, selection in the zone of overlap might be swamped by gene flow from allopatric regions (Bigelow 1965). This would seem particularly serious immediately after secondary contact, as the zone of overlap would presumably be narrow and the area of non-overlap large. Gene flow would thus be overwhelmingly *into* the zone of overlap. While Caisse and Antonovics (1978) showed that prezygotic isolation could evolve in sympatry despite gene flow if the alleles involved are neutral elsewhere, Sanderson (1989) showed that, as one might guess, the chances of reinforcement fall if these alleles are deleterious elsewhere. (The same problem arises if two populations do not physically overlap but exchange migrants at some rate. Reinforcement requires *some* gene flow, but not too much [Sanderson 1989; Servedio and Kirkpatrick 1997; Cain et al. 1999; Servedio 2000].)

Second, it is hard to see how the alleles underlying reinforcement could spread outside the region of overlap to the rest of the incipient species. One possibility, of course, is that they do not. Instead, the alleles conferring reinforcement could remain permanently trapped in the zone of contact. Alternatively, the populations might slowly migrate into a region of complete geographic overlap in which all individuals experience selection for prezygotic isolation. Finally, Howard (1993) suggested that *species*, not alleles, might spread from zones of contact: long selection in a zone of contact might yield reproductive isolation between individuals from the overlap zone and those from the remaining parental populations. Thus, a newly evolved species might ultimately migrate outwards, becoming sympatric with its ancestors. But it is not clear how this escapes the key problem: if swamping impedes reinforcement, it would probably also impede the evolution of reproductive isolation between individuals in the contact zone and those outside it.

REINFORCEMENT IS SELF-DEFEATING. Moore (1957, p. 336) and Spencer et al. (1986) emphasized another problem: reinforcement is self-defeating. Because the strength of selection for increased prezygotic isolation is proportional to the frequency of hybridization, any increase in prezygotic isolation automatically reduces the strength of selection for further reinforcement. Reinforcement thus pulls the rug out from under itself.

Spencer et al. saw this effect in their simulations. Among populations that survived extinction, mating behavior changed quickly early on, but then slowed as populations diverged. Spencer et al. (1986, p. 257) thus concluded that it is “extremely doubtful whether selection alone would lead to complete speciation.”

PRE- VERSUS POSTZYGOTIC ISOLATION. Finally, while there is a selective advantage to increasing prezygotic isolation, there is also an advantage to *decreasing* postzygotic isolation. Why not, then, eliminate the cost of hybridization by improving the fitness of hybrids, precluding reinforcement?

This idea has received little attention (but see Sanderson 1989). This may reflect the common intuition that it is hard to “undo” postzygotic isolation, at least in its intrinsic form (Muller 1939). But when postzygotic isolation is incomplete, it is obviously easy to purge incompatible alleles, leaving only compatible ones. Moreover, even when postzygotic isolation is complete (the case of reproductive character displacement), modifier alleles might accumulate that lessen the effects of genic incompatibilities. In addition, there is no reason for thinking that it is hard to reverse *extrinsic* postzygotic isolation, no matter how strong.

SUMMARY. The above objections posed an enormous challenge to the theory of reinforcement, a challenge that was not immediately met. Consequently, by the late 1980s the weight of opinion swung against reinforcement. The theory appeared dying, if not dead.

The revival of reinforcement

The third and most recent phase in the history of reinforcement featured a sudden reversal of fortunes. Over a span of a few years, reinforcement went from seeming nearly impossible to seeming likely. This reversal reflected a burst of theoretical work, a burst that was triggered not by perceived flaws in the above objections but by the appearance of new data sets suggesting that reinforcement—or at least something *looking* like reinforcement—acts in nature (Coyne and Orr 1989a; Howard 1993). The patterns revealed by these studies stood in stark contrast to those expected under existing theory. As we will see, the theory that followed identified and incorporated several factors ignored by previous theory.

The first of these new studies was that of Liou and Price (1994). Their computer simulations roughly followed those of Spencer et al. (1986). Populations were sympatric and ecologically distinct (i.e., a population could not go extinct by competitive exclusion, though it could by hybridization). Mate choice featured a male trait (z) and a female trait (y), each underlain by multiple loci. Females had an absolute preference function such that females of phenotype y preferred males of phenotype $z = y$. In a departure from Spencer et al., Liou and Price studied the full range of hybrid fitness.

