Host shifts, genetic models of sympatric speciation and the origin of parasitic insect species

G.L. Bush (1) and S.R. Diehl (2)

1. Department of Zoology, Michigan State University, East Lansing, Michigan 48824, USA

2. Department of Zoology, University of Texas, Austin, Texas 78712, USA

Abstract

Host shifts in parasitic insects may involve changes in host selection and survival genes. These shifts occur most often when the old and new host grow sympatrically and usually require genetic changes in both host selection and survival genes. Several proposed genetic modes of sympatric speciation are assessed with respect to their relevance to host race formation and speciation in parasites. By assuming random host selection and ignoring the fact that parasites frequently mate on their hosts these models are inappropriate for host-parasite systems. An alternative model of sympatric host race formation and speciation is proposed which incorporates more realistic assumptions about host and mate choice which significantly alter the pattern of gene flow and selection. Other ecological and genetic factors that influence host shifts are also considered. From this perspective host races and species are likely to arise sympatrically in a wide range of parasitic organisms.

Key-words: host race formation, speciation, genetic models, evolution

It is now a well recognized phenomenon: a plant introduced into a new area frequently remains free of serious pests and other organisms for varying lengths of time. Inevitably, however, the new plant succumbs first to a trickle, then to a flood of invading species. Eventually a more or less stable biotic guild develops consisting of insects, nematodes, fungi, and microorganisms all interacting at various levels of complexity and importance. At this point the new host may have acquired an array of new host races and even species that have evolved during the course of adaptation. In some aspects the process resembles succession in plant communities. In this paper we wish to focus from a genetic perspective on the conditions under which such host shifts occur and how they may lead to host race formation and speciation.

We define a host race as a population of a species living on and showing a preference for a host which is different from the host or hosts of other populations of the same species. Host races represent a continuum between forms which partially interbreed to those that rarely exchange genes (modified from Bush, 1969). We view parasites in the general conceptualization of Price (1980) to include a wide variety of species which gain their resources from other species, usually to the latter's detriment. The term parasite would encompass a range of life forms from microorgan-
isms and fungi to arthropod parasitoids of plants and animals. Parasitic species thus defined make up a significant percentage of the world’s faunal diversity (Bush, 1975).

There are four adaptive genetic scenarios that might occur once a new host plant is introduced into an area. The following genetic changes may be required for host shifts:
- none
- survival gene only
- host recognition gene only
- host recognition and survival genes

We will discuss the evolutionary consequences of each case separately and then focus on alternative genetic models of sympatric host race formation and speciation.

Adaptive genetic scenarios

Case I. No adaptive genetic changes required for host shift If the parasite is capable of recognizing and surviving on the new host plant, then the host range of the insect will simply be expanded to include the new host. No new mutations need to occur at least initially that improve survival or host selection prior to the host shift. This may be true even if the new host is of lower nutritional quality than the original host, due to an escape from competition, predation, or parasitism. Although several factors such as induction or host plant phenology might partially reduce gene flow between populations on the different hosts for non-genetic reasons, in the absence of mutation and selection which improve fitness on the new host, it is unlikely that these populations will diverge to any great degree. However, if mutations which improve host recognition and survival on the new plant do arise, host race formation and speciation could evolve. The situation then resembles Case IV discussed later.

Case II. A change in survival genes is required for a host shift In this situation, few insects or their progeny can survive on the introduced host unless a mutation occurs in a survival gene that makes it possible to develop on the new host. This type of genetic host-parasite interaction involving resistance and survival genes is well known (Day, 1974) and is best exemplified in insects by the highly coevolved gene-for-gene relationships between resistance genes in wheat cultivars and survival genes in the Hessian fly. As long as these flies mate at random and are unable to distinguish resistant from susceptible genotypes of the host plants, survival genes will be retained as a polymorphism. Host races are unlikely to evolve in the face of such gene flow unless assortative mating by host preference develops subsequent to the shift.

