TARGET REVIEW
The genic view of the process of speciation

CHUNG-I WU

Department of Ecology and Evolution, University of Chicago, Chicago, IL, USA

Abstract

The unit of adaptation is usually thought to be a gene or set of interacting genes, rather than the whole genome, and this may be true of species differentiation. Defining species on the basis of reproductive isolation (RI), on the other hand, is a concept best applied to the entire genome. The biological species concept (BSC; Mayr, 1963) stresses the isolation aspect of speciation on the basis of two fundamental genetic assumptions – the number of loci underlying species differentiation is large and the whole genome behaves as a cohesive, or coadapted genetic unit. Under these tenets, the exchange of any part of the genomes between diverging groups is thought to destroy their integrity. Hence, the maintenance of each species' genome cohesiveness by isolating mechanisms has become the central concept of species. In contrast, the Darwinian view of speciation is about differential adaptation to different natural or sexual environments. RI is viewed as an important by product of differential adaptation and complete RI across the whole genome need not be considered as the most central criterion of speciation. The emphasis on natural and sexual selection thus makes the Darwinian view compatible with the modern genic concept of evolution. Genetic and molecular analyses of speciation in the last decade have yielded surprisingly strong support for the neo-Darwinian view of extensive genetic differentiation and epistasis during speciation. However, the extent falls short of what BSC requires in order to achieve whole-genome 'cohesiveness'. Empirical observations suggest that the gene is the unit of species differentiation. Significantly, the genetic architecture underlying RI, the patterns of species hybridization and the molecular signature of speciation genes all appear to support the view that RI is one of the manifestations of differential adaptation, as Darwin (1859, Chap. 8) suggested. The nature of this adaptation may be as much the result of sexual selection as natural selection. In the light of studies since its early days, BSC may now need a major revision by shifting the emphasis from isolation at the level of whole genome to differential adaptation at the genic level. With this revision, BSC would in fact be close to Darwin's original concept of speciation.

Introduction

What are species? This is undoubtedly one of the big questions in biology. What this perspective intends to do is to formulate the genetic process of differentiation between diverging populations of sexually reproducing organisms. Some, but not all, aspects of differentiation

Correspondence: Chung-I Wu, Department of Ecology and Evolution, University of Chicago, Chicago, IL 60637, USA. Tel.: 773 702 2565; fax: 773 702 9740; e-mail: ciwu@uchicago.edu

may be manifested as reproductive isolation (RI) (e.g. divergence in the spermatogenic programmes may lead to hybrid male sterility). If that process can be understood, it may then be possible to define the earliest stage in which species can be considered as formed. According to the biological species concept (BSC; Mayr, 1963), such a stage can be objectively and naturally delineated when *genome-wide* RI is attained (defining RI on the basis of only a portion of the genome is logically incompatible with the central tenets of BSC; discussed later). This

position is based on specific assumptions about the genetic basis of biological differentiation between species. If the assumptions are not fulfilled, species may not be as objectively definable as BSC prescribes. From this view point, our current understanding of the molecular and genetic basis of the process of speciation will be presented. The possible stages in which species may be delineated will be briefly outlined.

There are many operational species definitions for practical use (Coyne, 1994; Harrison, 1998; and other chapters in Howard & Berlocher, 1998; Jiggins & Mallet, 2000). The biological relevance of these definitions would be clearer in the evolutionary context. This perspective is not to compare them, or to add another one, but to address the underlying process of speciation. A detailed treatment of the species concept in the light of this process will be given elsewhere.

It has also been long recognized that the genetics of speciation may be quite distinct between plants and animals (Grant, 1971). Most of the views presented here are based on animal studies. Some may not be applicable to plants (e.g. level of genetic divergence between nascent species; Bradshaw *et al.*, 1995; Doebley *et al.*, 1997) whereas others may be common knowledge among plant evolutionists (Rieseberg *et al.*, 1999; Rieseberg & Burke, 2001). In the animal literature, this perspective relies heavily on Drosophila studies. Although the general patterns of postmating isolation appear comparable across animal taxa (Wu & Davis, 1993; Wu & Palopoli, 1994; Orr, 1997; Presgraves & Orr, 1998), there is insufficient information on other characters. In fact, the sharp contrast in their genetic bases of sexual isolation between two Drosophila systems (Doi *et al.*, 2001; Ting *et al.*, 2001) serves as a sobering reminder on how little we know about the genetics of speciation. Finally, many of the conceptual issues addressed here have been reviewed recently from a different (Coyne & Orr, 1998) or similar (Schilthuizen, 2000) perspective.

Genes as the unit of species differentiation

That genes, or complexes of genes, are the units of evolution has been a widely adopted perspective in Biology. The study of speciation, however, is a conspicuous exception. Because the speciation process is still conceptualized at the level of individual (or whole genome), BSC which defines species as 'groups of interbreeding natural populations that are reproductively isolated from other such groups' (Mayr, 1963) remains the gold standard. To see the difference between genes and individuals, consider the simpler case of random-mating Mendelian population, which is meaningful only at the genic level. For example, human populations may be panmictic at most loci and, indeed, mating is essentially random with respect to the ABO or MN bloodtype and most microsatellites. Nevertheless, mate choice in human is distinctly nonrandom at the individual level. Mating with respect to loci that govern aspects of human morphology, such as height, is likely nonrandom. It may be possible that, at a great majority of loci, mating is random but a very tiny fraction of the genome is more than sufficient to dictate highly nonrandom mating at the individual level.

Like random mating within population, differentiation between populations is more meaningful at the genic than individual level. Imagine two hypothetical populations of grasshoppers on the north and south-facing slopes of a mountain range, which does not impede migration. Assume that the A, B and C alleles of three different loci are adapted to the northern slope and the a, b, and c alleles, to the southern slope. If local selection is sufficiently strong relative to gene flow, the two populations would differentiate at the three loci. Then, what about the rest of the genomes, which are equally fit in both habitats? A very low level of migration, down to some individuals per generation, is sufficient to prevent population differentiation (Crow & Kimura, 1970; Slatkin, 1987). Therefore, under most parapatric conditions, the genomes of the two populations would be mosaic in their extent of differentiation. Genomic regions very near the three loci of differential adaptation would be differentiated whereas the rest would not.

An important distinction can, therefore, be made between two classes of loci – those that directly affect differential adaptation and those that do not. Differential adaptation is a well defined form of divergence in which the alternative alleles have opposite fitness effects in the two populations (because the migration of these alleles to the other population is maladaptive, differential adaptation also implies restricted gene flow). It is plausible that, during the process of speciation, genes of differential adaptation would account for only a small fraction of the genome. I shall refer to them as 'speciation genes' and the rest as 'marker loci'. As genes directly responsible for differential adaptation and RI are rarely known except in some cases (Lee & Vacquier, 1992; Lee *et al.*, 1995; Metz & Palumbi, 1996; Ting *et al.*, 1998; Wang *et al.*, 1999), most of the genetic studies on species or population differentiation are based on the marker loci, including most allozymes, microsatellite and mitochondria DNA. Although the marker loci are informative about demographic events (Avise, 1994), the whole process of speciation depends primarily on the speciation genes (e.g. Wang *et al.*, 1999; Ting *et al.*, 2000).

Separating the two classes of genes, Fig. 1 outlines the very basic features of the process of speciation at the genic level. For simplicity, we may assume that the populations come into contact at the specified stage (although the specific mode does not really matter). In stage I, population differentiation has taken place at a small number of loci responsible for functional divergence. This could be in their mate recognition system as
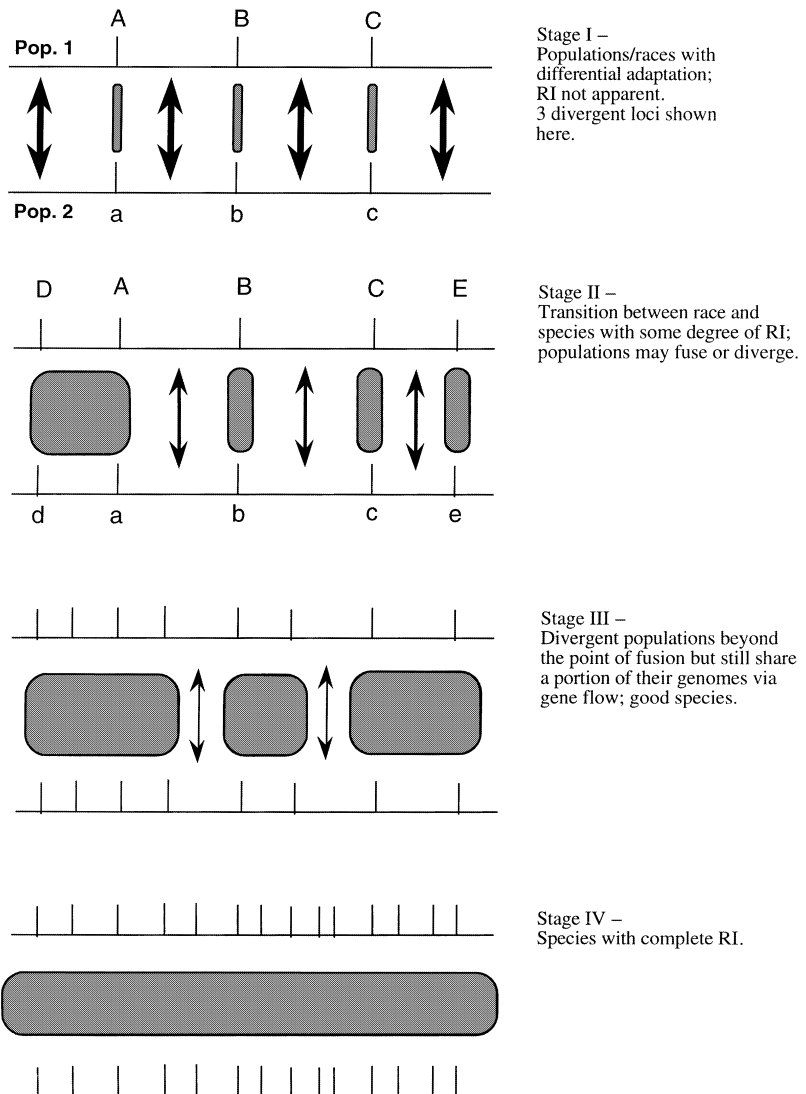
Fig. 1 Conceptualized stages of species differentiation at the genic level – illustrated are chromosomes of diverging populations, which come into contact at the specified stage and experience net gene flow in parts of their genomes. Only loci directly contributing to differential adaptation are shown and alleles at such loci are assumed to be differently fixed (A, B, C, etc. in population 1 and a, b, c, etc. in population 2). As the populations diverge, there will be more loci involved in differentiation. Between these differentiation loci, there are certainly many more sites of neutral divergence, which can be shared between populations. Double-headed arrow represents net gene flow (with migration overwhelming local selection) and is shown to decrease in magnitude as speciation progresses. Shaded block indicates the prevention of gene flow by local selection at and near the loci of differential adaptation. The fate of diverging populations (fusion or not) at a given stage can be inferred if there are sufficient data on the extent of divergence; see text for details.