They showed that reinforcement occurs reasonably often if hybrid fitness is low and populations already differ in female preference and male character. Although reinforcement is most likely when hybrids have zero fitness, this is not required. Reinforcement *does* occur in the face of gene flow. Liou and Price also found that any factor that decreases the chances of extinction increases the odds of reinforcement. Such factors include those boosting the carrying capacity or the intrinsic rate of increase. As expected from Felsenstein (1981), greater recombination hinders reinforcement when hybrids have non-zero fitness; reinforcement, however, does remain possible with recombination. Liou and Price (p. 1451) concluded, “There is a wide range of genetic and ecological conditions under which reinforcement rather easily occurs.” Indeed, they found that reinforcement occurs far more readily than in Spencer et al.’s study, even when using the same parameter values (Figure 10.4).

Given that this work signaled a reversal in opinion about reinforcement, it is worth understanding why Liou and Price’s results differed from those of Spencer et al. The reason seems clear. Liou and Price’s simulations allowed for sexual selection, while Spencer et al.’s did not. Recall that assortative mating gives rise to genetic correlations between male and female trait values (since females with large y values tend to mate with males with large z values). Because these associations arise automatically, it may not be obvious why they appear in Liou and Price but not in Spencer et al. The reason is that Liou and Price’s simulations explicitly followed the loci encoding male and female traits, while Spencer et al.’s simulations did not. By recording the genotypic values of males and females who mated, and assuming a constant heritability, Spencer et al. calculated *separate* responses to selection for male and female traits. But this approach does not allow a correlation between male and female genotypic values, precluding runaway sexual selection.

The important point is that sexual selection boosts the odds of reinforcement because *two* forces now drive the evolution of female preference: direct

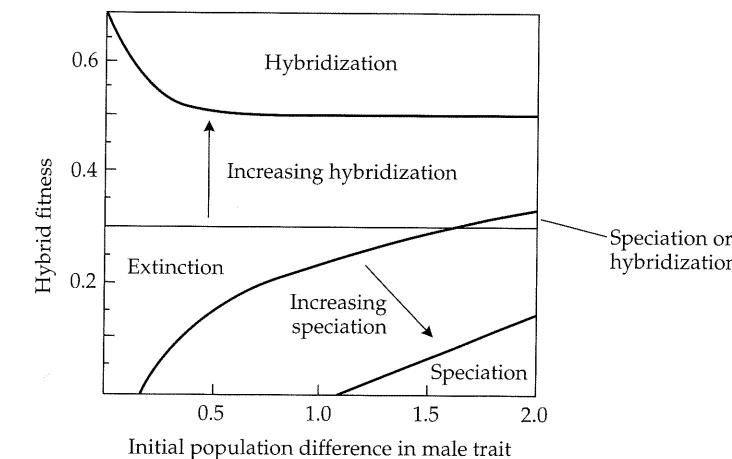


Figure 10.4 The parameter space over which reinforcement is likely to occur. (From Liou and Price’s 1994 computer simulations.)

selection on females (reflecting the price paid for choosing a male of the wrong type), and indirect selection on females (reflecting a correlated response to sexual selection on males). Sexual selection also helps overcome the self-defeating nature of reinforcement. Although the intensity of natural selection for reinforcement decreases as reinforcement progresses, the intensity of sexual selection need not.

Later work by Kelly and Noor (1996) also emphasized the ease of reinforcement. Kelly and Noor focused on the evolution of mate discrimination, not preference. Maladaptive hybridization might, in other words, favor a narrowing of female taste—such that males that were previously acceptable become unacceptable—not a shift in mean female preference. Although Kelly and Noor's simulations were complex, their question was simple: Can a rare allele that increases female discrimination increase in frequency when hybrids suffer low fitness? They found that the answer is often yes. More surprisingly, they found that this remains true even when selection against hybrids is weak. The question is why.

In retrospect, the answer seems straightforward. Because reinforcement here reflects the evolution of greater female discrimination, not divergence in female preference, reinforcement involves the fixation of the *same* discrimination allele in both populations. And, as Felsenstein (1981) emphasized, this reduces the challenge posed by recombination: in a “one-allele” model of reinforcement, linkage disequilibrium need not be maintained in the face of recombination. Thus, while Liou and Price showed how the self-defeating nature of reinforcement could be overcome, Kelly and Noor showed how selection-recombination antagonism could be overcome. In each case, the mechanism involved—sexual selection, or female discrimination—seems biologically plausible.