Case III. A change in a host recognition gene is necessary for shift to a new host plant Parasites can sometimes be forced to oviposit and feed upon a non-host species with no obvious effect on survival or fitness (Dethier, 1954; Fraenkel, 1969) suggesting survival genes are not always involved in host utilization. Yet the potential host is never used in nature even when growing side-by-side with a normal host.
The new host in this case is rejected because it does not provide some essential host recognition cue. Alternatively, a potential host may be unacceptable because of intense competition, predation, or parasitism (Smiley, 1978; Gilbert, 1979), although not all examples can be explained in terms of such 'ecological monophagy.'

When alteration in a host recognition gene results in a shift to a new host plant and there are no differences in parasite fitness on either host, it is unlikely that sympatric speciation will occur. When closely related, host-specific species can be forced to utilize each other's host with no apparent reduction in fitness, it may be reasonable to conclude that speciation was the result of allopatric isolation or perhaps occurred via the 'runaway process' of sexual selection (Kirkpatrick, 1982; Lande, 1981).

Case IV. A change in host recognition and survival genes is required for a host shift. A shift from one host plant to another may require genetic alterations of loci responsible for both host recognition and survival (Huettel & Bush, 1972). The genetics of such a shift have been explored by Bush (1975) who concluded that host shifts of this nature are the most likely candidates for producing new host races and sibling species in parasitic groups. For this reason we will discuss the implications of host race formation involving both host recognition and survival genes in the context of previously proposed mathematical models of speciation.

Genetic models of host race formation and speciation

Several authors have developed mathematical models of the genetics of sympatric speciation (Maynard Smith, 1966; Bazykin, 1969; Caisse & Antonovics, 1978; Pimm, 1979; Felsenstein, 1981). Although none of these models were directed specifically to resolve the problems of speciation in parasites, they do relate to the problem in varying degrees. One of the most widely cited of such models is that of Maynard Smith (1966). One variation of his model involves genetically controlled host selection, with mating occurring on the host. Although this approach may apply realistically to parasites, a full analytical and quantitative treatment of the specific conditions necessary for the evolution of reproductive isolation by this mechanism was not presented. Thus, it is difficult to assess how common such a mode of speciation might be under natural conditions.

More recently, Felsenstein (1981) has developed a similar haploid model of sympatric speciation based on the interaction of two diallelic viability (survival) loci (B and C) which preferentially adapt individuals to either of two niches (host 1 and 2), and a third assortative mating locus (A) which under the right conditions may bring about reproductive isolation. By computer simulation, this model attempts to define the minimum parameters of selection intensity and genetically determined assortative mating behaviour which are necessary for the evolution of reproductively isolated populations. Its conclusion is that in the absence of linkage between assortative mating and viability loci, conditions favouring sympatric speciation may be restrictive, requiring either high selection intensities or strong expression of assortative mating genes. However, there are a number of reasons for questioning the relevance of this mathematical model to the actual biological processes involved in host
shifts by parasitic organisms.

The central problem is that this model fails to take into consideration the possible association between mate selection and host selection behaviours. It is well known that a wide range of parasitic organisms frequently utilize their host as a 'rendezvous' site for mating and courtship (Bush, 1975; Price, 1980). In Felsenstein's model, these two processes are totally uncoupled. Individuals mate assortatively but then randomly disperse with respect to the A locus to either of the two hosts available irrespective of alleles at their viability loci (B, C). Thus, even when conditions result in the development of two distinct species, each continues to randomly select and utilize both of the available hosts equally. This occurs despite the fact that selection during each generation eliminates non-viable genotypes that happen to settle on the wrong host. Such a situation clearly bears little resemblance to proposed models of sympatric speciation via host race formation, where mate and host selection are coupled and emerging species are largely confined to the utilization of a single host (Bush, 1975).

