GENES AND SPECIATION

Chung-I Wu and Chau-Ti Ting‡*

It is only in the past five years that studies of speciation have truly entered the molecular era. Recent molecular analyses of a handful of genes that are involved in maintaining reproductive isolation between species (speciation genes) have provided some striking insights. In particular, it seems that despite being strongly influenced by positive selection, speciation genes are often non-essential, having functions that are only loosely coupled to reproductive isolation. Molecular studies might also resolve the long-running debate on the relative importance of allopatric and parapatric modes of speciation.

NEO-DARWINIAN (SYNTHESIS)
The modern theory of evolution that combines both natural selection and population genetics, in which the Darwinian concept of spontaneous variation is explained in terms of mutation and genetic recombination.

CO-ADAPTATION
Selection by which harmoniously interacting genes accumulate in the gene pool of a population.

**Department of Ecology and Evolution, University of Chicago, Chicago, Illinois 60637, USA.*
‡Department of Life Science, National Tsing Hua University, Hsinchu, Taiwan 300, Republic of China.
Correspondence to C.-I W.
e-mail: ciwu@uchicago.edu
doi:10.1038/nrg1269

What is a species? How do new species come into existence? These questions are some of the most enduring in biology and remain controversial today. Under many commonly accepted species definitions, speciation can be viewed as the process by which two identical populations diverge genetically to the point at which their subsequent merger would not be possible. Species are therefore both genetically distinct and independent. Although distinctness is often observable (in morphology, for example), independence usually is not.

How do species evolve to become phenotypically distinct? What are the underlying genes? What forces drive their divergence between species? Some insights into these questions might come from studying genes that cause reduced fitness in hybrids that are intermediate in phenotype between two species. The fitness reduction can range from ecological maladaptation or behavioural aberration to inviability or sterility. The loci that underlie such reductions in fitness might be considered 'speciation genes', which are important in driving the nascent species to become independent genetic entities. Below, we review recent progress in the characterization of speciation genes.

Operationally, species are often delineated by distinct phenotypes, such as distinct plumage in birds. It was the introduction of the concept of reproductive isolation (RI) that redirected the emphasis to the independent nature of species[1–4]. According to Mayr[3], species are "groups of interbreeding natural populations that are reproductively isolated from other such groups"[3]. RI therefore refers to the independence of gene pools, among which new mutations and allele frequency changes are not shared. A central question about RI is whether this independence, or non-sharing, should apply to every locus in the genome.

Those who argue for the primacy of RI in speciation[1,3] are essentially arguing for a 'whole-genome' concept[5] (see BOX 1). If we apply the concept of RI to only a portion of the genome, where would we draw the line? Does it make sense to say that 75% of the genome is reproductively isolated? Indeed, in this NEO-DARWINIAN view of RI, genetic changes between species are seen to be so strongly CO-ADAPTED that few genes can be integrated into the genome of another species. So, almost all regions in the genome are either part of such a 'cohesive' network or are closely linked to an element in such a network. In either case, gene flow across nascent species boundaries is effectively eliminated[6].

By contrast, the alternative is a genic view of speciation, as explained in FIG. 1 (see also REFS 5,7). Although rarely recognized as such, these different perspectives on the genetic architecture of species differences are the genesis of the long-running debate on the geographical mode of speciation; that is, whether speciation most commonly occurs when the diverging populations are in allopatry, parapatry or sympatry (see BOX 1). The central issue is whether two populations can evolve into good species while they continue to exchange genes during the process. Under the whole-genome view that gives primacy to RI in speciation[3], gene flow between diverging populations would have such negative effects on those populations that strict geographical barriers would have to be the prelude to speciation. The alternative genic

view (FIG. 1), in which a subset of genes determines RI or differential adaptation and which allows the possibility of gene flow between nascent species, underlies the concept of parapatric speciation. Although the idea that most speciation events occur in allopatry has dominated thinking in the second half of the last century, the genetic evidence has been marginal. Support has come mainly from ecological or biogeographical data (see Howard and Berlocher[8] for a review). Importantly, at the genic level, allopatric speciation provides a well-defined null hypothesis. For the first time, we might have the wherewithal to test this model definitively, owing to the genomic tools that have been developed in the past five years.

So, even an apparently ecological issue, such as whether speciation most frequently occurs in allopatry or in parapatry, depends on which of the genic or whole-genome views of speciation is most correct. Speciation research has just crossed the threshold into the molecular era. It is these molecular studies that will ultimately resolve whether the genic view or the whole-genome view provides a more accurate representation of speciation.