These pro-reinforcement conclusions have been confirmed in the latest and most systematic theoretical studies of reinforcement, those of Servedio and Kirkpatrick. These authors have considered the effect of many factors on the probability of reinforcement: one-way versus two-way gene flow (Servedio and Kirkpatrick 1997); assortative versus preference-based mate choice (Servedio 2000); many versus few loci causing hybrid incompatibilities (Kirkpatrick and Servedio 1999); extrinsic versus intrinsic postzygotic isolation (Kirkpatrick 2001); and postmating, prezygotic versus postzygotic fitness problems (Servedio 2001). Their qualitative conclusion is simple: reinforcement can occur under a broad range of conditions. Indeed, reinforcement can occur under both two-way and one-way gene flow (though it is easier in the former); whether mating is assortative or preference-based (though it is often easier in the former); or whether many versus few loci cause hybrid incompatibilities (though it is slightly easier in the former). Moreover, reinforcement can occur regardless of the nature of selection against hybridization (extrinsic isolation will serve as well as intrinsic; indeed, hybrids might merely be unattractive to potential mates from the parental species [Coyne and Orr 1989a]).

In summary, recent theory leads to two broad conclusions, one satisfying and the other troubling. The satisfying conclusion is that reinforcement seems

both possible and plausible. Indeed, while the details of the models differ wildly—complicating systematic assessment of the precise conditions under which reinforcement is likely—virtually all recent models suggest that reinforcement is easier than imagined a decade ago (Turelli et al. 2001; Kirkpatrick and Ravigné 2002). The troubling conclusion concerns the extraordinarily labile nature of this body of theory. While early theory mostly denied reinforcement and later theory mostly affirmed it, what occurred in the interim was the appearance of large empirical surveys suggesting that something that looks like reinforcement occurs. Put bluntly, theory said it could not happen until the data said it probably did. While a charitable interpretation of this history would emphasize the healthy dialectic between theory and data, we suspect that a more sober one would emphasize the often-misunderstood role of theory in evolutionary biology. Evolutionary theory is not, at least usually, in the business of telling us what is and is not biologically possible, a matter that depends on the particular suite of assumptions made. Rather, evolutionary theory tells us what does and does not follow from a given set of verbal assumptions. These roles were often confused in the history of reinforcement theory.

Alternative Explanations

The data reviewed earlier show that prezygotic isolation is often stronger in sympatry than allopatry. And the theory just reviewed shows that reinforcement is possible. But it does not follow that reinforcement is the correct explanation of the enhanced isolation seen in nature. Instead, several alternative explanations are possible. The primary task now confronting students of reinforcement is thus clear: to distinguish among these competing hypotheses.

As a step to this goal, we critically review several alternative hypotheses. Our discussion partly follows that of Noor (1999), although we consider several alternatives that he did not. In the end, we argue that reinforcement may yield unique (or nearly unique) predictions that allow us to distinguish it from the alternatives.

Publication bias

One of the most obvious but least discussed alternative explanations of the pattern of enhanced isolation is that it reflects a publication bias (though see Howard 1993). Findings that are consistent with reinforcement are exciting and so seem more likely to see the light of day than the opposite findings. Fortunately, this bias cannot explain all of the data. As Noor (1999) pointed out, a pattern of enhanced isolation appeared in Coyne and Orr's (1989a) literature survey of *Drosophila* long after the data were originally collected. Indeed, this is one of the advantages of such broad surveys: they are both less vulnerable to publication bias in the primary literature (as most data were collected for other reasons) and, as large and time-consuming enterprises, are likely to be

published whatever the conclusion. We agree with Noor (1999) that the pattern of enhanced isolation is probably real, whatever its cause.

Differential fusion

Templeton (1981) suggested that enhanced isolation might reflect differential fusion, not reinforcement. According to this idea, different taxa coming into secondary contact might show a range of prezygotic isolation. Taxa that are strongly isolated can persist in sympatry while those that are weakly isolated cannot. Instead, the latter fuse, leaving us with nothing to study. In the end, a pattern of greater prezygotic isolation in sympatry emerges although there has been *no* response to individual selection for increased isolation.