in the Zimbabwe – cosmopolitan behavioural races of Drosophila (Wu *et al.*, 1995) or the mimicry colour patterns of the Heliconius butterflies (Mallet *et al.*, 1998). Upon secondary contact, gene exchange would be extensive but, at the loci of functional divergence, exchange would be restricted. There is little disagreement that the populations belong to the same species because intermediate forms would be common in the absence of extrinsic barriers (for this stage to be considered as incipient speciation, it is implicitly assumed that there do not exist intermediate environments to which the intermediate forms are equally or better adapted; otherwise, it would be a simple case of population differentiation not germane to speciation). As generally recognized, a species may not necessarily be a homogeneous evolutionary unit (Butlin, 1998). For example, the two behavioural races of *Drosophila melanogaster* have

diverged at more than 15 loci governing mating behaviours (Hollocher *et al.*, 1997b; Ting *et al.*, 2001) with mild fitness reduction in the hybrids (Alipaz, 2000).

In stage II, further differentiation has occurred and genic substitutions within each population could be coadapted in the physiological sense. As a result, intergroup crosses would yield some genotypic combinations that are hybrid inviable or sterile, as is the case of the Bogota vs. mainland populations of *D. pseudoobscura* (Prakash, 1972; Orr, 1989; Schaeffer & Miller, 1991; Wang *et al.*, 1997). Alternatively, divergence may not have yielded hybrid incompatibility; instead, the ecological or behavioural divergence is so strong that hybrids are not as fit as the parents in their natural environments. The butterflies *Heliconius himera* × *H. erato* (Mallet *et al.*, 1998), the Rhagoletis host races (Feder *et al.*, 1997; Berlocher, 1998), the aphid races (Via *et al.*, 2000),

Darwin's finches (Grant & Grant, 1998) and the stickle-backs (Schluter, 1998) may be such examples. It has been common to distinguish the nonadaptive hybrid inviability/sterility from the ecological/behavioural misfit (Rice & Hostert, 1993). This distinction is rather artificial because it presupposes the genic basis of, say, gametogenic divergence (which results in hybrid sterility) is nonadaptive. As discussed later, this presupposition has neither conceptual nor empirical basis. The other assertion that the physiological incompatibility is more genetically 'hardwired' than the ecological divergence has not been tested either.

In stage II, a very substantial portion, or even a great majority, of genomic regions would remain shared and undifferentiated. Populations at this stage are sometimes given the subspecies status to reflect the ambiguity in their species status. Will they diverge or fuse? It is possible that, without ecological perturbation, they may continue to diverge to become fully fledged species. However, upon massive hybridization, the subspecies or species may fuse and emerge as a new hybrid form, or hybrid swarm (Mayr, 1963, p. 521). Habitat destruction or land bridge formation, for example, can facilitate such massive hybridization. It would be most interesting to see whether recently evolved species groups such as the African cichlid flocks (Meyer, 1993) still retain the capacity to fuse.

In contrast, by stage III, the diverging groups have passed the point of no return and it is inconceivable that they will ever fuse (the projection of nonfusion does not entail 'seeing the future'; a thorough understanding of the genetics of differentiation and RI may suffice). In complete sympatry in any habitat, they will either coexist or experience competitive exclusion. In parapatry, a narrow hybrid zone may be common. At this stage, the accumulation of 'speciation genes' has resulted in extensively divergent (between populations) and coadapted (within population) gene complexes. The groups are thus divergent in at least some aspects of reproductive biology, sexual behaviour, and/or morphology. Although there would be no disagreement over their species status, the genomes of these good species are by no means completely isolated. Gene-sharing by introgression could persist for a long period of time and effectively retard or nullify species differentiation over some portion of the genome. Many closely related 'good species' in allopatry or sympatry are at this stage. *Drosophila simulans* and its two sibling species, *D. mauritiana* and *D. sechellia*, are one such example – premating isolation is incomplete and hybrid females in F1 and subsequent generations are often fertile (Coyne, 1984; Lachaise *et al.*, 1988; Wu & Palopoli, 1994). The two species of Bombina toads that maintain a stable hybrid zone may be another example (Szymura & Barton, 1991). Stage II and III are where the concept of RI is ambiguous and has been inconsistently applied.

Finally, complete RI has evolved and the two gene pools would cease sharing alleles at any part of their genomes by means of breeding. At stage IV, premating isolation is often strong in nature and F1 hybrids of both sexes are usually inviable or sterile. When F1 or F2 are fertile, occasional introgressions are quickly eliminated by selection. A stage IV, genetic analysis is not possible except with some 'trickeries' (Muller & Pontecorvo, 1942; Orr, 1992; Davis *et al.*, 1996; Coyne *et al.*, 1998) and the level of functional divergence at the inception of this stage can be extremely high (Sawamura *et al.*, 2000).

The genic view of species

In the light of the genic basis of the full process of speciation, when do we consider species formed? Would the delineation of species be like the demarcation between adult and child – there simply is not a clearcut transition between phases? At the extreme, one may insist on the strict application of BSC with complete RI across the whole genome; thus, only stage IV of Fig. 1 would fulfil the criteria. The concept of complete isolation does not distinguish loci governing traits of speciation from other loci. By this concept, it is dubious whether *D. simulans* and *D. mauritiana* could even be classified as species because of the continual gene flow in the past [Solignac & Monnerot, 1986; Ballard, 2000; Ting *et al.*, 2000; were the two species in sympatry at present, gene flow would likely be much higher than estimated (Kliman *et al.*, 2000)]. Given the overwhelming evidence of functional divergence (Coyne, 1984; Coyne & Charlesworth, 1989; Davis *et al.*, 1994; Wu & Palopoli, 1994; Coyne, 1996; Davis & Wu, 1996; Hollocher & Wu, 1996; True *et al.*, 1996, 1997; Zeng *et al.*, 2000), declassifying them as good species seems farfetched.

The essence of any evolution-based species concept should be as follows: speciation is the stage where the populations will not lose their divergence upon contact and, furthermore, will be able to continue to diverge. Thus, the very essence does not have to include RI. In other words, speciation is the stage where (i) the gene pools at loci of differential adaptation would not mix even when the extrinsic barriers to gene exchange are removed; and (ii) the spread of advantageous mutations are sometimes (or often, but not always) restricted to the population in which they originate. Because the rest of the genome may mix upon contact, the genomes of the diverging populations are expected to be mosaic in their extent of differentiation.

It needs to be emphasized that mere differential adaptation, including that along a cline, does not constitute species in this genic view. The two central criteria are: (i) not losing the differentiation, and (ii) being able to continue to diverge. Therefore, the production of heterozygotes and all sorts of hybrids, which can either transfer the genes of differential adaptation between populations or form a self-sustained hybrid population with a mixture of adaptive characters, certainly represents the loss of differential adaptation. Most geographical

races, or polytypic species, are in this category and belong to the same species. In this genic view, our ability to judge whether the divergence will be lost upon contact, and hence our ability to determine species status, is a function of our knowledge of the genetics, ecology and reproductive biology of the diverging populations in question. It seems futile to attempt to devise a generally applicable species definition that demands little prior knowledge of the genetics and ecology of the taxa.

The evolutionary trajectory (fusion vs. speciation) of populations of stage II in contact depends on how strongly genes are coadapted within each nascent species and how often the coadapted gene complexes are broken up. Both the genetic architecture of differentiation (the number of loci involved, the degree of fitness reduction associated with various gene combinations, etc.) and the extent of gene flow between nascent species are crucial. This is an area where theories may figure prominently in the future (Felsenstein, 1981; Barton & Hewitt, 1985; Charlesworth *et al.*, 1987; Barton, 1992; Orr, 1995; Turelli & Orr, 1995; Gavrilets, 1997, 2000; Kondrashov *et al.*, 1998; Dieckmann & Doebeli, 1999; Kondrashov & Kondrashov, 1999). The models of Charlesworth *et al.* (1987) and Gavrilets (1997, 1999, 2000) are especially relevant to the concept developed here. The genetic architecture of differential adaptation also determines the fate of new advantageous mutations. At stage II, some advantageous mutations may permeate through the population boundary, making no contribution to their further differentiation but others may interact positively with genes of the same population and negatively with those of the other population. Such mutations will spread in only one population, contributing to further adaptive divergence.

