ELSEVIER

What does *Drosophila* genetics tell us about speciation?

James Mallet

Galton Laboratory, Department of Biology, University College London, 4 Stephenson Way, London, UK, NW1 2HE

Studies of hybrid inviability, sterility and 'speciation genes' in *Drosophila* have given insight into the genetic changes that result in reproductive isolation. Here, I survey some extraordinary and important advances in *Drosophila* speciation research. However, 'reproductive isolation' is not the same as 'speciation', and this *Drosophila* work has resulted in a lopsided view of speciation. In particular, *Drosophila* are not always well-suited to investigating ecological and other selection-driven primary causes of speciation in nature. Recent advances have made use of far less tractable, but more charismatic organisms, such as flowering plants, vertebrates and larger insects. Work with these organisms has complemented *Drosophila* studies of hybrid unfitness to provide a more complete understanding of speciation.

Drosophila and the understanding of speciation

The study of speciation was brought into the modern age and onto a firmer, more genetic basis by drosophilists such as Sturtevant, Dobzhansky and others in T.H. Morgan's lab during the 1930s and 1940s. Dobzhansky's reproductive isolation species concept, later incorporated into Mayr's biological species concept, is based in large part on the discovery of pairs of sibling species of *Drosophila*: genetically close and morphologically almost identical, but showing strong reproductive isolation in the form of hybrid sterility, hybrid inviability and assortative mating. Of course, studies in many other less tractable organisms, such as flowering plants, vertebrates and larger insects, have also contributed to the understanding of speciation [1]. Here, I focus on studies of *Drosophila* hybrid inviability and sterility, and discuss their role in our views of speciation.

Reproductive isolation and speciation

Statements such as 'speciation…occurs by the evolution of any of several forms of reproductive isolation, including the intrinsic sterility and inviability of hybrids' [2] are tautologically true if species are defined as reproductively isolated populations. (If species are instead viewed as phylogenetically distinct taxa, or as distinguishable genomic clusters in sympatry, the statement becomes somewhat more logically defensible.) However, it is not necessarily true that genes for hybrid unfitness are causes

of speciation. Most hybrid unfitness probably arose long after speciation, by which time hybrid production in nature had already ceased. Understanding speciation is not simply a matter of studying reproductive isolation or enumerating 'speciation genes.' Instead, we must investigate the relative strengths of different modes of reproductive isolation, and their order of establishment [1].

'Reproductive isolation' is the product of all barriers to hybridization or gene flow between populations. The term

Glossary

Allopatric: two populations that are completely geographically isolated are said to be allopatric (in terms of gene flow, $m \approx 0$). This situation is not very different from distant populations in parapatric contact, and therefore leads to the same population genetic consequences with respect to speciation.

Assortative mating: preferential mating among similar forms when in mixed or sympatric contact with other forms.

dN/dS: the ratio of non-synonymous (dN) to synonymous (dS) changes in a DNA sequence, normally estimated along the branches of a phylogeny or genealogy. If protein evolution is neutral, $dN/dS \approx 1$. When synonymous changes are more common than amino acid changes ($dN/dS < 1$), conservative selection against protein change can be inferred; if non-synonymous changes are more common ($dN/dS > 1$), 'positive selection' is likely.

Haldane's Rule: the tendency for the heterogametic sex of hybrids to suffer greater inviability or sterility than the homogametic sex. The rule was discovered in the 1920s by J.B.S. Haldane, who showed that hybrid incompatibilities were usually greater in the heterogametic sex, regardless of whether it was male (as in *Drosophila* and mammals) or female (as in birds and butterflies).

Intrinsic hybrid incompatibility: hybrid inviability or sterility expressed independently of the environment.

Introgression: transfer of genes between distinct taxa (including separate species).

Parapatric: two geographic entities that abut at the boundaries to their range are said to be parapatric. Although the populations in overlap zones can exchange genes as though they were sympatric ($m \approx 0.5$), most populations are distant from this boundary, and therefore exchange few or no genes with the alternate form ($m \approx 0$).

Positive selection: where non-synonymous DNA base substitutions encoding amino acid differences (dN) have diverged more rapidly between species than synonymous base changes (dS). Synonymous changes are approximately neutral because they have no effect on amino acid sequence and are often more common than are non-synonymous changes owing to conservative selection on the functional protein.

Reinforcement: evolution of assortative mate choice as an adaptation to avoid the cost of producing unfit hybrids. The term derives from the idea that evolution of mate choice reinforces other isolating mechanisms.

Sibling species: species actually or nearly indistinguishable on the basis of morphology, but which nonetheless maintain distinct genotypic clusters in sympatry and between which strong reproductive barriers exist.

Speciation gene: this 'perhaps unfortunate term' [62] is commonly used by drosophilists to mean a gene that causes intrinsic hybrid incompatibility. The term is confusing because it strongly implies a role in speciation, which is often not the case: 'by "speciation gene", we merely mean any gene that reduces hybrid fitness…many [such] incompatibilities may accumulate *after* the attainment of complete isolation' [63].

Sympatric: two populations are sympatric if they overlap spatially so that the potential fraction of gene exchange is nearly $m \approx 0.5$ (Box 1). If there are strong behavioural or ecological barriers to mating, actual gene flow can be much lower.

Corresponding author: Mallet, J. (j.mallet@ucl.ac.uk).

is normally used vaguely, and is not easily defined in population genetic terms (Box 1). To avoid confusion, it is best to break down 'reproductive isolation' into: (i) natural selection against immigration, for example between ecological niches [3]; (ii) mate choice or assortative fertilization leading to a lack of gene flow between populations; and (iii) selection against zygotes that form when hybridization occurs. 'Reproductive isolation' is therefore a portmanteau term for a variety of types of disruptive selection and (lack of) gene flow. Instead of forming a unitary species-level trait, these components are separate results of evolution that occur readily within, as well as between, species. 'Reproductive isolation' is perhaps easiest to interpret (although difficult to estimate) as a dimensionless selection:gene flow ratio that determines equilibrium genetic differentiation between sympatric populations or species (Box 1).