While this seems a powerful explanation of enhanced isolation, several arguments have been made against it. Coyne and Orr (1989a) argued that the hypothesis cannot explain two patterns seen in *Drosophila*. First, differential fusion predicts that levels of prezygotic isolation in sympatry are a *subset* of those in allopatry: strongly isolated taxa that can persist in sympatry must, by hypothesis, preexist in allopatry. But in *Drosophila*, the high levels of prezygotic isolation seen in sympatry are not seen in allopatry, weighing against differential fusion. As Noor (1999) points out, however, this argument is not conclusive. Strongly isolated allopatric taxa might, after all, be rare. If, for instance, only 1 in 100 cases of secondary contact involves taxa that are isolated enough to persist, our sampling of allopatric taxa would probably not detect them.

Second, differential fusion would seem to predict that both pre- *and* postzygotic isolation would be greater in sympatry than allopatry: either barrier lowers the odds that taxa fuse. But the *Drosophila* data show a striking enhancement of pre-, but not postzygotic, isolation in sympatry. Again, this pattern seems more consistent with reinforcement. Although this argument has been weakened somewhat by Gavrilets and Boake's (1998) demonstration that prezygotic isolation prevents fusion more effectively than postzygotic isolation, postzygotic isolation should still have *some* effect under differential fusion. It does not appear to.

Noor (1999) offered a novel objection to differential fusion. While it might, in principle, explain an *among*-species pattern of enhanced isolation (as in Coyne and Orr 1989a), it has a hard time explaining a *within*-species pattern (as in Noor 1995) unless gene flow among populations is very low. The reason is that if populations of a species all enjoy reasonable levels of gene flow before secondary contact, they would probably not differ substantially in levels of prezygotic isolation when coming into secondary contact with another species—and thus differential fusion cannot occur. This is an important point and, at least in one case of enhanced isolation—that of *Drosophila pseudoobscura* and *D. persimilis*—we know that levels of gene flow between populations are high (Schaeffer and Miller 1992).

Thus, while we cannot rule out differential fusion, and while it surely occurs at some rate in nature, there is some reason to believe that it is not the leading explanation of enhanced isolation.

Direct ecological effects

If an ecological variable both allows coexistence of two species and fortuitously increases mate discrimination in one or both of them, a pattern of enhanced isolation might result.

As Noor (1995, 1999) notes, however, this hypothesis suffers an awkward problem: why should ecological factors always *increase* mate discrimination? Why don't the ecological conditions found in sympatry just as often decrease discrimination? One possible answer is that sympatric conditions *do* decrease versus increase mate discrimination equally often, but in the former case, taxa hybridize and so are lost to fusion. But in this case, the ecological effects hypothesis collapses into a special case of the differential fusion hypothesis.

The ecological hypothesis suffers another problem. It is unclear how well ecological factors can explain the actual observations. In *Drosophila*, at least, comparisons of the strength of prezygotic isolation in sympatric versus allopatric taxa are made under uniform laboratory conditions (i.e., vials) in which ecology is almost certainly beside the point.

Ecological character displacement

There is, though, a special sense in which ecological effects might systematically give rise to enhanced isolation. Under this “ecological character displacement” hypothesis, natural selection at the level of individuals *does* act in sympatry, yielding heightened isolation, but selection acts not to reduce maladaptive hybridization but to reduce niche overlap between species. Response to such selection might *incidentally* change mate signals, yielding greater prezygotic isolation.

We know of only one attempt to experimentally distinguish reinforcement from ecological character displacement. Rundle and Schlüter (1998) compared the mate discrimination of sympatric versus allopatric stickleback females. Importantly, allopatric females were taken from populations that resemble sympatric females both morphologically and ecologically. Rundle and Schlüter found that sympatric females remain choosier than allopatric females despite this attempted leveling of morphological and ecological differences.

A special case of the ecological character displacement hypothesis has received a good deal of attention. According to this “noisy neighbors” hypothesis, sympatric species that produce similar courtship songs might interfere with each other acoustically, leading females to waste time and resources tracking males of the wrong species. Thus—even if species do not hybridize—natural selection might favor song divergence in sympatry. Such interference could yield enhanced isolation, at least when examining proxies for prezygotic isolation like song (Otte 1989; Noor 1999). As noted earlier, there is in fact evidence that courtship song sometimes differs between species in sympatry more than expected by chance: Otte (1989) examined divergence in male song in Hawaiian crickets of the genus *Laupala*. Gathering data on 46 cricket taxa distributed over 68 sites where at least two species coexist, he showed that the “spacings” between sympatric songs are greater than expected by chance. Some force has

displaced song in sympatry. The question is whether this force is selection against acoustic interference or selection against maladaptive hybridization.