We may summarize the genic view of species as follows: species are groups that are differentially adapted and, upon contact, are not able to share genes controlling these adaptive characters, by direct exchanges or through intermediate hybrid populations. These groups may or may not be differentiated elsewhere in the genome.

BSC – the whole genome as a cohesive unit

BSC stresses the role of isolation (first geographical and extrinsic, then biological and intrinsic) in speciation. During the process, and at the conclusion, of speciation, the genomes of the diverging populations are entirely insulated from each other. RI as a whole-genome concept has been amply underscored in the seminal writings on BSC (Dobzhansky, 1937; Mayr, 1963, Chap. 17). Recently, in order to reconcile the observations of gene flow between nascent or even good species, it has been suggested that RI may not have to be a whole-genome phenomenon. However, if we allow more than a trivial portion of the genome to permeate through the boundary between nascent species, then RI would lose its logical robustness. For example, if two taxa are exchanging genes over 30% of their genomes, are they reproductively isolated? How about 60%? The whole-genome cohesiveness of BSC in fact derives its logical consistency from the three tenets given below.

The first tenet (that speciation is a whole-genome phenomenon) rests on two strong assumptions about the genetic architecture of speciation, namely the number of genes involved is large and the genetic changes within each species are strongly coadapted, forming a cohesive unit (Dobzhansky, 1937; Mayr, 1963). In an influential address, Mayr (1959) challenged the genic view of species differentiation and labelled it as 'beanbag genetics'. The genic view of population genetics mainly deals with simple genetic changes and ignores genic interactions whereas, Mayr contended, genic interaction as the norm of the genetics of speciation. This is where the high-resolution genetic analyses of the last decade become most relevant (see Wu & Hollocher, 1998, for a review).

The second tenet, that complete RI and cessation of gene flow between diverging populations is the prerequisite for speciation, is a logical extension of the first tenet. Coadaptation of genes across the whole genome would certainly preclude the sharing of random fragments of genomes between species. Conversely, exchanging genomic fragments would also jeopardize the integrity of the diverging genomes. A genic concept of speciation, such as the one outlined in Fig. 1, would be incompatible with the two tenets. Indeed, Mayr (1963, Chap. 6) devoted a whole chapter on the rarity of breakdown of isolating mechanisms in nature, which he suggested is the prima facie evidence for genomic cohesion.

The third tenet, implicit in BSC, pertains to the stage at which speciation is considered complete. As the process of speciation is continuous, one might expect demarcating the stages 'before and after' speciation subjective, as Darwin (1859) noted. BSC is able to prescribe a transition in this process because it assumes that the whole-genome coadaptation, and hence RI, evolves rapidly once it gets started. In this view, there exists a natural transition, which characterizes speciation.

Conflicts (gene vs. genome and adaptation vs. isolation) and consequences

Although BSC assumes a highly coadaptive genetic architecture, which led directly to the conception of whole-genome isolation, RI was always considered as a by product of differential adaptation by Mayr (1963, p. 548–554). The process of speciation since Darwin (1859) has been about differential adaptation to different environments or diverging mating systems (Darwin, 1875). Although Darwin (1859, Chap. 8) did discuss RI and recognized its role in preventing diverging populations from becoming homogenized, he viewed RI as a

secondary phenomenon. In BSC, because whole genome coadaptation and genome-wide isolation are treated as two sides of the same coin, it is only necessary to study either, but not both, in order to understand speciation. Indeed, RI eventually became the dominant side and the thinking of adaptation was largely eclipsed as BSC subtly evolved. This imbalance has brought in several popular, albeit somewhat illogical, concepts about speciation.

(1) Non-adaptive 'causes' of speciation – In a popular interpretation of BSC, the 'causes of speciation' denotes the event that completes RI. 'Speciation genes' thus refer to those that cause RI and are sometimes assumed to belong to a special class of genes unrelated to normal adaptation. On the other hand, if RI is viewed as a by product of differential adaptation as portrayed in Fig. 1, there is really no 'causes' of speciation; speciation genes are simply genes that determine aspects of differential adaptation and whose by product happens to be fitness reduction in the hybrids.

There are indeed a number of special genetic events whose primary phenotype appears to be RI. These events are often nonadaptive for the host's genome (or at least of no known adaptive value). The invasion of cytoplasmic symbionts causing hybrid incompatibility (Werren, 1998) is the primary example. When a population of host acquires a certain cytoplasmic symbiont such as the bacteria Wolbachia, crosses between males of this population and females of an uninfected one yield inviable progeny. If this second population later acquires a different symbiont of cytoplasmic incompatibility, then RI will be complete between them at the time of invasion of the second symbiont. Mayr's lack of enthusiasm for such an event is entirely consistent with his emphasis on differential adaptation (Schilthuizen, 2000). Transposable element invasion and meiotic drive divergence may be other such nonadaptive examples (Hurst & Pomiankowski, 1991; Hurst & Schilthuizen, 1998). Chromosomal changes, including translocations, inversions and polyploidization, have also occupied a large section in the speciation literature (White, 1954). The interest has been their potential in causing problems in meiosis in interspecific hybrids and, hence, postmating RI. Whether these changes may or may not be advantageous is not known and was not the primary reason for the interest. Indeed, it is now feasible to create lines in model organisms that are reproductively isolated from the rest of the species by transforming them with a toxin gene and its repressor. The transformed lines are normal but hybridization with the untransformed lines will result in hybrid inviability because of the derepression of the toxin gene. Will we then have created a new species? The suggestion seems absurd in the sense of biological differentiation but does not contradict the isolation concept.

In the Darwinian view, the genetic events discussed above are special cases because they do not result in differential adaptation to different environments or sexual systems. In BSC, they are an important category; in fact, they may be the only known causative genetic events of speciation.

(2) Does selection or drift drive speciation? In the view of differential adaptation, speciation is driven by the same forces that drive adaptation to changing environments or mating structure. It is essential to define the adaptive characters and the underlying genes. In the case of postmating isolation, the challenge is most revealing. As hybrid inviability and sterility are negative by products, what could be the normal function and selective advantage of the underlying genic changes? How does RI arise as a correlate of the genetic divergence? These questions can only be approached after genes have been identified.

In the last two decades, the formulation of the question of postmating isolation is strikingly different as it follows the isolation concept. The possible normal function and selective advantage are not considered, the only selection being that against the inviable or sterile genotypes. Most efforts have thus been directed towards the question: how can mutation and genetic drift drive the underlying genetic changes without being retarded by evolving into the unfit genotypes? To state it simply, in the total genotypic space, two pure species occupy different adaptive peaks and the space in between are filled with unfit genotypes, representing F1, F2 breakdown, etc. How did evolution proceed from one adaptive peak to the other? There are many solutions to this problem (Dobzhansky, 1937; Muller, 1940; Nei et al., 1983; Cabot et al., 1994; Orr, 1995; Turelli & Orr, 1995, 2000; Gavrilets, 1997). The purpose of this perspective is not to review them but to point out that few incorporate adaptive advantage into the model.

Not knowing what the normal functions of these genes may be, some authors have focused on genetic drift (e.g. Nei et al., 1983; Gavrilets, 1997) whereas others attempt to address the effect of the fixation of speciation genes on patterns of RI and bypass modelling the process of fixation itself (e.g. Turelli & Orr, 1995, 2000). Charlesworth et al. (1987) presented a pioneering effort at modelling RI as a by product of adaptive evolution by contrasting the evolutionary dynamics of X-linked vs. autosomal genes. Recently, Gavrilets (1999, 2000) and Gavrilets et al. (2000) have started developing explicit models incorporating differential adaptation, which speeds up speciation well beyond what mutation-drift can sustain. We shall return to the molecular evidence in the end.

That genetic drift has played a big role in modelling RI is more by default and expediency than by reality. It is somewhat ironic that, even silent substitutions are increasingly believed to be non-neutral (Li, 1987; Akashi, 1997), changes that can cause severe fitness reduction in some genetic backgrounds are widely suggested to be driven by genetic drift (that the modelling is a legacy of Dobzhansky's (1937, 1970) genetic analysis on speciation is doubly ironic). Note that the

introduction of genetic drift into these population genetic models of RI is for very different reasons than those of Mayr (1963) or Carson & Templeton (1984) in their proposal of genetic revolution through founder populations, the latter being about evolution from one to another local adaptive peak.

The quest to explain the origin of RI may be analogous to the attempt at explaining the evolution of the Major Histocompatibility Complex (MHC), which was discovered as molecules for graft rejection in transplantation surgery. It would not have been possible to understand the evolution of graft rejection without first knowing about MHC and its adaptive function in immune surveillance. Graft rejection, like postmating isolation, is a by product and its origin cannot be understood on its own.