Here, we review the exciting new molecular studies that are beginning to clarify the genetic basis of speciation. First, we discuss studies that are starting to show how many genes underlie RI and how they interact together. We then go on to summarize some important recent studies that have begun to shed light on the nature of a 'speciation gene' before discussing attempts to assess the population-genetic mode of speciation on a genome-wide scale.

The genetic architecture of speciation

The number of genes underlying RI and the interactions among them define the genetic architecture of speciation. Studies of *Drosophila* species have provided the most extensive data on the genetic architecture of RI and, therefore, we focus on these studies (see REFS 9–11 for more detailed reviews). In general, even closely related sibling species of *Drosophila* show extensive functional-genetic differentiation. For example, standard quantitative trait locus analysis indicates that there are approximately 20 genes that control the difference in the shape of the genital arch between *D. simulans* and *D. mauritiana*[12].

Box 1 | The concept and classification of reproductive isolation

The basic concept

Reproductive isolation (RI) is a population-genetic concept that refers to the non-exchange of genes between two species that are in contact with each other. The cessation of gene flow is the result of the genetic properties of the two species in question, and not the result of extrinsic barriers that prevent contact. One common mistake is to equate RI with a reduction in fitness in the hybrids between two forms. As long as some hybrids can still pass genes back and forth between the two forms, even at a very low rate, there is effectively only one gene pool. Furthermore, RI is a whole-genome concept; if two diverging populations still share a common gene pool for some parts of the genomes, these populations cannot be considered to be reproductively isolated (but see Mayr[73] for an alternative view).

Pre-mating versus post-mating isolation

The mechanisms of RI can be divided into those that act before and those that act after mating. Pre-mating isolation refers to the absence of interspecific hybrids because the two species do not mate owing to ecological or behavioural factors. In post-mating isolation, members of the two species do mate but the hybrids are inviable, sterile or ecologically maladapted. Obviously, pre-mating isolation does not incur as much cost in reproductive efforts as post-mating isolation does. Mechanisms of post-mating isolation can be further divided into those that act before fertilization (pre-zygotic) and those that act after fertilization (post-zygotic).

The geographical modes of speciation

In allopatric speciation, the diverging populations have to be geographically separated with no gene flow between them for divergence to take place (see also FIG. 3). Parapatric speciation refers to populations that are usually geographically separated, but connected by gene flow, during speciation (see also FIG. 3). Sympatric speciation is the extreme form of parapatric speciation, with almost unrestricted gene flow and few differences in the ecological niches of the diverging populations.

The genetic basis of RI

There are three broad classes of genetic element that contribute to RI: extra-chromosomal, chromosomal and genic. Extra-chromosomal elements include cytoplasmic symbionts and transposable elements that can cause RI. Chromosomal elements include rearrangements that can cause aberrant meioses in interspecific hybrids. However, the most common mode of chromosomal speciation is polyploidization, in which the genome size is doubled by having two sets of the same genome (autotetraploidy) or by having two sets of different genomes (allotetraploidy). Genic RI is caused by the incompatibilities between the genes of the diverging species (see also BOX 2).

The definition of a speciation gene

In the genic view of speciation[5], speciation genes are those that contribute to RI, often in the form of hybrid inviability, sterility or behavioural aberration. This definition can include genes that cause isolation owing to physiological, behavioural or even ecological factors. For example, an INTROGRESSION hybrid might be viable and fertile but less cold-resistant. The underlying genes would still fail to successfully migrate across the incipient species boundary.

INTROGRESSION
The integration of a genomic region from one species into the genome of another species. Even a few percent of the introgressed genome can lead to hybrid incompatibility.

Similarly, at least 15 genes control the differences in mating behaviour between two behavioural races in *D. melanogaster*[13,14]. However, in at least one study of the differences between *Drosophila* species, the genetics of sexual isolation seem to be much simpler, with only a few loci being involved[15] (see REF. 13 for a discussion of this contrast).