However, some of us have argued that speciation is not simply a matter of the evolution of reproductive isolation. A better criterion for whether speciation has occurred is the existence of stable multilocus genetic differentiation in the face of potential gene flow [1,4,5]. Under this view, a species pair exhibits a bimodal genotypic distribution in a local area. Separate species exist when selection dominates gene flow, so that distinguishable genotypic or genomic clusters are stable in sympatry [6–8] or in hybrid zones between parapatric forms [9].

Complete reproductive isolation will, of course, enable rapid divergence into separate genotypic clusters, but reproductive isolation among many taxa that people call species is incomplete. Hybridization and introgression occur regularly [1,10,11]. Speciation theorists are also among those who use multilocus genetic criteria as evidence for speciation, for example the existence of linkage disequilibrium [12] or, equivalently, multiple locus genetic divergence [13] or the presence of distinct genotypic clusters [14,15]. The use of genotypic rather than isolation criteria for species does not deny that reproductive isolation triggers the origin of sexual species. Because reproductive isolation is not the same as speciation, we need to decide which components of reproductive isolation are important. We must infer the sequence of genetic changes, and their relative effects on the stability of divergent populations to gene flow in sympatry.

These considerations show why we cannot easily understand speciation by examining divergent allopatric populations. For example, *Drosophila pseudoobscura* produces infertile male hybrids when males from the USA are mated to females from Bogotá, Colombia [16]. The two populations are separated by > 2000 km. If these forms were to come together today, three outcomes are possible: (i) one of the species could go extinct; (ii) genes causing hybrid sterility could be selected against, and would be gradually selected out while the populations fused and other genes mixed into a single population; or (iii) the two populations could stabilize and remain separate in spite of geographic overlap, perhaps as a result of the evolution of strong mating isolation. On balance, (i) and (ii) seem most likely, but speciation would be 'successful' only for (iii), in which the hybrid sterility would contribute to speciation. Therefore, we cannot decide whether these populations have already speciated in an evolutionary sense, or whether this kind of hybrid

---

Box 1. 'Reproductive isolation': proof of equivalence to genetic differentiation in sympatry

'Reproductive isolation' is generally considered to be a combined effect of all barriers to gene flow between divergent populations that are in contact. It therefore includes 'prezygotic isolation' [all features that prevent fertilization and gene flow (m)] and 'postzygotic isolation' [disruptive natural selection (s) against genes that flow, including that against zygotes formed by cross-fertilization]. Prezygotic isolation can be expressed as the inverse of gene flow, or $1/m$, and 'reproductive isolation' as the product of the two, or s/m. Because s and m both have units of fractions of the population per unit time, the units cancel, and we are left with reproductive isolation that is equivalent to a dimensionless parameter measuring the relative strength of the two processes. Haldane [53] showed that equilibrium gene frequencies at a disruptively selected single locus depend on this fraction. This measure of reproductive isolation is therefore equivalent to the strength of genetic differentiation in overlapping populations.

If selection is weak, gene flow swamps disruptive selection and populations will not diverge. When selection is strong, and if s/m is greater than some critical value, gene frequencies will diverge approximately in proportion to s/m (Haldane's isolation theory [53]). This critical value of reproductive isolation will enable the persistence of separate populations with divergent gene frequencies (i.e. genotypic clusters). If these populations overlap spatially, they are viewed as incipient or actual separate species. Thus, populations identified by reproductive isolation will usually be equivalent to those identified as genotypic clusters. Arguably, 'speciation' would be clarified if we abandoned the term 'reproductive isolation' and concentrated on selection and gene flow as separate processes.

However, Haldane's isolation theory applies most easily to single genes. Weakly selected genes can have subcritical values of s/m and show little differentiation, even when the genome as a whole is considered 'reproductively isolated' and strongly divergent at other genes. Parts of the genome can be 'reproductively isolated,' whereas other parts are not [54]. Taxa can be viewed as distinct species provided they form distinct clusters at some genomic regions: even without gene flow, human DNA is 98.77% identical to chimpanzee DNA, yet we regard humans and chimps as different species because of the remaining 1.23%.

Alternatively, reproductive isolation in a whole-genome, 'barriers to gene flow' sense can be calculated as: 1–(1–prezygotic isolation)(1–postzygotic isolation) [1,3,17]. Reproductive isolation is therefore equivalent to $1 - m(1 - s)$. However, studies of reproductive isolation in this vein (and this can also include studies of s/m) can usually go only to the F_1 incompatibility level. F_2 and subsequent hybrid breakdown are often ignored because selection parameters on all the loci become disconnected, and fitnesses become extremely variable. (For an approximate estimate of multi-generation components of genome-wide selection and gene flow in *Heliconius*, see Ref. [55] and Box 2.)

Many studies of 'reproductive isolation' in *Drosophila* do not quantify the ecological and/or sexual selection (Box 2) that is potentially more important in speciation than is intrinsic hybrid incompatibility. Reproductive isolation defined as a 'barrier to gene flow' is therefore almost impossible to evaluate with respect to its effect on maintenance of genotypic clusters, because all the types of selection on all the separate genes, and on gene flow, must be integrated over many generations. For this reason, sympatric pairs are judged in practice to be separate species if they show multiple character divergence in separate clusters, rather than by means of a calculation of 'reproductive isolation'.

incompatibility is important in speciation. [It is true that (iii), although probably a rare outcome of many secondary contacts, might still contribute most to actual cases of speciation. However, the possibility is hard to evaluate.]