Consider the case where genetic variability exists at a locus governing both habitat and mate preference as in many parasites. Selection can act on this locus in at least two fundamentally different ways. First, selection for assortative mating can prevent recombination between two interacting viability loci as in Felsenstein's model. However, if mate and host selection are biologically coupled activities as they are in many parasitic organisms, and host selection genes simultaneously specify mate choice, selection can operate through a second and potentially more efficacious process. An association (i.e. linkage disequilibrium) between an allele at a viability locus which confers superior fitness on one of the two hosts with an allele at a second locus specifying a preference for this same host will be directly favoured. This occurs not only due to the reduction of less viable recombinants (as in Felsenstein's model), but also because of the obvious benefit which accrues to individuals who select the host which is most suited to the alleles present at their viability loci.

Taken from this perspective, perhaps a more appropriate mathematical model of the biology of parasite host shifts is the migration modification model of Balkau & Feldman (1973). This model envisions two populations and two loci. One locus represents a viability locus, with alternative alleles each adapted to a different habitat. The second locus segregates for alleles influencing the tendency for individuals to migrate from one habitat to another. In the context of host shifts, this would be analogous to an induction gene which increases the tendency to remain on whichever host they previously fed upon. It was found that if an allele which reduced the tendency to migrate was introduced at a low frequency, it would always increase and thus bring about at least partial reproductive isolation between the two populations.

Felsenstein's model poses more severe restrictions on the development of reproductive isolation. He suggests that this is so because reproductive isolation is due to the substitution of the same allele (which reduces migration) into both populations in the Balkau and Feldman model versus the greater difficulty of substituting two different alleles for assortative mating into each of the two populations in Felsenstein's model. Although this may be true, the greater part of the difference between the two models may lie in the fact that Balkau's and Feldman's model has the
additional advantage of coupling the ecologically important variable of mate selection with habitat preference so that genotypes remain in the environments to which they are best adapted.

Although the model developed by Balkau & Feldman (1973) represents a step in the right direction, additional modifications are necessary in order to more realistically simulate the biological properties of parasitic organisms. We wish to point out how Felsenstein’s model might be modified to accommodate cases where host race formation comes about as a consequence of different host choice alleles (A,a). For the remainder of the paper we will refer to this modification as the parasite model. Since it is assumed that individuals mate on their host, such differences in host choice automatically bring about at least partial reproductive isolation. It is also assumed that prior to the introduction of the new host the allele influencing preference for this host (a), as well as the alleles at the two loci which enhance survival on the new host (b, c), are maintained in the population on the original host at low frequency by mutation pressure. At this stage in the model we also assume that the genetic control of host choice by the host choice locus A,a is imperfect and that some errors occur. We expect that partially reproductively isolated host races will evolve at this stage but more conclusive results must await a more quantitative treatment now in progress. In addition, if genetic variability exists (or arises through mutation) at a fourth locus (D, d) capable of improving host choice still further, the alleles at this locus will be directly selected for non-random association (disequilibrium) with the ecologically related alleles already present at the other loci. Thus, by direct adaptive selection for non-random association between such ecologically coupled alleles (ABCD and abcd) host choice mechanisms may be perfected and consequently lead to speciation in these parasitic organisms.

At this point we wish to compare the attributes of the new species formed via the Felsenstein and parasite models in order to assess their biological relevance. In the parasite model, the initial population has been split to form reproductively and ecologically isolated species with each utilizing only one host. In the Felsenstein model, however, reproductive isolation occurs without any corresponding ecological isolation as individuals continue to ‘compete perfectly with each other’ (Felsenstein, 1981, p. 137). The closest real biological example which we can imagine fitting the attributes of this latter model would perhaps be some kind of plankton species which is incapable of influencing into which of several alternative habitats it is passively dispersed. If this species exhibited genetic variability at viability loci correlated with the different habitats and could mate assortatively by some mechanism such as seasonality, the conditions would satisfy the assumed biological properties inherent in the Felsenstein model. However, such a creature bears little resemblance to most of the parasitic organisms for which sympatric modes of speciation have been proposed (Bush, 1975; White, 1978).