The most extensive studies of the genetic architecture that underlies RI have focused on hybrid male sterility between *D. simulans* and *D. mauritiana* and hybrid inviability between the more distantly related *D. simulans* and *D. melanogaster*. To dissect the genetic architecture of RI, such studies attempt to POSITIONALLY CLONE genes that are involved in hybrid incompatibility (see FIG. 2). The number of genes that contribute to hybrid male sterility between *D. simulans* and *D. mauritiana* is estimated to be at least 50 when introgressed in the homozygous state[10]. This type of architecture can be characterized as 'weak effect–strong interaction' — that is, each individual gene has little effect on its own, but in combination, these genes cause reproductive incompatibility[10]. Functional studies of one of these interacting genes (*OdsH*; see below) have indeed shown that it has an extremely weak effect. Despite the weak individual effects of these genes, it was estimated that the hybrids between these species are still sterilized 15 times over[16].

It is important to point out that, in the above comparisons between *D. simulans* and *D. mauritiana*, there is virtually no hybrid inviability despite extensive hybrid male sterility. The extreme rapidity with which hybrid male sterility has evolved, relative to hybrid inviability or female sterility, has become a hallmark of post-mating isolation. Both the causes and the consequences are important[17,18]. The contrast in the genetic basis of hybrid sterility compared with hybrid inviability is notable not only for the difference in the number of loci involved but also in how the genes interact. Studies of two divergent species in which hybrid inviability has evolved (*D. melanogaster* and *D. simulans*) have shed more light on these differences[19]. DEFICIENCY MAPPING and ALLELE COMPLEMENTATION (FIG. 2b) were used to study hybrid male sterility[20,21] and hybrid inviability[19,22] between these species. Intriguingly, the complementation test showed that no single locus rescued fertility of the hybrid male. On the other hand, single genes with a substantial effect seemed common for hybrid inviability (see also below).

In summary, with respect to RI between closely-related species, *Drosophila* studies have shown that genetic differentiation is often extensive. Although other studies have indicated a much simpler genetic architecture outside *Drosophila*, the apparent discrepancy might be the result of a difference in the mapping resolution[10].

The molecular genetics of speciation

To understand the molecular basis of RI, three key questions need to be addressed. First, what genes contribute to RI? Second, what are the normal functions of those genes? Third, how did these normal functions diverge among different populations, leading to RI? The fact that the genes that underlie post-mating isolation must have normal functions that are distinct from their role in RI is an important point to keep in mind. After all, the function of RI genes could not possibly be to sterilize their carriers. In this way, the analysis of RI is not unlike the study of graft rejection in organ transplantation: the biology of the major histocompatability complex, which underlies graft rejection, certainly has much wider and more profound implications than the phenomenon of graft rejection itself. In modelling the evolution of RI, the practice has been to consider only the RI phenotype without addressing the underlying function (for example, see BOX 2). The reason for such a glaring omission is clear — the identities of speciation genes, and so, their normal functions, have not been known until very recently.

Now, there are a handful of studies in which the identities of speciation genes have been shown: each of which we discuss. By definition, a speciation gene is one that can be shown to cause some degree of ecological, sexual or post-mating isolation between young, or even nascent, species. Although there have been other claims of speciation genes being identified, we consider the five that we discuss to be the only studies that have truly identified the molecules involved. There are many excellent studies that focus on genes that differ between species and the molecular interactions between them but that do not address their phenotypic effects on the whole organism (for example, see REF. 23). Until these further studies have been done, such genes cannot be classified as speciation genes.

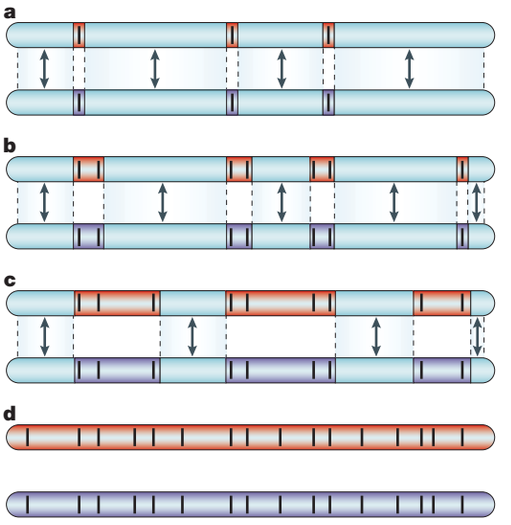


Figure 1 | **The genic view of species differentiation.** The two horizontal bars represent the genomes of two diverging populations. When they start to differentiate, only a few loci (indicated by black lines) are differentially adapted and genes at such loci are not exchanged between populations (**a**). Gene flow continues in the rest of the genome (arrows). Although the regions of differential adaptation expand, the amount of gene flow between the two genomes is gradually reduced owing to linkage with such regions (indicated in red/purple) (**b**,**c**) until the two populations are completely reproductively isolated and are therefore considered to be separate species (**d**). Modified with permission from REF. 5 © (2001) Blackwell Science.