Examining causes of speciation in currently sympatric species is also tricky, especially if the species are old. In crosses between species pairs such as *Drosophila simulans* and *Drosophila melanogaster*, many genes interact to produce hybrid inviability or sterility. However, disruptive ecological adaptation, not accompanied by lab-demonstrable 'reproductive isolation,' might often be the trigger for speciation, followed by the evolution of strong assortative mate choice, via 'reinforcement' after initial divergence. Once species have separated, hybrid inviability and/or sterility can evolve by the slow fixation of negatively epistatic differences in each population, as though in allopatry, leading to the present situation of strongly reproductively isolated species.

Drosophilists recognize these difficulties [1]. To determine the causes of speciation, the study of recently diverged, sympatric populations provides major insights. Here, genes currently causing reproductive isolation are likely to be the same as those that originally caused the forms to diverge. Populations that remain distinct in spite of regular, ongoing hybridization are of particular interest. Unfortunately, such situations seem rare in the well-studied *Drosophila*. Another useful procedure is to study the time course of speciation by means of comparative analysis. Meanwhile, extraordinary recent advances in the molecular and genetic understanding of *Drosophila* speciation genes have begun to open up the genetic basis of hybrid inviability and sterility. This cutting-edge work is the main subject of this review.

Comparative methods

In 1989, Coyne and Orr collated a vast literature on lab pre-mating and post-mating isolation, together with details of genetic distance (a surrogate measure for time since divergence) between many *Drosophila* species pairs. The data, updated in 1997 [17], demonstrate several important patterns: (i) assortative mating, hybrid inviability and sterility all increase gradually, if noisily, with genetic distance; (ii) hybrid unfitness, particularly sterility, usually affects males first, an example of Haldane's Rule; (iii) geographically separate populations (geographic races as well as species pairs) develop assortative mating at about the same speed as they gain postzygotic hybrid incompatibility; and (iv) sympatric pairs of populations (they remain distinct, so these are all considered full species) develop assortative mating more rapidly than do allopatric populations, and more rapidly than they gain postzygotic hybrid incompatibility. Many closely related, sympatric species are strongly isolated overall, but show little or no hybrid unfitness in the lab. The speed of evolution of hybrid unfitness, meanwhile, differs little between sympatric and allopatric populations.

Discovery (iv) suggests that speciation is more rapid when populations overlap. Adaptive 'reinforcement' appears to have been the cause. In essence, this is sympatric speciation, although the authors interpret it differently [1,17]. Because hybrids between many recently diverged species are fertile and viable, much hybrid unfitness driving this reinforcement must have been ecological, and not readily detectable in the lab. Previous to this study, reinforcement had been in doubt, but the evidence from hundreds of species of *Drosophila* was a major factor that led to a revival of reinforcement as a viable hypothesis. [An alternative explanation cannot be excluded entirely: divergent populations that made contact with low levels of post-mating isolation would have required strong assortative mating, or they would have fused into a single population (Box 1). Thus, we might be seeing a survival bias towards highly assortatively mating populations with low genetic divergence only among the sympatric populations.] Lab experiments, especially in *Drosophila* [18], show that reinforcement is possible. Recently, further evidence has come from in-depth studies showing that *Drosophila* species pairs have stronger mating isolation in sympatry than in allopatry [19,20].

Nature of genes underlying hybrid inviability and sterility

The many lab studies of *Drosophila* reproductive isolation in the past 10–20 years [17] have also led to a good understanding of genes underlying hybrid incompatibility and Haldane's Rule (Table 1). These genes rarely cause incompatibility in the parent species, so hybrid incompatibility must usually be a result of interactions between two or more genes (negative epistasis, or 'Dobzhansky–Muller incompatibilities') in different species. Today, an important contender to explain Haldane's Rule is a version of Muller's original 'dominance theory,' in which recessive genes on the X chromosome of one species interact with dominant genes on the autosomes of the other to cause incompatibilities [21]. The F_1 male hybrid, which is hemizygous with a single maternally derived X chromosome, expresses these deleterious effects, whereas, in the F_1 female, recessive incompatibilities are masked by the paternal X [21]. As expected under dominance theory, Haldane's Rule applies also in organisms such as birds and Lepidoptera: the female is the hemizygous sex and also the sex most likely to suffer hybrid incompatibilities [22,23]. In *Drosophila*, but not birds and Lepidoptera, another important process ('faster male') also goes in the direction of Haldane's Rule: hybrid male sterility evolves much faster than does hybrid inviability or female sterility, perhaps as a result of sexual selection via sperm competition [24]. Another possible explanation of 'faster male' evolution is genomic conflict over the sex-ratio during gametogenesis in the heterogametic sex [25].