(3) Gene flow during speciation – in the strict isolation concept, any degree of gene flow is perceived to be disruptive of genome cohesiveness and capable of reversing the process of speciation. The studies of hybrid zones (Barton & Hewitt, 1985; Harrison, 1990, 1993; Szymura & Barton, 1991; Arnold & Emms, 1998; Butlin, 1998; Mallet *et al.*, 1998) and sympatric speciation (Bush, 1975, 1998; Schliewen *et al.*, 1994; Feder *et al.*, 1997; Berlocher, 1998; Dieckmann & Doebeli, 1999; Kondrashov & Kondrashov, 1999), in conjunction with the genetic analysis of RI, may be gradually changing that view (Schilthuizen, 2000). Gene flow will be discussed further.

Resolution of conflicts

As BSC evolved in the years to account for new observations, two tiers of conflicts have built up within the concept itself. These conflicts are now in serious need of resolution. The first conflict is whether to view the gene or the genome as the unit of speciation; in other words, whether each genetic substitution would interact with future changes everywhere in the genome. The second conflict is whether to view speciation primarily as the process of differential adaptation or the evolution of RI. In its original formulation, BSC was a 100% genomic concept. It also viewed RI as a by product of differential adaptation. In the recent past, the whole genome concept could not be abided by and, by expediency, RI has usually been studied without reference to adaptation.

The attempts to resolve these conflicts bring us to the fine-mapping and cloning of genes of RI. Whether the gene or genome is the unit of speciation can only be answered when the speciation history of genes of RI can be contrasted with those of the rest of the genome (Hilton *et al.*, 1994; Palopoli *et al.*, 1996; Wang *et al.*, 1997; Ting *et al.*, 2000). More importantly, if RI is a by product of differential adaptation, the underlying genes should reveal the process and mechanisms of adaptation. This is especially true for genes of hybrid male sterility which is the manifestation of divergence in the male
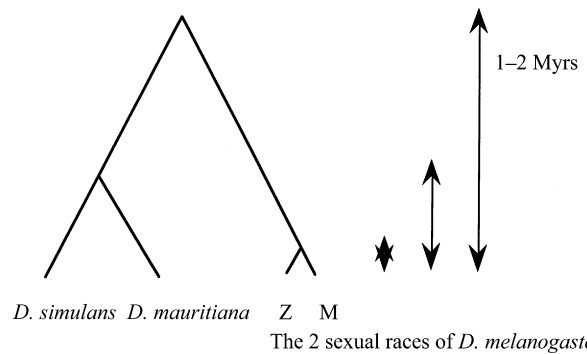

Fig. 2 The levels of divergence among the three *Drosophila* species most intensely analysed for the genetics of speciation. There are two sexual races (Z and M) in *Drosophila melanogaster*.

reproductive system, and its evolution is probably driven by sexual selection (Eberhard, 1985; Coulthart & Singh, 1988; Wu & Davis, 1993; Wu *et al.*, 1996; Presgraves & Orr, 1998; Ting *et al.*, 1998; Wu & Hollocher, 1998).

In this perspective, I shall rely mainly on the genetic analysis among the four species in the *D. melanogaster* clade but the conclusion does not depend only on these studies. In Fig. 2, three levels of divergence are illustrated. The lowest one is between the two behavioural races, Z and M, of *D. melanogaster*, which are at stage I of Fig. 1. The next level is between *D. simulans* and *D. mauritiana*, which are at stage III and represent the most extensively analysed pair of species for the genetics of RI. Both are sibling species of *D. melanogaster* which provides the genetic information necessary for the analysis of their hybridization. *Drosophila simulans* is cosmopolitan in distribution and *D. mauritiana* is endemic to the island of Mauritius. There is not report of *D. simulans* on Mauritius but this species is present abundantly on the nearby island of Reunion. F1 hybrid females do not suffer much in viability or fecundity and are inseminated by males of the pure species with little discrimination. F1 males are sterile. The third level is between *D. melanogaster* and *D. simulans*, which produce only sterile or inviable hybrids and are in stage IV.

The first conflict: is the gene or genome the unit of speciation?

The conflict centres around the two tenets of BSC: (i) the genomes between species are extensively divergent; (ii) little genetic exchange is possible during incipient speciation. Surprisingly (i) is probably right but (ii) is not, as described below.

(i) Extensive genetic divergence between species and coadaptation within species
In Table 1, I summarize the extent of genetic divergence underlying different traits of speciation and at various

	Stage I Z–M races*	Stage III D. simulans–D. mauritiana†	Stage IV D. simulans– D. melanogaster§
Sexual behaviours	≥15	–	–
Male sterility	≈0	≥120	>200
Inviability or female sterility	≈0	<10	>10
Genitalia morphology	≈0	≥19‡	–

Table 1 The estimated number of genes contributing to different traits of speciation at levels of species divergence shown in Fig. 2. A gene is counted only when it has a major effect at least in some genetic backgrounds. Stage I, III and IV refer to those defined in Fig. 1.

Dashes indicate no data of comparable resolution.

*Hollocher *et al.* (1997b), Alipaz (2000), Ting *et al.* (2001), Takahashi *et al.* (2001).

†Coyne (1984), Coyne & Charlesworth (1989), Wu & Palopoli (1994), Cabot *et al.* (1994), Davis *et al.* (1994), True *et al.* (1996), Coyne (1996), Hollocher & Wu (1996), Davis & Wu (1996), Ting *et al.* (1998).

‡True *et al.* (1997), Zeng *et al.* (2000).

§Muller & Pontecorvo (1942), Coyne *et al.* (1998), Sawamura *et al.* (2000).

levels of speciation. All these are sibling species. There are three salient features about the studies summarized in Table 1. First, the number of genes involved in speciation is substantial, even between sibling species. Secondly, the extent of gene interaction underlying RI was also unexpected (not shown in Table 1; see Palopoli & Wu, 1994; Cabot *et al.*, 1994). In fact, there is no real evidence in the literature for the simplest genetic model of postmating isolation, the so-called Muller-Dobzhansky model in which only one gene from each species interacts to cause RI (Palopoli & Wu, 1994; Wu & Palopoli, 1994; Perez & Wu, 1995; Turelli & Orr, 1995, 2000; Wu & Hollocher, 1998). Conspecific and heterospecific genic interactions are both indispensable for hybrid incompatibility, a feature important in later discussions on introgression. Thirdly, different speciation traits have evolved at highly disparate rates, notably the difference between hybrid male sterility and hybrid inviability/ female sterility. This is an indication that RI is not driven only by mutation and genetic drift; each trait must be experiencing different selective pressure.

The above results have indeed been surprising to 'beanbag' geneticists. Mayr's assertion that species represent thoroughly reconstituted genomes is indeed closer to the experimental evidence than many geneticists' belief that speciation is largely the result of some major changes (e.g, Goldschmidt, 1940; Nei, 1975; Raff & Kaufman, 1983; Brakefield *et al.*, 1996). However, this triumph is insufficient to support the tenets of BSC that whole-genome RI is absolutely central to the process of speciation.

(ii) Genetic exchanges during incipient speciation – speciation genes vs. the rest of the genome

Imagine the extreme case where all loci freely recombine. It is obvious that genes not directly involved in differential adaptation and/or RI can permeate through the nascent species boundary without impediment. Even if we assume that as many as 10% of the genes in the genome are involved in speciation at the incipient stage,

the bulk of the genome should still resume to sharing a common gene pool once the diverging populations come into contact. Of course, genes exist in linkage groups. Hence the key to the question of the cohesiveness of the entire genome is linkage and recombination.

Consider stage II of Fig. 1 where alleles at loci A–E are differentially adapted between the two diverging populations. When allele A is introgressed from population 1 to population 2, it will be gradually eliminated because of reduced fitness. Whether alleles at other neutral loci co-introgressed with A will be able to persist in population 2 depends on their being dissociated from allele A before elimination. Therefore, selection intensity against the introgressed alleles and recombination distance are both important.

Let the migration rate into population 2 each generation be m. Because of selection against speciation genes from population 1, their effective migration rate would be extremely low. Other genes linked to them will have a reduced effective migration rate, m'. In the simplest case of one speciation gene and one neutral locus that recombine with the rate r (Barton & Bengtsson, 1986; N. Takahata, pers. comm.),

$$m' = \frac{mr(1 - s)}{s + r(1 - s)} \tag{1}$$

Note that as long as four $Nm' > 1$, where N is the effective size of population 2, divergence between populations at the neutral locus would not be able to continue (Crow & Kimura, 1970; Slatkin, 1987). Therefore, unless selection against the introgression of speciation genes is very strong, only a small fraction of the genome near them would be affected. The question then is how strong such selection is, and how large the region affected would be.

Selection against genes associated with RI depends on the underlying genetic interactions and is generally much weaker than might have been conjectured. A single gene upon introgression into another species rarely, if ever, induces hybrid sterility or inviability in

animals (Wu & Palopoli, 1994 and references cited in Table 1). As a general rule, a specific constellation of genes are jointly needed to give rise to hybrid incompatibility. For example, if the A and B allele of Fig. 1 would jointly cause 90% hybrid sterility, A or B by itself might reduce fertility by <1%. This is the evidence for strong epistasis among conspecific genes cited earlier. In other words, many genes that are part of hybrid incompatibility interactions can in fact be individually introduced into another species when the constellation is broken by recombination. Although each may not be fixed upon introgression, it is eliminated slowly, allowing other cointrogressed genes to be dissociated from it.

In summary, all three parameters of eqn (1) have to assume rather extreme values (strong selection against introgressed genes, little migration and low recombination) to make the whole genome 'congeal' during speciation. Such theoretical arguments rely on many assumptions and we shall turn to empirical observations below.