POSITIONAL CLONING
The procedure by which we identify and isolate genes on the basis of their location in the genome, involving detailed genetic and physical maps of chromosomes.

DEFICIENCY MAPPING
Uses chromosomes that have different sections deleted to locate the position of a gene of interest. Without the deficiency, the normal functional gene usually masks the effect of (that is, complements) the defective or foreign copy that we wish to identify.

ALLELE COMPLEMENTATION
A test of whether a wild-type phenotype can be restored with two given alleles in a diploid genome.

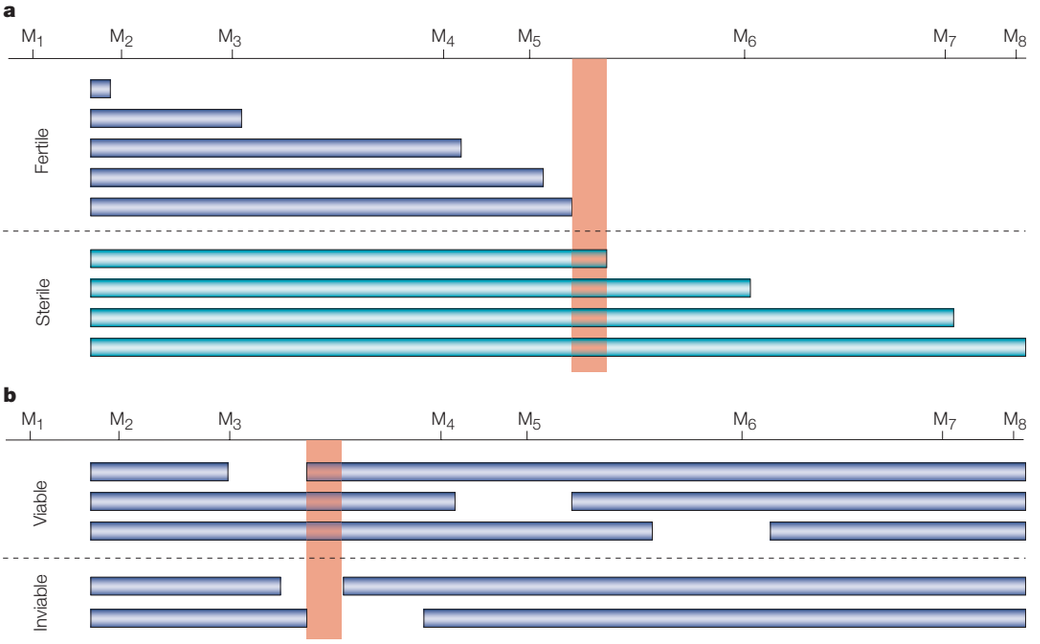


Figure 2 | **Positional cloning of hybrid incompatibility genes. a** | Recombination mapping: the coloured horizontal bars indicate introgressed chromosomes. The vertical orange bar marks the position of the gene that is responsible for hybrid sterility (for example, the *Drosophila* gene *OdsH*). Different sterile and fertile hybrids are typed for markers on the introgressed chromosomes that are known to be polymorphic between the two species (M_1–M_8). The pattern of marker distribution among the different hybrids allows the location of the hybrid sterility gene to be mapped at a resolution that is dependent on the concentration of informative markers in the region. **b** | Deficiency mapping: the phenotypes that are associated with various overlapping chromosome deficiencies (indicated by the gaps in the bars) show the position (marked with the vertical orange bar) of the hybrid incompatibility gene (for example, the *Drosophila* gene *Nup96*). Viable and inviable hybrids with different chromosome deletions are typed for markers that are known to be polymorphic between the two species (M_1–M_8). The pattern of marker distribution among the different hybrids allows the location of the hybrid inviability gene to be mapped at a resolution that is dependent on the concentration of informative markers in the region.

Although the definition of speciation genes includes those with strong or weak effects on ecological, behavioural or physiological differences, most of these initial studies have concentrated on genes that have large effects on physiological characteristics. Four of these five examples focus on post-mating isolation. The final example, which concerns ecological adaptation between behavioural races, is therefore of considerable interest.