If dominance theory is correct, Haldane's Rule is obeyed because each species fixes recessive incompatibilities faster than dominant incompatibilities. If dominant incompatibilities were fixed first, both sexes would suffer equally, because hemizygous and heterozygous X-chromosomal sexes would each express incompatibilities. It is still somewhat mysterious why recessive incompatibilities should be fixed faster. Turelli and Orr [21] originally proposed that, by analogy with single gene deleterious mutations, excess recessive incompatibilities are caused by losses of function, which are generally recessive in

Table 1. Examples of 'speciation genes' from the *Drosophila melanogaster* species group

Gene[a]	Species cross	Chromosomal location (species)	Epistasis Interacts with:	Deleterious effect on hybrids	Normal function	Positive selection?	Refs
OdsH ('Odysseus'; i.e. *Ods-site homeobox*)[b] FBgn0026058 Dmel_CG6352	simulans × mauritiana	X-16D (*mauritiana*)	Autosomal genes (presumably)	When introgressed to *simulans*, dominant male sterility results. Misexpression in apex of testis of sterile male hybrids	Presumed transcription factor containing homeobox domain. Expression normally embryonic; in *melanogaster*, null mutants mildly reduce fertility of very young males	Very rapid evolution (100–1000× as fast as other homeobox genes, 8× higher nucleotide substitution in homeodomain than in intron; $dN/dS = 10.0$ in branch to *mauritiana*, and 3.0 in branch to *simulans*)	[64,65]
Hmr (*Hybrid male rescue*) FBgn0001206 Dmel_CG1619	simulans × melanogaster	X-9D3 (*melanogaster*)	Unidentified genes on *melanogaster* X and autosomal regions, *Dsim\Lhr* (*Lethal hybrid rescue* FBgn0010058, *D. simulans* 2-88.4) and *In(1)AB* (FBab0003876)	*melanogaster* Hmr^+ in hybrids with *simulans* or *mauritiana* causes male and high-temperature female lethality; loss-of-function mutations or transgenes from *simulans* or *mauritiana* rescue hybrid viability of male progeny of *simulans* male × *melanogaster* female crosses	ADF1- or MYB-like transcription factor containing MADF domain. Unknown normal function	Very strong, many regions with $dN/dS > 1$, ~12 in exon 5 at 3' end. Selection is relatively ancient, perhaps during or soon after the origin of the two species	[35,66,67]
Nup98 (*nucleoporin 96–98*) FBgn0039120 Dmel_CG10198	simulans × melanogaster	3R-95B1-B5 (*simulans*)	Unknown recessive gene(s) on X chromosome; autosomal genes including *Dsim\Lhr* (FBgn0010058, *simulans* 2-88.4)	Loss-of-function mutations in *melanogaster* cause hybrid lethality in *melanogaster* female × *simulans* male crosses; the *simulans* allele is therefore a recessive hybrid lethal	Nuclear pore complex protein. Part of dicistronic gene *Nup98–Nup96*, which is alternatively transcribed	Strong in early part just after cleavage point of *Nup96* (normally highly conserved in eukaryotes)	[68]
zhr (*Zygotic hybrid rescue*) FBgn0004840[c]	simulans × melanogaster	X-62.5+/-4.0 (*melanogaster*)	*Dsim\mhr* in *D. simulans* (maternal hybrid rescue, FBgn0012874)	Rescues embryonic lethality of female progeny in *simulans* × *melanogaster* crosses, or in *sechellia* or *mauritiana* × *melanogaster* crosses. Female lethality suppressed by zhr^- deletions.	Gene is non-coding, presumed to bind chromatin proteins needed for chromatin function	Unknown	[34,69]
						(errata in red)	

[a]Accession numbers are given for Flybase (FB) and Genbank.
[b]Especially the 3 kb including exon 314 only.
[c]Sequence unknown.

interacting sets of genes as well as in single genes. Regardless of whether this is correct, experimental introgressions from *Drosophila mauritiana* and *D. sechellia* into *D. simulans* have shown that incompatibility loci (most of which are recessive and interact with dominant autosomal genes in the other species) are common on autosomes as well as on sex chromosomes [26,27], as expected under dominance theory.

In the F_1 generation, incompatibilities are expressed only if X-linked and in hemizygous hybrids, leading to Haldane's Rule. Another major finding is that introgressions producing hybrid male sterility are more common than those producing inviability or female sterility, confirming that 'faster male' evolution is important in *Drosophila* [28].

More recently, Presgraves [2] reasoned that if recessive–dominant interactions were more common than dominant–dominant interactions (as they must be if dominance theory explains Haldane's Rule), then recessive–recessive interactions might be more common still. Using the genomic technology available for *Drosophila*, Presgraves tested this by crossing female *D. melanogaster* with small autosomal deletions to males of *D. simulans*. Normally, this cross is lethal to hybrid males, but here males were rescued using the *Lhr* mutation (*Lethal hybrid rescue*, see 'epistasis' column in Table 1) from *D. simulans*, which suppresses the lethal hybrid effect of the hemizygous *melanogaster* X. Thus, Presgraves tested for the presence of autosomal recessive *simulans* alleles, unmasked by the deletions, which interacted with recessive genes on the hemizygous *melanogaster* X. These new interactions then killed rescued males to cause male deficits. Presgraves convincingly showed that recessive–recessive lethal interactions were approximately eightfold more common than the recessive–dominant incompatibilities normally involved in Haldane's Rule [2].

The numbers of 'speciation genes' affecting reproductive isolation between species

Experimental introgressions show that many factors contribute to hybrid incompatibility in the *melanogaster–simulans* group [2,25–28]. Fine-scale mapping of a small region of the *mauritiana* X chromosome [29] revealed that only 3% of the genome contained at least six separate factors, each able to cause complete hybrid sterility in crosses with *simulans*. From these and other data, at least 120 distinct recessive genes or regions were inferred to be scattered around the genome, all capable of causing complete sterility in hybrid males [24]. Approximately 190 gene regions can cause male hybrid lethality in *simulans–melanogaster* crosses [2]. We have the surprising result that many divergent genes with actions similar to those in Table 1 can kill or sterilize hybrids hundreds of times over. Not all these gene regions could have been responsible for the original speciation event; in fact, these crosses demonstrate that most divergence at incompatibility genes must have accumulated within each lineage after speciation.