We have summarized some of the major differences between the Felsenstein and parasite models of sympatric speciation in Table 1. Complementary to our discussion above, Taylor (1976) has pointed out the ways in which host selection (or more generally habitat or niche selection) may facilitate the maintenance of a polymorphism. This conclusion has been corroborated experimentally for Drosophila by Jones & Probert (1980). These results emphasize the importance of host selection to
Table 1. Major differences between the Felsenstein and parasitic sympatric speciation models. Both models of sympatric speciation are haploid and deterministic with discrete generations, assume infinite population sizes, no linkage (initially), equal niche (or habitat or host) resource sizes, and separate density-dependent regulation of population size within each niche. In habitat or host 1, host-related viability or survival alleles $BC$ are favoured, and in habitat or host 2, $bc$ is at an advantage, while intermediate genotypes ($Bc$ and $bC$) are of intermediate fitness in each habitat and of lower fitness on average in the total population of both habitats combined.

<table>
<thead>
<tr>
<th></th>
<th>Felsenstein model</th>
<th>Parasite model</th>
</tr>
</thead>
<tbody>
<tr>
<td>host selection</td>
<td>random</td>
<td>non-random</td>
</tr>
<tr>
<td>mating</td>
<td>assortative</td>
<td>on host</td>
</tr>
<tr>
<td>locus $A$</td>
<td>assortative</td>
<td>host</td>
</tr>
<tr>
<td>locus $D$</td>
<td>assortative</td>
<td>host selection</td>
</tr>
<tr>
<td>mating genes</td>
<td>mating modifier</td>
<td>selection gene</td>
</tr>
<tr>
<td>selection on $A$</td>
<td>initially weak</td>
<td>initially strong and direct</td>
</tr>
<tr>
<td>on $D$</td>
<td>very weak</td>
<td>strong</td>
</tr>
<tr>
<td></td>
<td>and indirect</td>
<td>and direct</td>
</tr>
</tbody>
</table>

Populations which utilize two or more hosts, and thus further question the relevance of conclusions derived from a model such as Felsenstein's which assumes that host selection is a purely random process.

One final weakness of the Felsenstein model which we wish to contrast with the parasite model concerns his use of assortative mating genes to bring about reproductive isolation. To begin with, it is not explicitly stated at what frequency the alleles at this locus are initially set during simulation runs with this model. And it has been pointed out elsewhere that a polymorphism for assortative mating in the initial population is likely to exist only under a restricted set of conditions (Moore, 1979). If an assortative mating allele is introduced into a population at an initially low frequency, individuals possessing this allele will be at a disadvantage due to the difficulty of finding appropriate mates (Wilson, 1982). While in the parasite model the individuals initially colonizing a new host may suffer a similar disadvantage due to absence of mates, they also experience a large counterbalancing increase in fitness due to the probable absence of competitors (Pimm, 1979; Gibbons, 1979), predators, and parasites on the new host.

There is no direct selection acting to increase a rare assortative mating allele which corresponds to these important ecological interactions. Instead, selection for assortative mating in Felsenstein's model is confined to act only by the reduction of intermediate genotypes $Bc$ and $bC$ whose fitness is lower on average than either $BC$. 302
and be (assuming that the viability loci interact multiplicatively with each other).
Thus the selection on the locus D,d which in Felsenstein's model modifies (enhances) the degree of assortative mating is even more indirect since it must, in turn, act via any disequilibrium already established between A and the B or C viability loci. This again contrasts sharply with the more direct effect brought about by the different alleles at the D,d host choice locus in the parasite model. Felsenstein (1981, p.132) justifies the choice of the particular assortative mating modification mechanism used in his model "... because it places no direct selection on the modifier, and in hopes that the result will prove to be insensitive to the specific modification scheme." In contrast, we have argued that this procedure lacks applicability in that the use of an assortative mating scheme at all (as opposed to a host choice mechanism) ignores important aspects of the biology of many organisms and especially parasites.