Melanoma formation in Xiphophorus species hybrids (Xmrk-2). Many species in the fish genus *Xiphophorus* have spots on their skin that are composed of black pigment cells. In interspecific hybrids between *X. maculatus* (platyfish) and *X. helleri* (swordtail), these spots sometimes spontaneously develop malignant melanomas[24–26]. A two-locus Dobzhansky-Muller (DM)-type model (see BOX 2) has been proposed to explain the formation of malignant melanomas. In this model, overexpression of the *Tu* gene causes these melanomas to form. The second locus involved, called the *R* gene, is a suppressor that negatively controls *Tu*. The platyfish contains both *Tu* and *R* genes, whereas the swordfish contains neither. In the backcross F_2 hybrids, a quarter of the offspring produce melanomas owing to the presence of *Tu* but the absence of the *R* gene.

The X-linked *Tu* locus was subsequently mapped to a candidate gene, *Xmrk-2* (REFS 27–29). *Xmrk-2* encodes a transmembrane growth factor of the RECEPTOR TYROSINE KINASE SUPERFAMILY that is important in signal transduction. Its closest homologue in humans is the epidermal growth factor receptor (*EGFR*)[30]. All the features of *Xmrk-2* are consistent with those of the dominant ONCOGENE that causes the melanomas in the hybrid fish. In particular, mutations at the *Xmrk-2* locus abolish the *Tu* phenotype and the overexpression of *Xmrk-2* gives rise to a high frequency of tumour formation.

In the adjacent genomic region, another *EGFR* homologue, *Xmrk-1*, was found in all *Xiphophorus* fish. *Xmrk-1* and *Xmrk-2* are therefore duplicated genes. However, *Xmrk-1* transcripts can be found in all tissues, whereas *Xmrk-2* transcripts are only abundant in the melanomas of the hybrids. *Xmrk*-2 apparently originated from non-homologous recombination between *Xmrk-1* and an adjacent *D* locus[31]. So, this hybrid locus has the regulatory region from the *D* locus and most of the coding regions from *Xmrk-1*. The *R* gene represses *Xmrk-2* as well as the *D* locus. Another important difference between *Xmrk-1* and *Xmrk-2* is the two amino-acid replacements in the extracellular domain, which shows ligand-independent activation[32]. So, divergence after gene duplication is important in the differentiation of these species and this might be a common feature of speciation genes (see also below).

RECEPTOR TYROSINE KINASE SUPERFAMILY
One of the important cell-surface receptors that interacts with water-soluble ligands.

ONCOGENE
A gene that induces uncontrolled cell proliferation.

MATERNAL EFFECT
The effect of the maternal genotype on the phenotype of the offspring, or the zygotes, usually at the embryonic stage (see also zygotic effects).

ZYGOTIC EFFECT
The effect of the zygotes' own genotype on their own phenotype.

Xmrk-2 induces tumour formation only in the hybrids in which the *R* gene is absent, in accordance with the classical DM model of post-mating isolation. In other species, such as the medaka, overexpression of *Xmrk-2* also causes embryonic lethality. The constitutively expressed *Xmrk-2* activates a transcription factor, STAT5, and subsequently upregulates several downstream targets[33]. So, overall, there is strong circumstantial evidence that *Xmrk-2* is a speciation gene. However, the multiple alleles at each locus in the natural populations of each species, all of which cause a different degree of hybrid phenotype, require further study. It is possible that some of these alleles are not becoming fixed but rather are simply deleterious mutations in the process of being removed by purifying selection. Those deleterious alleles that are destined to be removed would not contribute to species differentiation.

Hybrid male sterility in Drosophila species (OdsH). The *Odysseus* (*OdsH*) gene from *D. mauritiana* causes complete male sterility when co-introgressed with the adjacent segment into *D. simulans*. Genetic mapping of the male sterility locus (FIG. 2a) allowed the initial identification of *OdsH* as this RI locus[34]. Recent transgenic studies have confirmed this identification[35]. *OdsH* is a homeobox gene from a family of transcription factor-encoding genes that are known to be slowly evolving. Curiously, *OdsH* has been evolving rapidly within the *D. melanogaster* subgroup even though its homologues from other species are extremely conservative.

The most direct approach to assess the normal function and the phenotypic effect of a specific gene is to knock it out. Interestingly, the deletion of *OdsH* from *D. melanogaster* results in no obvious adverse phenotype[35]. So, at least at this crude level of observation, *OdsH* is dispensable. However, a more detailed examination showed a subtle effect: males missing *OdsH* suffer a 40% fertility reduction when they are two days old and mate repeatedly. This fertility reduction lessens to 20% and 8% in the next two days, also under sperm-exhaustion conditions. After five days, the role of *OdsH* in fertility enhancement vanishes. One interpretation of these findings is that the role of *OdsH* is to accelerate the maturation of sperm. So, only very young males under sperm-exhaustion conditions are affected.