Drosophila melanogaster and *D. simulans* provide rich information about hybrid unfitness that might be involved in speciation, but they have probably diverged too far for useful inference of initial causes. Another pair, *D. pseudoobscura* and *D. persimilis*, split $\sim$ 550 000 years ago, comparable to the ages of the three largely allopatric sibling species *D. simulans*, *D. mauritiana* and *D. sechellia*, whereas the Bogotá and USA races of *D. pseudoobscura* are younger, 100 000–200 000 years old [11,30]. *Drosophila pseudoobscura* and *D. persimilis* are partially sympatric. They differ in genes that cause F_1 hybrid male sterility, backcross hybrid inviability, species-specific female choice of males and F_1 hybrid male courtship dysfunction. Females from areas of overlap prefer to mate with their own species more strongly than do non-sympatric females. Reinforcement is therefore implicated as a cause of strengthened mate discrimination [19]. These differences map chiefly to inverted regions on the X chromosome and central parts of the second and third chromosomes. *Drosophila pseudoobscura–persimilis* hybrids are occasionally found in the wild, and there is evidence that gene flow homogenizes interspecific variation in mtDNA and several nuclear genes [11].

In the *D. pseudoobscura–persimilis* pair, an initial ability for divergent genes involved in speciation to persist in sympatry was probably enhanced by recombination suppression due to chromosomal inversions [31]. The only major exception is the finding of a dominant factor affecting mate choice, located within an uninverted genomic region. This allele causes assortative mating in both species and might have been recruited after stable sympatric divergence was established [32,33]. In addition, as few as 15 interacting loci on the Bogotá X chromosome contribute to male F_1 sterility in crosses between male USA and allopatric females from Bogotá within *D. pseudoobscura* [16]; the reverse cross produces fertile males. The results suggest that relatively few 'speciation genes' separate the *D. pseudoobscura* group taxa, and that chromosomal inversions might, by preventing recombination, enable differences causing speciation to build up in spite of sympatry [31].

Hybrids are apparently killed many times over in *D. simulans–melanogaster* and related crosses by multiple genes. Surprisingly, several single gene mutations or chromosomal abnormalities have been found that rescue hybrid viability or fertility. Normally, when females of *D. melanogaster* are crossed with males of *D. simulans*, female offspring are typically viable but sterile, whereas hybrid males usually do not survive to pupal stage. In the reciprocal cross, poorly viable daughters and sterile sons result. It is therefore extraordinary that three mutations can rescue hybrid sons of *melanogaster* mothers [*Lhr* in *D. simulans*; *Hybrid male rescue*, *Hmr*, and the inversion *In(1)AB* in *D. melanogaster*], and two mutations can rescue the viability of females in the reciprocal cross (*maternal hybrid rescue*, *mhr*, in *D. simulans*; and *Zygotic hybrid rescue*, *Zhr*, in *D. melanogaster*) [34,35]. The fact that single mutations can rescue viability and fertility redundantly effected by dozens of genes suggests that speciation genes have been overcounted, and that the many incompatibility genes belong to relatively few pathways [16,34]. Meanwhile, 'rescue' genes are themselves good candidates to be 'speciation genes' because their non-rescuing alleles cause unfitness.

Advances in *Drosophila* genomics have now enabled a number of 'speciation genes,' including some hybrid rescue genes, to be characterized (Table 1). Of the four loci highlighted, three are known to have diverged under 'positive' selection. This is a major discovery because if these or similar genes caused the initial burst of divergence leading to speciation, it would implicate natural selection as a major cause of speciation.

What we do not know

We have learnt a great deal about the genetics of reproductive isolation during the past decade by opening up the 'black box' of *Drosophila* genes causing hybrid unfitness. This work is leading towards a better understanding of the role of negative epistasis in evolution. 'Speciation genes' are good models for the kinds of locus that might trigger speciation, especially as reinforcement (which builds on hybrid unfitness) now seems to be widespread. Unfortunately, we do not yet know why positive selection occurs at these loci, or whether they had a role in initial divergence. Furthermore, they probably evolved after speciation was essentially complete. These genes for reproductive isolation are unlikely to have been involved in speciation.

In species other than *Drosophila*, hybrid unfitness is less tractable, and we know little about the genetics of hybrid inviability and sterility. Speciation genes are known in only a few species pairs, and only a few comparative analyses exist [22,23,36,37]. Yet in these other species, some incompatibilities can be attributed to factors other than the negatively epistatic problems on which drosophilists have focused. Much recent interest has been on adaptation to different environments, symbionts, or foods in organisms as diverse as flowering plants, true fruit flies, butterflies and fishes (e.g. Refs [37–45]; Box 2). Direct ecological selection against movement between niches should, in theory, be a more effective cause of genetic divergence (or 'reproductive isolation' in the sense of Box 1) than selection against hybrids in classic postzygotic isolation, because selection is against the pure foreign type, rather than against only a partially foreign, and partially native, hybrid genome [3].

Haldane's Rule and other examples of hybrid unfitness are probably explained mainly by environment-independent negatively epistatic interactions ('Dobzhansky–Muller incompatibilities') between autosomes and the X chromosome, but it is unlikely that this traditional model is generalizable as the major cause of hybrid incompatibility or speciation. Ten years ago, drosophilists did not believe that meiotic drive by selfish genetic elements was important in speciation; today, genomic conflict is again implicated in reproductive isolation [16]. In *Saccharomyces* yeasts, hybrid sterility is due to heterozygous disadvantage of chromosomal rearrangements, coupled with a directly deleterious effect of DNA divergence on chromosomal pairing [46–48], rather than necessarily involving negative epistasis. Reproductive isolation might also result naturally from pleiotropic, ecological effects on assortative mating and hybrid fitness. Classic environment-independent hybrid inviability or sterility is often weak by comparison (Box 2). Speciation is a complex, multilocus process, and we might well be suspicious of ideas that concentrate on intrinsic postzygotic reproductive isolation.