From Fig. 1, one would expect that, at the incipient stages of speciation, the diverging genomes are mosaic with respect to the genealogies of different gene regions. This can often be indirectly inferred. For example, loci pertaining to the phenotypic characters of human racial groups are likely to be strongly geographically patterned whereas most loci unrelated to such characters exhibit little differentiation among racial groups (Nei & Roychoudhury, 1993; Barbujani *et al.*, 1997). The behavioural races of *D. melanogaster* also show strong incongruities between the phenotypic and genic differentiation. Although central–southern African lines can be distinguished from the cosmopolitan ones by their mating preferences (Wu *et al.*, 1995; Hollocher *et al.*, 1997a) such differences are largely determined by the autosomes (Hollocher *et al.*, 1997b), variations at most autosomal loci overlap extensively among populations (Tsaur *et al.*, 1998; Andolfatto, 2001). Endler (1977) reviewed many cases of strong spatial differentiation in spite of gene flow. Recent Drosophila studies suggest that the divergence in phenotypic characters can occur within a few kilometers (Capy *et al.*, 2000) or even shorter distance (Korol *et al.*, 2000), apparently in the presence of substantial gene flow. Lewontin & Krakauer (1973) have earlier suggested that different degrees of population differentiation among random loci may indicate local adaptation for the more highly differentiated ones. This approach has been used by Wang *et al.* (1997).

The diverging genomes that are mosaic in their gene genealogies can be observed at the species level as well. For example, Darwin's finches have been exchanging genes between species whose divergent morphological characters are stable (see Grant & Grant, 1996, 1998). Although the sharing of DNA variation at random loci apparently did not reverse the divergent evolution in phenotype among species, the strict adherence to the whole-genome isolation concept has made some authors question the species status of Darwin's finches (Freeland & Boag, 1999). Other examples may include the benthic and limnetic species of sympatric sticklebacks (Schluter, 1998; Rundle *et al.*, 2000) and the Rhagoletis species group (Berlocher, 2000).

The above examples are indirect inferences because the genes underlying the phenotypic divergence have not been identified. Direct evidence can only be obtained with the prior identification of the speciation gene, followed by a population genetic analysis near it. The Ods gene of hybrid male sterility between *D. mauritiana* and *D. simulans* is one such example. In the coding region of Ods, gene genealogies sort cleanly by species and *D. simulans* is unambiguously clustered with *D. mauritiana*, with *D. sechellia* as an outgroup (Ting *et al.*, 2000). This is not true for most other genes that have been surveyed so far (Caccone *et al.*, 1996) suggesting that the Ods speciation gene indeed has a distinct genealogical history. What is most interesting is that the DNA pattern merely 2 kb away shows a genealogical history like the rest of the genome. That a speciation gene would leave such a small footprint of influence on the evolutionary dynamics of the genome is commensurate with theoretical considerations (Maynard Smith & Haigh, 1974; Fay & Wu, 2000) as well as other empirical observations. For example, one of the key loci affecting the domestication of maize bears the signature of artificial selection in the regulatory region but not the coding region (Wang *et al.*, 1999). In a recent study aimed at identifying the gene responsible for the pheromonal difference between *D. melanogaster* populations, Takahashi *et al.* (2001) were able to show strong differentiation between races at a desaturase locus. Again, this differentiation is restricted to a region of only a few kb.

A rigorous proof of the genic nature of speciation shown in Fig. 1 demands three pieces of evidence: (i) the genome is mosaic with respect to the speciation history of each gene; (ii) variations at speciation loci are not shared between diverging species whereas other variations may permeate through the nascent species boundary; (iii) species divergence at the speciation loci tends to be more ancient than at other loci, inferrable from silent substitutions of these genes. Without (iii), the pattern of (ii), i.e. the sharing of variation at random loci, can be accounted for by ancestral polymorphisms, whereas the lack of sharing at speciation loci can be explained by selective sweep. The third piece of evidence has rarely been demonstrated and is only partially evident between *D. mauritiana* and *D. simulans* (Ting *et al.*, 2000, and references therein).

Another type of observation germane to the genic view of speciation is the natural mixing of species genomes upon secondary contact. In plants, such mixing has been shown to have the potential of creating new species, or recombinational speciation (Rieseberg *et al.*, 1995). Whether this has happened in animals can only be speculated (Grant & Grant, 1998) although the creation

of dog breeds and other domesticated animals may provide a clue (Darwin, 1859). Most of the documented genome mixing in animals is the introgression of genetic materials between species. Although some of the introgressions may be advantageous, for example, the t-complex between Mus (Hammer *et al.*, 1989) or a chromosomal inversion between Anopheles species (Besansky *et al.*, 1994), the main question is whether the bulk of the neutral variations can remain shared. The large body of literature on hybrid zones provides some of the best evidence of gene sharing across race or species boundary (Barton & Hewitt, 1985; Harrison, 1990, 1993; Shaw *et al.*, 1990; Szymura & Barton, 1991; Arnold & Emms, 1998; Butlin, 1998; Mallet *et al.*, 1998). Nevertheless, the true extent of introgression can be difficult to gauge. When there is no divergence at any locus in the entire study range, it is often conservatively attributed to the lack of divergence before secondary contact although the homogenization after the contact could produce the same pattern. Narrow hybrid zones may also tend to be cases of low introgression whereas very broad hybrid zones may be of less interest to investigators. In some studies, even allozymes themselves were suspected to be under local selection (Szymura & Barton, 1991).

In summary, the diverging genomes during (or even after) speciation can be quite 'porous' with respect to gene flow at nonspeciation loci.

The second conflict: is RI the byproduct of differential adaptation?

In the Darwinian view, RI is a by product of differential adaptation, whereas BSC sees the former as the prerequisite for the latter. To resolve the conflict over such a fundamental concept, we may choose either of the two approaches. First, we can analyse the genetics of adaptive differences between species or races and then study the possible roles the underlying genes may play in RI (Jones, 1998; Macnair & Gardner, 1998). The approach is of great value in its own right but the proportion of adaptive differences that have pleiotropic effects on the fitness of hybrids remains largely unknown.

The second approach, which is more direct, is to clone genes that are responsible for RI and determine whether and how the changes were driven by adaptive evolution. If we view hybrid incompatibility as an adaptive valley, flanked by the adaptive peaks of the two pure species, the focus is on how the adaptive peak of the common ancestor split (Wu & Hollocher, 1998). What forces drive the two peaks apart? How many genetic events does it take to separate them far enough to create a fitness valley? How many are adaptive changes and how many are random events?

A strong indication that the evolution of RI may be a by product of either natural or sexual selection is the relative rate of evolution towards hybrid male sterility vs. inviability. Hybrid male sterility reflects the rapid differ-

entiation in the genetics of spermatogenesis and hybrid inviability reflects the differentiation in everything else. This may be the most interesting facet of Haldane's rule of RI (Wu *et al.*, 1996; Presgraves & Orr, 1998). If RI is decoupled from differential adaptation and is primarily driven by mutation, one would expect the evolution of hybrid inviability to outpace hybrid male sterility by 10:1, in terms of the number of loci involved [this ratio in Drosophila is known on the basis of the mutagenic potentials of inviability vs. male sterility (see Wu & Davis, 1993)]. However, the realized evolutionary patterns are overwhelmingly in the opposite direction with a ratio skewed further from 1:10, hence a discrepancy of two orders of magnitude (hybrid female sterility is unremarkable in its rate of evolution). This has been summarized in Table 1 which contrasts the number of loci contributing to different traits of speciation. Why do the genes for sperm-making evolve so rapidly? A plausible suggestion is sexual selection (Coulthart & Singh, 1988; Wu *et al.*, 1996; Presgraves & Orr, 1998; Ting *et al.*, 1998).

Outside of the *D. melanogaster* clade where detailed genetic analysis is not feasible, we may rely on the patterns of species hybridization. In general, the number of species pairs yielding hybrid male sterility is much larger than those yielding inviability or hybrid female sterility. For Drosophila, the observed numbers are 199:14:3 (for male sterility:male inviability:female sterility) and, for mammals, 25:0: 0 (Wu & Davis, 1993). Even in mosquitoes whose males remain homogametic and where the recessivity of genes would confound the rapid evolution of males' reproductive characters, a careful analysis has nonetheless revealed the rapid evolution of hybrid male sterility (Presgraves & Orr, 1998).

These observations suggest that the biology of males' reproductive systems diverges much faster than other systems (see also Andersson, 1994; Endler & Basolo, 1998). This is an extension of Eberhard (1985) treatment of the rapid evolution of animals' male genitalia. Sexual selection operates forcefully on males' reproductive functions, of which male genitalia is a conspicuous example. Divergence in spermatogenic programmes should fall within the same framework and the manifestation of such divergence is hybrids' inability to make sperm.

Recent molecular studies have strongly supported the view of sexual selection driving the rapid evolution of male reproductive genes. The most interesting example is the protamine genes of primates. Protamine is a main component of the DNA-protein complex in the nucleus of sperm and are functionally analogous to the histones in the somatic cells. Whereas histones are the slowest evolving proteins in the mammalian genome, protamine genes are the fastest evolving ones (Wyckoff *et al.*, 2000). There are many other examples from the studies of mammals (Sutton & Wilkinson, 1997), marine invertebrates (Lee & Vacquier, 1992; Metz & Palumbi, 1996) and Drosophila (Civetta & Singh, 1998; Nurminsky *et al.*,

1998; Ting *et al.*, 1998; Tsaur *et al.*, 1998). Fast evolution, of course, is not equivalent to positive selection. Detailed molecular evolutionary and population genetic analyses are usually necessary to distinguish between the relaxation of negative selection and the operation of positive selection (Fay & Wu, 2000; Wyckoff *et al.*, 2000).