In principle, the noisy neighbor and the reinforcement hypotheses can be distinguished. If taxa hybridize at some rate, both hypotheses remain viable. But if taxa do *not* hybridize, reinforcement cannot occur, though noisy neighbor displacement can. Thus, enhanced isolation among non-hybridizing taxa is more plausibly explained by noisy neighbor effects. Unfortunately, this test is imperfect, as we can usually only ask if taxa hybridize *now* (Noor 1999). But the fact that taxa do not presently hybridize does not mean they did not do so in the past. Later, we suggest a test that may allow one to better distinguish noisy neighbor effects from reinforcement.

Runaway sexual selection

Day (2000) showed that a type of runaway sexual selection could yield results similar to those expected under reinforcement. He considered a model in which female choice is costly and a population is distributed over an ecological gradient. This gradient causes different male trait values to be favored in different locations. Day showed that, although runaway selection does not usually occur when female choice is costly, it *does* occur in spatial models, even when female choice is costly. More important, if the male trait shows a sharp transition along the ecological gradient, a pattern resembling enhanced isolation can arise: females mate randomly near the ends of the gradient but assortatively near the transition zone.

As Day notes, sexual selection and reinforcement can, in principle, be distinguished. The sexual selection model holds only with a single population arrayed along an ecological gradient—not with two once-allopatric populations that come into contact. Thus, the runaway sexual selection model cannot account for enhanced isolation between taxa showing *postzygotic* isolation. More precisely, the model could explain such cases only if one believes that postzygotic isolation arose within a single continuously distributed population, not in allopatry. This seems to us unlikely, especially for intrinsic postzygotic isolation.

Sympatric speciation

Finally, we note a radical alternative: sympatric speciation. A pattern of strong prezygotic (but not postzygotic) isolation among young sympatric (but not allopatric) species might arise by the splitting of taxa in sympatry. Sympatric speciation, after all, is likely characterized by the rapid evolution of mate discrimination, not postzygotic isolation (at least of the intrinsic type).

The attractiveness of this hypothesis depends on the plausibility of sympatric speciation as determined by independent lines of evidence (Chapter 4). For present purposes, we assume that sympatric speciation is plausible, and merely ask if its signature can be distinguished from that of reinforcement. When dealing

with data at the species level (as in the Coyne and Orr data sets), one approach is to take advantage of Noor's (1997b) biogeographic contrast as shown in Figure 10.2. If the outgroup species A speciated sympatrically from the common ancestor of B and C, one would predict strong prezygotic isolation between A and B *and* between A and C. If, however, all speciation were allopatric but reinforcement occurred upon geographic contact, one would predict stronger prezygotic isolation between species A and B than between species A and C. Noor's (1997b) results are, of course, consistent with this second prediction.

Matters are different at the *population level* (i.e., when explaining a pattern of enhanced isolation among sympatric populations of two species). The clearest scenario involves a cline. Here, there are at least two ways to distinguish reinforcement from sympatric speciation. The first concerns the key difference between the models: reinforcement posits that a cline reflects secondary contact, while sympatric speciation does not. One might therefore test for telltale signs of secondary contact: as Barton and Hewitt (1989) emphasized, secondary contact is supported if two taxa show congruent clines over many characters. Second, reinforcement and sympatric speciation make different assumptions about postzygotic isolation (Kirkpatrick and Ravigné 2002): the former requires its pre-existence, while the latter does not. Sympatric speciation is therefore unlikely if two young taxa show intrinsic postzygotic isolation in both allopatric and sympatric populations.

Distinguishing the Alternatives

Although we have described tests that allow us to distinguish reinforcement from a particular alternative, what we would most like to have is a way to distinguish reinforcement from *all* the alternatives. We would like, in other words, to have a pattern that could serve as a signature of reinforcement. We emphasize that the search for such a signature does not require us to believe that enhanced isolation is *always* caused by reinforcement. Some of the above alternatives can, and surely do, act. For instance, even if reinforcement is real and reasonably common, differential fusion surely acts as well. What we are looking for, then, are predictions that allow us to infer the action of reinforcement, *not* predictions that allow us to conclude that the alternatives never act.

Noor (1999) suggested such a prediction. If heightened prezygotic isolation in sympatry is due to natural selection for greater choosiness, the alleles underlying enhanced isolation should, he argued, be mostly dominant. This prediction follows from "Haldane's sieve," the fact that dominant favorable mutations enjoy greater probabilities of fixation than recessive favorable mutations. Noor reviewed the evidence bearing on this prediction from *Drosophila*, finding it mixed.