The wealth of comparative *Drosophila* data has reawakened interest in reinforcement. Yet recent experimental results suggesting reinforcement have largely, with some exceptions [19,20], come from other groups. Ecological parameters, such as host plants, might be as important in *Drosophila* [49,50] as in other insects [7,8], but there are few studies. Similarly, some of the youngest, most assortatively mating sympatric species show little intrinsic hybrid unfitness within as well as outside *Drosophila* [9,17,23,43,44] (Box 2). In most such cases, speciation probably had little to do with sterility or inviability measurable in the lab. It is embarrassing that, although many isofemale lines have been trapped in the wild since the pioneering work of Dobzhansky, the natural foods and larval habitats of *D. pseudoobscura* and *D. persimilis* are virtually unknown (M. Noor, personal communication). Although we have many lab studies on hybrid inviability and reinforcement in this group, we have no idea how the two species can coexist in nature.

Positive selection explains the rapid evolution of *Drosophila* genes expressed in testes and male accessory

---

Box 2. Ecological speciation: the example of *Heliconius*

Many organisms have now been used to infer the ecological basis of speciation (e.g. Refs [37–45]), although genetic studies are still somewhat primitive compared with those in *Drosophila*. Studies of butterflies in the *Heliconius melpomene* species group show how ecological studies, coupled with knowledge of genetics, can evaluate the relative importance of different kinds of reproductive isolation. *Heliconius melpomene* and close relatives, particularly *Heliconius cydno*, are readily distinguishable via differences in colour patterns. These patterns are adaptations for Müllerian mimicry with other distasteful *Heliconius* species [42], or with ithomiine butterflies [56]. Sympatric pairs of *melpomene* group species occasionally hybridize in the wild [10], resulting in some fertile hybrids, usually the males, whereas females are often sterile, in concordance with Haldane's Rule [55]. Thus, there is potential for gene flow, and indeed introgression has been demonstrated recently via genealogical analysis of sympatric *H. melpomene* and *H. cydno* [57].

Geographic races of *Heliconius* are also strongly differentiated in mimicry, even within each species, and these differences are determined by the same major loci that control interspecific wing colour differences [58,59]. Races meet in narrow hybrid zones maintained by strong selection for mimicry [56]. Colour patterns also act as mating cues, and colour pattern divergence is therefore associated with coevolved mate preferences, leading to pleiotropic assortative mating. This forms the major barrier to gene flow between *H. cydno* and *H. melpomene*, and is enhanced by probable reinforcement in sympatry [42,60]. Selection against hybrids due to mimicry has not been measured in this pair, but it must be strong judging by results from inter-racial hybrid zones ($s \approx 0.64$) of the same order as selection due to female sterility ($s \approx 0.70$ overall). The main barrier, however, is sexual. Assortative mating, sexual selection against hybrids and limited habitat overlap reduce gene flow to $m < 0.001$ [55].

Therefore, the value of overall reproductive isolation s/m is $\sim [1 - (1 - 0.64)(1 - 0.70)]/(0.001) \approx 892$ [$s/m \ll 1$ implies reproductive continuity, whereas $s/m \gg 1$ indicates strong isolation (Box 1), as here]. The ecologically generated effects, including colour pattern and assortative mating, account for approximately a thousand times more reproductive isolation than does hybrid sterility. Therefore, the handful of genes determining colour pattern differences, together with unknown other factors influencing assortative mating, such as pheromonal divergence, are 'speciation genes' in the true sense of being among the prime movers of speciation. The example is paralleled in recent comparative studies of species for which ecological differentiation, hybrid inviability and sterility, and genetic distance were quantified [3,61].

gland proteins [51], and also of genes known to affect hybrid sterility, and it seems likely that this can partly explain 'faster male' evolution. Strong 'conspecific sperm precedence,' which, in essence, is a form of post-mating assortative fertilization [52], might be due to rapid divergence of reproductive proteins in *Drosophila* and other organisms. These are probably examples of sexual selection driving hybrid unfitness [24]. But it is hard to pin down specific cases and whether this divergence is a cause or merely an incidental result of speciation. By contrast, sexual selection for outrageous male visual, pheromonal, sonic or mechanical traits (including a few notable cases in *Drosophila*) is well established. Female choice explains these traits, so they are also directly implicated in assortative mating that will ease speciation.

Conclusions

Drosophila has enormous potential for studying speciation, yet some of its advantages (e.g. ease of rearing in half-pint milk-bottles, availability of inbred stocks and genomic data) might be among the reasons why the ecological basis of speciation is less well developed than in more intractable taxa. Fly genome projects provide essential resources in studying genetic mechanisms in *Drosophila* itself, as well as candidate speciation genes for other species. *Drosophila* 'speciation genes' themselves are useful as general models of negative epistasis, even though most are not causes of speciation in their own right.

Drosophila geneticists were among the first to plan a modern attack on the problem of the origin of species. Unsurprisingly, *Drosophila* has since contributed strongly to our understanding of speciation. It is perhaps inevitable that flies are not the most convenient organisms for all types of speciation research. Drosophilists have, in my view, overemphasized the importance of hybrid sterility and inviability expressed in the lab. This is not to say that more field and ecological research cannot be done; perhaps more attention should be paid to *Drosophila* in the wild. Conversely, if nondrosophilists want to challenge *Drosophila* as a system for understanding speciation, they must prepare for decades of development of genetic strains and genomic resources. In truth, useful contributions will continue to be made by studies using *Drosophila* as well as a diversity of other organisms.