In summary, the discussions above link the evolution of RI with the action of positive (sexual) selection, but only indirectly. Direct evidence would have to show that genes of RI themselves have been under the influence of directional selection. Lysin and bindin on the sperm of abalone and sea urchin, respectively, are two such examples. Both function in sperm–egg interactions and both have been shown to evolve rapidly because of positive selection (Lee *et al.*, 1995; Metz & Palumbi, 1996). Interestingly, the evolution of the receptor of lysin on the egg membrane does not show a comparable, elevated rate of amino acid substitution (Swanson & Vacquier, 1998). This observation thus supports the hypothesis of sexual selection over the classical lock-and-key interpretation (Eberhard, 1985). The other example is the Ods gene of speciation (Ting *et al.*, 1998), which has already been discussed in the context of species' genic history (Ting *et al.*, 2000). Ods is a major component of hybrid male sterility between *D. mauritiana* and *D. simulans* (Perez & Wu, 1995) and belongs to the homeodomain family of regulatory genes. Significantly, whereas the homeodomain is generally highly conserved, the homeodomain of Ods has evolved faster than even the neighbouring neutral DNAs between *D. mauritiana* and *D. simulans*. The molecular signature indisputably suggests positive selection driving this rapid evolution. There are many outstanding questions concerning the evolution of RI and differential adaptation. Molecular characterization of speciation genes may offer the promise of studying speciation at a deeper conceptual and mechanistic level.

Conclusion

The process of speciation is gene-based but RI is fundamentally a genomic concept. Speciation defined by the criteria of RI, as does BSC, would be inconsistent with the process of speciation itself. During the years, BSC has also decoupled the analysis of RI from the study of differential adaptation, contrary to what was originally conceived (Mayr, 1963). Empirical evidence has increasingly suggested the need to revise BSC by redirecting the focus from RI of the whole genomes to differential adaptation at the genic level. In the revised (or updated) BSC, RI would be treated as divergence in the underlying reproductive (spermatogenic or oogenic), developmental (e.g. embryogenic) or behavioural (sexual) traits. For example, the divergence in the reproductive characters would lead to hybrid sterility. This divergence can be used to delineate species even when RI remains quite incomplete. The revision pre-serves the essential character of BSC that species are divergent and incompatible entities (but considers the incompatibility as part of divergence, and at the genic level). A more thorough treatment of the practice and implications of this revision of BSC may be a worthy subject in the near future.

References

Akashi, H. 1997. Codon bias evolution in Drosophila: population genetics of mutation-selection-drift. *Gene* **205**: 269–278.

Alipaz, J.A. 2000. *The nature of incipient speciation in Drosophila.* PhD Thesis, Chicago, IL.

Andersson, M. 1994. *Sexual Selection.* Princeton University Press, Princeton.

Andolfatto, P. 2001. Contrasting patterns of X-linked and autosomal nucleotide variation in *Drosophila melanogaster* and *D. simulans. Mol. Biol. Evol.* **18**: 279.

Arnold, M.L. & Emms, S.K. 1998. Paradigm lost: natural hybridization and evolutionary innovations. In: *Endless Forms: Species and Speciation* (D. Howard & S. Berlochers, eds), pp. 379–389. Oxford University Press, Oxford.

Avise, J.C. 1994. *Molecular Markers, Natural History, and Evolution.* Chapman & Hall, New York.

Ballard, J.W.O. 2000. Comparative genomics of mitochondrial DNA in members of the Drosophila melanogaster subgroup. *J. Mol. Evol.* **51**: 48–63.

Barbujani, G., Magagni, A., Minch, E. & Cavalli-Sforza, L.L. 1997. An apportionment of human DNA diversity. *Proc. Natl. Acad. Sci.* **94**: 4516–4519.

Barton, N.H. 1992. On the spread of new gene combinations in the third phase of Wright's shifting balance. *Evolution* **46**: 551–557.

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

Barton, N. & Bengtsson, B.O. 1986. The barrier to genetic exchange between hybridising populations. *Heredity* **56**: 357–376.

Berlocher, S.H. 1998. Can sympatric speciation via host or habitat shift be proven from phylogenetic and biogeographic evidence? In: *Endless Forms: Species and Speciation* (D. Howard & S. Berlochers, eds), pp. 99–113. Oxford University Press, Oxford.

Berlocher, S.H. 2000. Radiation and divergence in the *Rhagoletis pomonella* species group: inferences from allozymes. *Evolution* **54**: 543–557.

Besansky, N.J., Powell, J.R., Caccone, A., Hamm, D.M., Scott, J.A. & Collins, F.H. 1994. Molecular phylogeny of Anopheles gambiae complex suggests genetic introgression betwen principal malaria vectors. *Proc. Natl. Acad. Sci.* **91**: 6885–6888.

Bradshaw, H.D., Wilbert, S.M., Otto, K.G. & Schemske, D.W. 1995. Genetic mapping of floral traits associated with reproductive isolation in monkeyflowers (*Mimulus*). *Nature* **376**: 762–765.

Brakefield, P.M., Gates, J., Keys, D., Kesbeke, F., Wijingaarden, P.J., Monteiro, A., French, V. & Carroll, S.B. 1996. Development, plasticity and evolution of butterfly eyespot patterns. *Nature* **384**: 236–242.

Bush, G.L. 1975. Modes of animal speciation. *Ann. Rev. Ecol. Syst.* **6**: 339–364.

Bush, G.L. 1998. The conceptual radicalization of an evolutionary biologist. In: *Endless Forms: Species and Speciation* (D. Howard & S. Berlochers, eds), pp. 425–438. Oxford University Press, Oxford.

Butlin, R. 1998. What do hybrid zones in general, and the Chorthippus parallelus zone in particular, tell us about speciation? In: *Endless Forms: Species and Speciation* (D. Howard & S. Berlochers, eds), pp. 367–378. Oxford University Press, Oxford.

Cabot, E.L., Davis, A.W., Johnson, N.A. & Wu, C.-I. 1994. Genetics of reproductive isolation in the *Drosophila simulans* clade: Complex epistasis underlying hybrid male sterility. *Genetics* **137**: 175–189.

Caccone, A., Moriyama, E.N., Gleason, J.M., Nigro, L. & Powell, J.R. 1996. A molecular phylogeny for the *Drosophila melanogaster* subgroup and the problem of polymorphism data. *Mol. Biol. Evol.* **13**: 1224–1232.

Capy, P., Veuille, M., Paillette, M., Jallon, J., Vouidibio, J. & David, J.R. 2000. Sexual isolation of genetically differentiated sympatric populations of *Drosophila melanogaster* in Brazzaville, Congo: the first step towards speciation? *Heredity* **84**: 468–475.

Carson, H.L. & Templeton, A.R. 1984. Genetic revolutions in relation to speciation phenomena: the founding of new populations. *Ann. Rev. Ecol. Syst.* **15**: 97–131.

Charlesworth, B., Coyne, J.A. & Barton, N.H. 1987. The relative rates of evolution of sex chromosomes and autosomes. *Am. Natur.* **130**: 113–146.

Civetta, A. & Singh, R. 1998. Sex-related genes, directional sexual selection and speciation. *Mol. Biol. Evol.* **15**: 901–909.

Coulthart, M.B. & Singh, R.S. 1988. Low genic variation in male reproductive tract proteins of *Drosophila melanogaster* and *D. simulans*. *Mol. Biol. Evol.* **5**: 167–181.

Coyne, J.A. 1984. Genetic basis of male sterility in hybrids between two closely related species of Drosophila. *Proc. Natl. Acad. Sci.* **81**: 4444–4447.

Coyne, J.A. 1994. Ernst Mayr on the origin of species. *Evolution* **48**: 19–30.

Coyne, J.A. 1996. Genetics of sexual isolation in male hybrids of *Drosophila simulans* and *D. mauritiana*. *Genet. Res.* **68**: 211–220.

Coyne, J.A. & Charlesworth, B. 1989. Genetic analysis of X-linked sterility in hybrids between three sibling species of Drosophila. *Heredity* **62**: 97–106.

Coyne, J.A. & Orr, H.A. 1998. The evolutionary genetics of speciation. *Phil. Trans. Roy. Soc. Lond. B* **353**: 287–305.

Coyne, J.A., Simeonedes, S. & Rooney, P.J. 1998. Relative paucity of genes causing inviability in hybrids between *D. melanogaster* and *D. simulans*. *Genetics* **150**: 1091–1103.

Crow, J.F. & Kimura, M. 1970. *An Introduction to Population Genetics Theory*. Harper & Row, New York.

Darwin, C. 1859. *On the Origin of Species*. Murray, London.

Darwin, C. 1875. *The Descent of Man, and Selection in Relation to Sex*. Murray, London.

Davis, A.W. & Wu, C.-I. 1996. The broom of the sorcerer's apprentice: the fine structure of a chromosomal region causing reproductive isolation between two sibling species of Drosophila. *Genetics* **143**: 1287–1298.

Davis, A.W., Noonburg, E. & Wu, C.-I. 1994. Complex genic interactions between conspecific chromosomes underlying hybrid female sterility in the *Drosophila simulans* clade. *Genetics* **137**: 191–199.