Unfortunately, this prediction suffers several problems. First, the degree of dominance of an allele may differ between pure species and hybrids, as Noor noted. Second, Haldane's sieve itself is not straightforward. Although domi-

nant mutations do enjoy higher probabilities of fixation when mutations are unique, probabilities of fixation are nearly independent of dominance when mutations come from the standing genetic variation and were previously deleterious (Orr and Betancourt 2001). (This seems likely in the case of reinforcement: “choosiness” alleles are probably kept rare in allopatry because they are costly in the absence of the other species.) Finally, Haldane’s sieve does not allow us to distinguish between reinforcement and ecological character displacement, including the noisy neighbors scenario. In *both* cases, enhanced isolation is due to natural selection at the individual level and so, in both cases, the alleles involved might be dominant.

Are there any other predictions that might distinguish reinforcement from the above alternatives? There seems to be at least one. In most species, females will pay a larger fitness cost for mating with the wrong species than will males. Thus, as Partridge and Parker (1999) emphasized, reinforcement should typically result in larger changes in female than male behavior, physiology, or morphology. In the simplest scenario, males of two species may have diverged in allopatry, but do not change at all upon secondary contact. Females, on the other hand, become choosier upon contact, as they are now confronted with two types of males. This sex asymmetry was already appreciated in some of the earliest papers on reinforcement (e.g., Sved 1981, p. 210), and on sexual selection in good genes models (e.g., Trivers 1972). But while Sved (1981), Trivers (1972), and Partridge and Parker (1999) all saw the point, none seem to have emphasized that it might provide a systematic way to distinguish reinforcement from the above alternative hypotheses. But it seems unlikely that any of these alternatives would predict female-specific changes, and some would seem to predict the opposite.

There is no reason, for instance, why a publication bias should result in more cases of female than male changes. Similarly, differential fusion would not seem to predict sex-differential effects. Fusion occurs because of gene flow and it is irrelevant whether that flow occurs because of female or male behavior. Similarly, there is no obvious reason why direct ecological effects would preferentially change female behavior or preference in sympatry. Similarly, fortuitous changes in mate choice as a result of ecological character displacement might affect either sex; indeed, with noisy neighbor effects, we would, if anything, predict more change in *male* not female traits in sympatry, as males are the sex that typically calls. Although we cannot rule out the possibility that some of the above hypotheses might, under some circumstances, lead to preferential change in female behavior or structure, the prediction does not seem to follow naturally from them. (The sole possible exception is sympatric speciation. We have argued on independent grounds, however, that sympatric speciation offers a fairly implausible explanation of heightened prezygotic isolation in sympatry when considering either species or population level data.) We suspect, therefore, that systematically greater change in female behaviors or structures in sympatry represents a reasonably robust signature of reinforcement. (Marshall et al. 2002 argue that conspecific gamete precedence might

eliminate much of the cost of hybridization in females and increase it in males. They conclude, however, that this will probably prevent reinforcement. Here, we are interested only in those species pairs that *do* show enhanced isolation.)

In fact, we can recognize three sub-predictions. The first is the obvious one that, among taxa pairs showing greater sympatric isolation, enhanced isolation should be due to changes in female components of mate choice more often than to male components of mate choice. (Note that we are not arguing that males of the two taxa do not differ; females are obviously choosing on the basis of something. Our prediction is that enhanced isolation reflects greater female than male *changes* in sympatry. We are also not arguing that male characters do not change in sympatry. We are arguing that enhanced isolation is due *more* to female changes than to male changes.) Second, in cases in which hybrids have low fitness in one direction of a species cross (A female × B male) and high fitness in the other (B female × A male), A females will be more discriminating than B females. Again, this follows from the fact that males pay little or no price for hybridization. Finally, reinforcement due to greater changes in *male* behaviors or structures might occur in those rare taxa in which males mate only once, or invest significantly in offspring or copulation (e.g., by nuptial gifts). There should be, in other words, exceptions that prove the rule. The finding that sympatric males of the butterfly genus *Heliconius* reject or ignore females having mimetic patterns from the wrong species might well represent such an exception (Jiggins et al. 2001b) as *Heliconius* males transfer costly spermatophores (Naisbit et al. 2001).

As far as we know, none of the above predictions has been tested systematically. Until this is done, or until other unique predictions are offered and tested, we will not be able to conclude with confidence if the pattern of enhanced isolation seen in sympatry is usually—or only rarely—caused by reinforcement.