References
1　Coyne, J.A. and Orr, H.A. (2004) *Speciation*, Sinauer Associates
2　Presgraves, D.C. (2003) A fine-scale genetic analysis of hybrid incompatibilities in *Drosophila*. *Genetics* 163, 955–972
3　Nosil, P. *et al*. (2005) Reproductive isolation caused by natural selection against immigrants from divergent habitats. *Evolution* 59, 705–719
4　Mallet, J. (1995) A species definition for the Modern Synthesis. *Trends Ecol. Evol.* 10, 294–299
5　Feder, J.L. (1998) The apple maggot fly, *Rhagoletis pomonella*: flies in the face of conventional wisdom. In *Endless Forms. Species and Speciation* (Howard, D.J. and Berlocher, S.H., eds), pp. 130–144, Oxford University Press
6　Kondrashov, A.S. and Mina, M.V. (1986) Sympatric speciation: when is it possible? *Biol. J. Linn. Soc.* 27, 201–223
7　Berlocher, S.H. and Feder, J.L. (2002) Sympatric speciation in phytophagous insects: moving beyond controversy? *Annu. Rev. Entomol.* 47, 773–815
8　Drès, M. and Mallet, J. (2002) Host races in plant-feeding insects and their importance in sympatric speciation. *Phil. Trans. R. Soc. B* 357, 471–492
9　Jiggins, C.D. and Mallet, J. (2000) Bimodal hybrid zones and the scale of a snail (reply to M. Schilthuizen). *Trends Ecol. Evol.* 15, 469
10　Mallet, J. (2005) Hybridization as an invasion of the genome. *Trends Ecol. Evol.* 20, 229–237
11　Wang, R.L. *et al*. (1997) Gene flow and natural selection in the origin of *Drosophila pseudoobscura* and close relatives. *Genetics* 147, 1091–1106
12　Felsenstein, J. (1981) Skepticism towards Santa Rosalia, or why are there so few kinds of animals? *Evolution* 35, 124–138
13　Gavrilets, S. *et al*. (2000) Dynamics of speciation and diversification in a metapopulation. *Evolution* 54, 1493–1501
14　Kondrashov, A.S. and Kondrashov, F.A. (1999) Interactions among quantitative traits in the course of sympatric speciation. *Nature* 400, 351–354
15　Dieckmann, U. and Doebeli, M. (1999) On the origin of species by sympatric speciation. *Nature* 400, 354–357
16　Orr, H.A. and Irving, S. (2001) Complex epistasis and the genetic basis of hybrid sterility in the *Drosophila pseudoobscura* Bogotá-USA hybridization. *Genetics* 158, 1089–1100
17　Coyne, J.A. and Orr, H.A. (1997) 'Patterns of speciation in *Drosophila*' revisited. *Evolution* 51, 295–303
18　Rice, W.R. and Hostert, E.E. (1994) Laboratory experiments on speciation: what have we learned in forty years? *Evolution* 47, 1637–1653
19　Noor, M.A. (1995) Speciation driven by natural selection in *Drosophila*. *Nature* 375, 674–675
20　Higgie, M. *et al*. (2000) Natural selection and the reinforcement of mate recognition. *Science* 290, 519–521
21　Turelli, M. and Orr, H.A. (1995) The dominance theory of Haldane's rule. *Genetics* 140, 389–402
22　Presgraves, D.C. (2002) Patterns of postzygotic isolation in Lepidoptera. *Evolution* 56, 1168–1183
23　Price, T.D. and Bouvier, M.M. (2002) The evolution of F1 postzygotic incompatibilities in birds. *Evolution* 56, 2083–2089
24　Wu, C-I. *et al*. (1996) Haldane's rule and its legacy: why are there so many sterile males? *Trends Ecol. Evol.* 11, 281–284
25　Tao, Y. and Hartl, D.L. (2003) Genetic dissection of hybrid incompatibilities between *Drosophila simulans* and *D. mauritiana*. III. Heterogeneous accumulation of hybrid incompatibilities, degree of dominance, and implications for Haldane's rule. *Evolution* 57, 2580–2598
26　Hollocher, H. and 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
27　True, J.R. *et al*. (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
28　Tao, Y. *et al*. (2003) Genetic dissection of hybrid incompatibilities between *Drosophila simulans* and *D. mauritiana*. I. Differential accumulation of hybrid male sterility effects on the X and autosomes. *Genetics* 164, 1383–1398
29　Davis, A.W. and 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
30　Kliman, R.M. *et al*. (2000) The population genetics of the origin and divergence of the *Drosophila simulans* complex species. *Genetics* 156, 1913–1931
31　Noor, M.A.F. *et al*. (2001) Chromosomal inversions and the reproductive isolation of species. *Proc. Natl. Acad. Sci. U. S. A* 98, 12084–12088

32 Ortíz-Barrientos, D. *et al*. (2004) The genetics of speciation by reinforcement. *PLoS Biol.* 2, e416

33 Ortíz-Barrientos, D. and Noor, M.A.F. (2005) Evidence for a one-allele assortative mating locus. *Science* 310, 1467

34 Sawamura, K. *et al*. (1993) Hybrid lethal systems in the *Drosophila melanogaster species complex*. *Genetica* 88, 175–185

35 Barbash, D.A. *et al*. (2000) The *Drosophila melanogaster* hybrid male rescue gene causes inviability in male and female species hybrids. *Genetics* 154, 1747–1771

36 Wittbrodt, J. *et al*. (1989) Novel putative receptor tyrosine kinase encoded by the melanoma-inducing *Tu* locus in *Xiphophorus*. *Nature* 341, 415–421

37 Bradshaw, H.D. and Schemske, D.W. (2003) Allele substitution at a flower locus produces a pollinator shift in monkeyflowers. *Nature* 426, 176–178