Despite these deficiencies, Felsenstein's model of sympatric speciation represents a considerable step forward in this field due to its thorough and quantitative approach to the modeling process. And it does serve to emphasize one very important aspect of the underlying genetic mechanism which we suspect will also apply to more realistic models of host shifts in parasitic organisms. This centers on the role of linkage among the loci involved in this process. Felsenstein found that a reduction in the recombinant fraction (i.e. increased linkage) between a viability locus (B or C) and the assortative mating locus (A) substantially reduced the selection intensity and degree of assortative mating necessary to bring about sympatric divergence. Linkage between the viability gene and the assortative mating gene also had the effect of promoting divergence in a similar model developed by Caisse & Antonovics (1978). On the other hand, increased linkage between the two viability loci (B, C) themselves made the restrictions on sympatric speciation considerably more severe in the haploid version of Felsenstein's model, but only slightly more severe in diploid versions of this model. Unfortunately, he did not examine the case where all three of these loci are linked.

Linkage relationships between host selection and survival genes can therefore play an important role in the sympatric divergence of populations utilizing different hosts or habitats. The closer the linkage the more likely distinct host races and eventually species will evolve. Such 'supergenes' arise as a result of chromosome rearrangements or transpositions by movable genetic elements which appear to be pervasive among organisms (White, 1973; Campbell, 1981). For example, it has been estimated that about one individual in five hundred is heterozygous for a new chromosomal rearrangement in organisms as diverse as lilies, grasshoppers, and man (White, 1978).

Although there are many ways linkage patterns can be established, it is not necessary to postulate that a chromosomal rearrangement bringing about tight linkage between these loci must arise de novo in each successive host shift. In the course of previous host shifts and speciation events in a parasite's evolutionary history, selection would have favoured close linkage of genes governing such ecologically correlated traits as host selection and host-associated viability differences. The lack of adequate genetic variability of linkage relationships among loci would not, therefore, be expected to exert serious constraints on the sympatric development of re
productive isolation. However, differences in patterns of linkage between host selection and viability genes may to some extent determine whether a parasite is monophagous, oligophagous, or polyphagous.

Other factors that influence host shifts

There are two ways a parasite can encounter and colonize a potential new host plant. Wind or some other agent can transport a parasite to a new area where it may, under favourable conditions, infest a new host plant. Alternatively, it can shift to a new host that has been introduced or has been growing in close proximity. Although long-distance transport and colonization of a new host have occurred, such events require special conditions. Not only must the insect be preadapted to recognize and survive on the new host, it must also be able to compete with the local fauna for resources. In parasites, many of which are monophagous (Price, 1980), it is far more likely that successful host shifts occur when a host plant is introduced into a new area free of its associated fauna and flora. Under these sympatric conditions local parasites would have opportunities to colonize and test their ability to survive on the new plant repeatedly during and after acquiring the right combination of host recognition and survival genes. Many attempts probably end in extinction and these events must be common following the introduction of a new host plant into an area (Bush, 1975).

There are several other genetic and ecological factors influencing host race formation that warrant attention. These include:
- the genetic structure and demography of a population such as degrees of inbreeding and effective population size (Templeton, 1981)
- the genetic control of diapause leading to allochronic isolation in association with host plant phenology (Tauber & Tauber, 1981)
- induction and other mechanisms of learning (Prokopy et al., p 123)
- the number of loci controlling host selection and survival abilities
- the taxonomic affinity of old and new host plants (Ehrlich & Raven, 1965; Bush, 1969) as well as their micro- and macro-abundance and distribution
- agricultural practices such as the widespread planting of monocultures and removal of defensive chemicals by selective breeding

In summary we conclude that in parasitic insects sympatric host race formation and speciation may occur frequently. But there is obviously a need for further empirical investigations and theoretical development that takes into account the unique biological qualities of host-parasite interactions.

References