Davis, A.W., Roote, J., Morley, T., Sawamura, K., Herrmann, S. & Ashburner, M. 1996. Rescue of hybrid sterility in crosses between *D. melanogaster* and *D. simulans*. *Nature* **380**: 157–159.

Dieckmann, U. & Doebeli, M. 1999. On the origin of species by sympatric speciation. *Nature* **400**: 354–357.

Dobzhansky, T. 1937. *Genetics and the Origin of Species*. Columbia University Press, New York.

Dobzhansky, T. 1970. *Genetics of the Evolutionary Process*. Columbia University Press, New York.

Doebley, J., Stec, A. & Hubbard, L. 1997. The evolution of apical dominance in maize. *Nature* **386**: 485–488.

Doi, M., Matsuda, M., Tomaru, M., Matsubayashi, H. & Oguma, Y. 2001. A locus for female mate discrimination causing sexual isolation in Drosophila. *Proc. Natl. Acad. Sci.* **98**: 6714–6719.

Eberhard, W.G. 1985. *Sexual Selection and Animal Genitalia*. Harvard University Press, Cambridge, MA.

Endler, J.A. 1977. *Geographical Variation, Speciation and Clines*. Princeton University Press, Princeton.

Endler, J.A. & Basolo, A.L. 1998. Sensory ecology, receiver biases and sexual selection. *Trends Ecol. Evol.* **13**: 415–420.

Fay, J.C. & Wu, C.-I. 2000. Hitchhiking under positive Darwinian selection. *Genetics* **155**: 1405–1413.

Feder, J.L., Roethele, J.B., Wlazlo, B. & Berlocher, S.H. 1997. Selective maintenance of allozyme differences among sympatric host races of the apple maggot fly. *Proc. Natl. Acad. Sci.* **94**: 11417–11421.

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

Freeland, J.R. & Boag, P.T. 1999. The mitochondrial and nuclear genetic homogeneity of the phenotypically diverse Darwin's ground finches. *Evolution* **53**: 1553–1563.

Gavrilets, S. 1997. Evolution and speciation on holey adaptive landscapes. *Trends Ecol. Evol.* **12**: 307–312.

Gavrilets, S. 1999. A dynamical theory of speciation on holey adaptive landscapes. *Am. Natur.* **154**: 1–22.

Gavrilets, S. 2000. Waiting time to and duration of parapatric speciation. *Roc. Roy. Soc. London (Series B)* **256**: 2483–2492.

Gavrilets, S., Li, H. & Vose, M. 2000. Patterns of parapatric speciation. *Evolution* **54**: 1126–1134.

Goldschmidt, R. 1940. *The Material Basis of Evolution*. Yale University Press, New Haven.

Grant, V. 1971. *Plant Speciation*. Columbia University Press, New York.

Grant, P.R. & Grant, B.R. 1996. Speciation and hybridization of island birds. *Philos. Trans. R. Soc. B* **351**: 765–772.

Grant, B.R. & Grant, P.R. 1998. Hybridization and speciation in Darwin's finches: the role of sexual imprinting on a culturally transmitted trait. In: *Endless Forms: Species and Speciation* (D. Howard & S. Berlochers, eds), pp. 404–424. Oxford University Press, Oxford.

Hammer, M.F., Schimenti, J. & Silver, L.M. 1989. Evolution of mouse chromosome 17 and the origin of inversions associated with t-haplotypes. *Proc. Natl. Acad. Sci.* **86**: 3261–3265.

Harrison, R.G. 1990. Hybrid zones: windows on evolutionary process. *Oxford Surv. Evol. Biol.* **7**: 69–128.

Harrison, R.G. (ed.) 1993. *Hybrid Zones and the Evolutionary Process*. Oxford University Press, Oxford.

Harrison, R.G. 1998. Linking evolutionary pattern and process: The relevance of species concepts for the study of speciation. In: *Endless Forms: Species and Speciation* (D. Howard & S. Berlochers, eds), pp. 19–31. Oxford University Press, Oxford.

Hilton, H., Kliman, R.M. & Hey, J. 1994. Using hitchhiking genes to study adaptation and divergence during speciation within the *Drosophila melanogaster* species complex. *Evolution* **48**: 1900–1913.

Hollocher, H. & Wu, C.-I. 1996. The genetics of reproductive isolation in the *Drosophila simulans* clade: X vs. autosomal effects and male vs. female effects. *Genetics* **143**: 1243–1255.

Hollocher, H., Ting, C.T., Pollack, F. & Wu, C.-I. 1997a. Incipient speciation by sexual isolation in *Drosophila melanogaster*: Variation in mating preference and correlation between sexes. *Evolution* **51**: 1175–1181.

Hollocher, H., Ting, C.T., Wu, M.L. & Wu, C.-I. 1997b. Incipient speciation by sexual isolation in *Drosophila melanogaster*: The genetics of the Zimbabwe race. *Genetics* **147**: 1191–1201.

Hurst, L.D. & Pomiankowski, A. 1991. Causes of sex ratio bias may account for unisexuality in hybrids: a new explanation of Haldane's rule and related phenomena. *Genetics* **128**: 841–858.

Hurst, G.D.D. & Schilthuizen, M. 1998. Selfish genetic elements and speciation. *Heredity* **80**: 2–8.

Jiggins, C.D. & Mallet, J. 2000. Bimodal hybrid zones and speciation. *Trends Ecol. Evol.* **15**: 250–255.

Jones, C.D. 1998. The genetic basis of Drosophila sechellia's resistance to a host plant toxin. *Genetics* **149**: 1899–1908.

Kliman, R.M., Andolfatto, P., Coyne, J.A., Depaulis, F., Kreitman, M., Berry, A.J., McCarter, J., Wakeley, J. & Hey, J. 2000. The population genetics of the origin and divergence of the *Drosophila simulans* complex species. *Genetics* **156**: 1913–1931.

Kondrashov, A.S. & Kondrashov, F.A. 1999. Interactions among quantitative traits in the course of sympatric speciation. *Nature* **400**: 351–354.

Kondrashov, A.S., Yampolsky, L.Y. & Shabalina, S.A. 1998. On the sympatric origin of species by means of natural selection. In: *Endless Forms: Species and Speciation* (D. Howard & S. Berlochers, eds), pp. 90–98. Oxford University Press, Oxford.

Korol, A., Rashkovestsky, E., Iliadi, K., Michalak, P., Ronin, Y. & Nevo, E. 2000. Nonrandom mating in *Drosophila melanogaster* laboratory populations derived from closely adjacent ecologically contrasting slopes at 'Evolution Canyon'. *Proc. Natl. Acad. Sci.* **97**: 12637–12642.

Lachaise, D.M., Cariou, M.-L., David, J.R., Lemeunier, F. & Tsacas, L. 1988. The origin and dispersal of the *Drosophila melanogaster* subgroup: a speculative paleogeographic essay. *Evol. Biol.* **22**: 159–225.

Lee, Y.-H. & Vacquier, V.D. 1992. The divergence of species-specific abalone sperm lysins is promoted by positive Darwinian selection. *Biol. Bull.* **182**: 97–104.

Lee, Y.-H., Ota, T. & Vacquier, V.D. 1995. Positive selection is a general phenomenon in the evolution of abalone sperm lysin. *Mol. Biol. Evol.* **12**: 231–238.

Lewontin, R.C. & Krakauer, J.K. 1973. Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. *Genetics* **74**: 175–195.

Li, W.-H. 1987. Models of nearly neutral mutation with particular implications for nonrandom usage of synonymous codons. *J. Mol. Evol.* **24**: 337–345.

Macnair, M.R. & Gardner, M. 1998. The evolution of edaphic endemics. In: *Endless Forms: Species and Speciation* (D. Howard & S. Berlochers, eds), pp. 157–171. Oxford University Press, Oxford.

Mallet, J., McMillan, W.O. & Jiggins, C.D. 1998. Mimicry and warning color at the boundary between races and species. In: *Endless Forms: Species and Speciation* (D. Howard & S. Berlochers, eds), pp. 390–403. Oxford University Press, Oxford.

Maynard Smith, J. & Haigh, J. 1974. The hitch-hiking effect of a favorable gene. *Genet. Res.* **23**: 23–35.

Mayr, E. 1959. Where are we? *CSH Sympos. Quant. Biol.* **24**: 1–14.

Mayr, E. 1963. *Animal Species and Evolution*. The Belknap press, Cambridge, MA.

Metz, E.C. & Palumbi, S.R. 1996. Positive selection and sequence rearrangements generate extensive polymorphism in the game recognition protein bindin. *Mol. Biol. Evol.* **13**: 397–406.

Meyer, A. 1993. Phylogenetic relationships and evolutionary processes in East African cichlid fishes. *Trends Ecol. Evol.* **8**: 279–284.

Muller, H.J. 1940. Bearing of the Drosophila work on systematics. In: *The New Systematics* (J.S. Huxleys, ed.), pp. 185–268. Clarendon Press, Oxford.

Muller, H.J. & Pontecorvo, G. 1942. Recessive genes causing interspecific sterility and other disharmonies between *Drosophila melanogaster* and simulans. *Genetics* **27**: 157.

Nei, M. 1975. *Molecular Population Genetics and Evolution*. North-Holland Publishing Co., Amsterdam.

Nei, M. & Roychoudhury, A.K. 1993. Evolutionary relationships of human populations on a global scale. *Mol. Biol. Evol.* **10**: 927–943.