38 Savolainen, V. *et al*. Sympatric speciation of palms on an oceanic island. *Nature* (in press)

39 Grant, P.R. and Grant, B.R. (1996) Speciation and hybridization in island birds. *Phil. Trans. R. Soc. B* 351, 765–772

40 Feder, J.L. *et al*. (1997) Selective maintenance of allozyme differences among sympatric host races of the apple maggot fly. *Proc. Natl. Acad. Sci. U. S. A.* 94, 11417–11421

41 Feder, J.L. *et al*. (1998) Sympatric host-race formation and speciation in *Rhagoletis* (Diptera: Tephritidae): a tale of two species for Charles D. In *Genetic Structure and Local Adaptation in Natural Insect Populations* (Mopper, S. and Strauss, S.Y., eds), pp. 408–441, Chapman & Hall

42 Jiggins, C.D. *et al*. (2001) Reproductive isolation caused by colour pattern mimicry. *Nature* 411, 302–305

43 Rundle, H.D. *et al*. (2000) Natural selection and parallel speciation in sympatric sticklebacks. *Science* 287, 306–308

44 Seehausen, O. (2003) Hybridization and adaptive radiation. *Trends Ecol. Evol.* 19, 198–207

45 Barluenga, M. *et al*. (2006) Sympatric speciation in Nicaraguan crater lake cichlid fish. *Nature* 439, 719–723

46 Hunter, N. *et al*. (1996) The mismatch repair system contributes to meiotic sterility in an interspecific yeast hybrid. *EMBO J.* 15, 1726–1733

47 Fischer, G. (2000) Chromosomal evolution in *Saccharomyces*. *Nature* 405, 451–454

48 Greig, D. *et al*. (2003) A role for the mismatch repair system during incipient speciation in *Saccharomyces*. *J. Evol. Biol.* 16, 429–437

49 Brazner, J.C. and Etges, W.J. (1993) Pre-mating isolation is determined by larval rearing substrates in cactophilic *Drosophila mojavensis*. II. Effects of larval substrates on time to copulation, mate choice and mating propensity. *Evol. Ecol.* 7, 605–624

50 Jones, C.D. (2005) The genetics of adaptation in *Drosophila sechellia*. *Genetica* 123, 137–145

51 Richards, S. *et al*. (2005) Comparative genome sequencing of *Drosophila pseudoobscura*: chromosomal, gene, and cis-element evolution. *Genome Res.* 15, 1–18

52 Price, C.S.C. (1997) Conspecific sperm precedence in *Drosophila*. *Nature* 388, 663–666

53 Haldane, J.B.S. (1932) *The Causes of Evolution*, Longmans, Green & Co.

54 Wu, C-I. (2001) The genic view of the process of speciation. *J. Evol. Biol.* 14, 851–865

55 Naisbit, R.E. *et al*. (2002) Hybrid sterility, Haldane's rule, and speciation *Heliconius cydno* and *H. melpomene*. *Genetics* 161, 1517–1526

56 Mallet, J. *et al*. (1998) Mimicry and warning color at the boundary between races and species. In *Endless Forms: Species and Speciation* (Howard, D.J. and Berlocher, S.H., eds), pp. 390–403, Oxford University Press

57 Bull, V. *et al*. (2006) Polyphyly and gene flow between non-sibling *Heliconius* species. *BMC Biol.* 4, 11

58 Naisbit, R.E. *et al*. (2003) Mimicry: developmental genes that contribute to speciation. *Evol. Devel.* 5, 269–280

59 Jiggins, C.D. *et al*. (2005) A genetic linkage map of the mimetic butterfly, *Heliconius melpomene*. *Genetics* 171, 557–570

60 Jiggins, C.D. *et al*. (2005) Assortative mating and speciation as pleiotropic effects of ecological adaptation: examples in moths and butterflies. In *Insect Evolutionary Ecology* (Fellowes, M. *et al*., eds), pp. 451–473, Royal Entomological Society

61 Funk, D.J. *et al*. (2006) Ecological divergence exhibits consistently positive associations with reproductive isolation across disparate taxa. *Proc. Natl. Acad. Sci. U. S. A.* 103, 3209–3213

62 Orr, H.A. (2005) The genetic basis of reproductive isolation: insights from *Drosophila*. *Proc. Natl. Acad. Sci. U. S. A.* 102, 6522–6526

63 Orr, H.A. and Presgraves, D.C. (2000) Speciation by postzygotic isolation: forces, genes and molecules. *BioEssays* 22, 1085–1094

64 Ting, C-T. *et al*. (1998) A rapidly evolving homeobox at the site of a hybrid sterility gene. *Science* 282, 1501–1504

65 Sun, S. *et al*. (2004) The normal function of a speciation gene, *Odysseus*, and its hybrid sterility effect. *Science* 305, 81–83

66 Barbash, D.A. and Ashburner, M. (2003) A novel system of fertility rescue in *Drosophila* hybrids reveals a link between hybrid lethality and female sterility. *Genetics* 163, 217–226

67 Barbash, D.A. *et al*. (2004) Functional divergence caused by ancient positive selection of a *Drosophila* hybrid incompatibility locus. *PLoS Biol.* 2, 0839–0848

68 Presgraves, D.C. *et al*. (2003) Adaptive evolution drives divergence of a hybrid incompatibility gene between two species of *Drosophila*. *Nature* 423, 715–719

69 Sawamura, K. and Yamamoto, M-T. (1993) Cytogenetical localization of *Zygotic hybrid rescue (Zhr)*, a *Drosophila melanogaster* gene that rescues interspecific hybrids from embryonic lethality. *Mol. Gen. Genet.* 239, 441–449