Nei, M., Maruyama, T. & Wu, C.-I. 1983. Models of evolution of reproductive isolation. *Genetics* **103**: 557–579.

Nurminsky, D.I., Nurminskaya, M.V., Aguiar, D.D. & Hartl, D.L. 1998. Selective sweep of a newly evolved sperm-specific gene in Drosophila. *Nature* **396**: 572–575.

Orr, H.A. 1989. Localization of genes causing postzygotic isolation in two hybridizations involving *Drosophila pseudoobscura*. *Heredity* **63**: 231–237.

Orr, H.A. 1992. Mapping and characterization of a 'speciation' gene in Drosophila. *Genet. Res.* **59**: 73–80.

Orr, H.A. 1995. The population genetics of speciation: the evolution of hybrid incompatibilities. *Genetics* **139**: 1805–1813.

Orr, H.A. 1997. Haldane's rule. *Ann. Rev. Ecol. Syst.* **28**: 195–218.

Palopoli, M. & Wu, C.-I. 1994. Genetics of hybrid male sterility between Drosophila sibling species: a complex web of epistasis is revealed in interspecific studies. *Genetics* **138**: 329–341.

Palopoli, M.F., Davis, A.W. & Wu, C.-I. 1996. Discord between the phylogenies inferred from molecular vs. functional data: uneven rates of functional evolution or low levels of gene flow? *Genetics* **144**: 1321–1328.

Perez, D.P. & Wu, C.-I. 1995. Further characterization of the hybrid sterility gene, Odysseus (Ods), in the Drosophila simulans clade: one gene is not enough. *Genetics* **140**: 201–206.

Prakash, S. 1972. Origin of reproductive isolation in the absence of apparent genetic differentiation in a geographic isolate of *Drosophila pseudoobscura*. *Genetics* **72**: 143–155.

Presgraves, D.C. & Orr, H.A. 1998. Haldane's rule in taxa lacking a hemizygous X. *Science* **282**: 952–954.

Raff, R.A. & Kaufman, T.C. 1983. *Embryos, Genes and Evolution*. Macmillan Publishing Co., New York.

Rice, W.R. & Hostert, E.E. 1993. Laboratory experiments on speciation: what have we learned in 40 years? *Evolution* **47**: 1637–1653.

Rieseberg, L.H. & Burke, J.M. 2001. The biological reality of species: gene flow, selection, and collective evolution. *Taxon* (in press).

Rieseberg, L.H., Fossen, C.V. & Desrochers, A.M. 1995. Hybrid speciation accompanied by genomic reorganization in wild sunflowers. *Nature* **375**: 313–316.

Rieseberg, L.H., Whitton, J. & Gardner, K. 1999. Hybrid zones and the genetic architecture of a barrier to gene flow between two wild sunflower species. *Genetics* **152**: 713–727.

Rundle, H.D., Nagel, L., Boughman, J.W. & Schluter, D. 2000. Natural selection and parallel speciation in sympatric sticklebacks. *Science* **287**: 306–308.

Sawamura, K., Davis, A.W. & Wu, C.-I. 2000. Genetic analysis by means of introgression into Drosophila melanogaster. *Proc. Natl. Acad. Sci.* **97**: 2652–2655.

Schaeffer, S.W. & Miller, E.L. 1991. Nucleotide sequence analysis of Adh genes estimates the time of geographic isolation of the Bogota population of *Drosophila pseudoobscura*. *Proc. Natl. Acad. Sci.* **88**: 6097–6101.

Schilthuizen, M. 2000. Dualism and conflicts in understanding speciation. *BioEssay* **22**: 1134–1141.

Schliewen, U.K., Tautz, D. & Paabo, S. 1994. Sympatric speciation suggested by monophyly of crater lake cichlids. *Nature* **368**: 629–632.

Schluter, D. 1998. Ecological causes of speciation. In: *Endless Forms: Species and Speciation* (D. Howard & S. Berlochers, eds), pp. 114–129. Oxford University Press, Oxford.

Shaw, D.D., Marchant, A.D., Arnold, M.L., Contreras, N. & Kohlmann, B. 1990. The control of gene flow across a narrow hybrid zone: a selective role for chromosomal rearrangement. *Genome* **68**: 1761–1769.

Slatkin, M. 1987. Gene flow and the geographic structure of natural populations. *Science* **236**: 787–792.

Solignac, M. & Monnerot, M. 1986. Race formation, speciation and introgression within *Drosophila simulans*, *D. mauritiana* and *D. sechellia* inferred from mitochondrial DNA analysis. *Evolution* **40**: 531–539.

Sutton, K.A. & Wilkinson, M.F. 1997. Rapid evolution of a homeodomain: evidence for positive selection. *J. Mol. Evol.* **45**: 579–588.

Swanson, W.J. & Vacquier, V.D. 1998. Concerted evolution in an egg receptor for a rapidly evolving abalone sperm protein. *Science* **281**: 710–712.

Szymura, J.M. & Barton, N.H. 1991. Genetic analysis of a hybrid zone between the fire-bellied toads, *Bombina bombina* and *B. variegata*, near Cracow in southern Poland. *Evolution* **40**: 1141–1159.

Takahashi, A., Tsaur, S.C., Coyne, J.A. & Wu, C.-I. 2001. The nucleotide changes governing pheromonal variation in Drosophila and their evolution. *Proc. Natl. Acad. Sci.* **98**: 3920–3925.

Ting, C.-T., Tsaur, S.-C., Wu, M.-L. & Wu, C.-I. 1998. A rapidly evolving homeobox at the site of a hybrid sterility gene. *Science* **282**: 1501–1504.

Ting, C.-T., Tsaur, S.-C. & Wu, C.-I. 2000. The phylogeny of closely related species as revealed by the genealogy of a speciation gene, Odysseus. *Proc. Natl. Acad. Sci.* **97**: 5313–5316.

Ting, C.T., Takahashi, A. & Wu, C.-I. 2001. Incipient speciation by sexual selection: concurrent evolution at multiple loci. *Proceedings Natl. Acad. Sci.* **98**: 6709–6713.

True, J.R., Weir, B.S. & Laurie, C.C. 1996. A genome-wide survey of hybrid incompatibility factors by the introgression of marked segments of *Drosophila mauritiana* chromosomes into *Drosophila simulans*. *Genetics* **142**: 819–837.

True, J.R., Liu, J., Stam, L.F., Zeng, Z.-B. & Laurie, C.C. 1997. Quantitative genetic analysis of divergence in male secondary sexual traits between *Drosophila simulans* and *D. mauritiana*. *Evolution* **51**: 816–832.

Tsaur, S.C., Ting, C.-T. & Wu, C.-I. 1998. Positive selection driving the evolution of a gene of male reproduction, Acp26Aa of Drosophila. II. Divergence vs. polymorphism. *Mol. Biol. Evol.* **15**: 1040–1046.

Turelli, M. & Orr, H.A. 1995. The dominance theory of Haldane's rule. *Genetics* **140**: 389–402.

Turelli, M. & Orr, H.A. 2000. Dominance, epistasis and the genetics of postzygotic isolation. *Genetics* **154**: 1663–1679.

Via, S., Bouck, A.C. & Skillman, S. 2000. Reproductive isolation between divergent races of pea aphids on two hosts. II. Selection against migrants and hybrids in the parental environments. *Evolution* **54**: 1626–1637.

Wang, R.L., Wakeley, J. & Hey, J. 1997. Gene flow and natural selection in the origin of *Drosophila pseudoobscura* and close relatives. *Genetics* **147**: 1091–1106.

Wang, R.-L., Stec, A., Hey, J., Lukens, L. & Doebley, J. 1999. The limits of selection during maize domestication. *Nature* **398**: 236–239.

Werren, J.H. 1998. Wolbachia and speciation. In: *Endless Forms: Species and Speciation* (D. Howard & S. Berlochers, eds), pp. 245–260. Oxford University Press, Oxford.

White, M.J.D. 1954. *Animal Cytology and Evolution*. Cambridge University Press, Cambridge.

Wu, C.-I. & Davis, A.W. 1993. Evolution of postmating reproductive isolation: the composite nature of Haldane's rule and its genetic bases. *Am. Natur.* **142**: 187–212.

Wu, C.-I. & Palopoli, M.F. 1994. Genetics of postmating reproductive isolation in animals. *Ann. Rev. Genet.* **28**: 283–308.

Wu, C.-I. & Hollocher, H. 1998. Subtle is nature: the genetics of species differentiation and speciation. In: *Endless Forms: Species and Speciation* (D. Howard & S. Berlochers, eds), pp. 339–351. Oxford University Press, Oxford.

Wu, C.-I., Hollocher, H., Begun, D.J., Aquadro, C.F., Xu, Y. & Wu, M.-L. 1995. Sexual isolation in *Drosophila melanogaster*: a possible case of incipient speciation. *Proc. Natl. Acad. Sci.* **92**: 2519–2523.

Wu, C.-I., Johnson, N.A. & Palopoli, M.F. 1996. Haldane's rule and its legacy: why are there so many sterile males? *Trends Ecol. Evol.* **11**: 281–284.

Wyckoff, G.J., Wang, W. & Wu, C.-I. 2000. Rapid evolution of male reproductive proteins in the descent of man. *Nature* **403**: 304–309.

Zeng, Z.-B., Liu, J., Stam, L.F., Kao, C.-H., Mercer, J.M. & Laurie, C.C. 2000. Genetic architecture of a morphological shape difference between two Drosophila species. *Genetics* **154**: 299–310.