Comparative analyses indicate that *OdsH* was duplicated in the *Drosophila* lineage from a neuron-expressed gene, *unc-4*, after it diverged from the mosquito lineage. Whereas *unc-4* in *Drosophila* has not diverged much in either sequence or expression from the ancestral state in the common ancestor it shares with mouse and *C. elegans*, *OdsH* has changed in both sequence and expression. Specifically, *OdsH* has been evolving away from the *unc-4* pattern of embryonic and neuronal expressions to a testicular role[35].

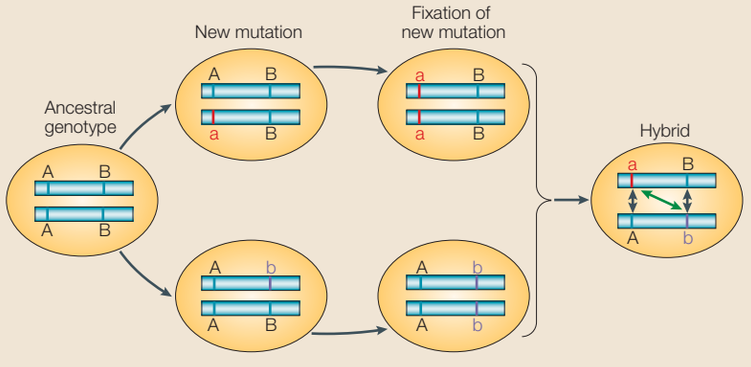
Because *OdsH* is divergently regulated between *D. simulans* and *D. mauritiana*, its expression in the testis of the sterile hybrids is highly misregulated. *OdsH* transcripts accumulate in very young spermatocytes. The pattern is not observed in either parental species or in the fertile introgression line, which differs from the sterile line by a 3-kb segment of *OdsH* (REF. 34). The expression of *unc-4* in the sterile introgression line is also normal. Therefore, the divergence in the sequence and the expression of *OdsH* might both contribute to the hybrid sterility.

Hybrid inviability in Drosophila species (Hmr). Another classical RI system in *Drosophila* is the hybrid incompatibility between *D. melanogaster* and *D. simulans*, two species that have been reproductively isolated more than 2.5 million years. Crosses between these species produce only inviable or sterile hybrids[36–38]. Five mutations that could rescue the inviable F_1 hybrid progeny have been found and several were elegantly characterized for their MATERNAL or ZYGOTIC EFFECTS[39–43]. Given the multi-locus nature of hybrid incompatibility, it was surprising that such hybrid-rescue mutations could be identified.

Among the hybrid-rescue mutations, the X-linked *Hmr* (hybrid male rescue) gene, which rescues the inviable hybrid males, was mapped and cloned[44,45]. Hmr was identified as a transcription factor in the myeloblastosis family[44]. The primary amino-acid sequence of Hmr contains two DNA-binding protein motifs that indicate its role in transcription regulation. There were many amino-acid substitutions between the sibling species in the DNA-binding domains of Hmr. So, this pattern indicates that positive selection might drive the rapid evolution of *Hmr*. However, the *Hmr* mutation

Box 2 | Dobzhansky-Muller model of hybrid incompatibility

How does genic incompatibility between species evolve without simultaneously causing defects in pure species? A popular explanation is the Dobzhansky-Muller (DM) model of hybrid incompatibility. In the ancestral population, the genotype is *AA BB*. When the population is split into two, *A* evolves into *a* in one population and *B* evolves into *b* in the other. *a* and *b* are mutually incompatible. As the a–b interaction is not present in the pure species, the evolution of incompatibility is possible. Detailed genetic analysis on hybrid male sterility, however, has shown that hybrid incompatibility often involves conspecific genic interactions as well[9,10,18,20]. The DM model is far too simple in that respect. In the far right (hybrid), the divergence process is indicated by the black double-headed arrows and the incompatibility is indicated by the green double-headed arrow. The relationship between the black and green arrows should be the essence of the evolution of reproductive isolation. A deficiency of the DM model is that it does not consider the divergence process that is indicated by the black arrows but focuses instead on the incompatibility, which is a byproduct of that divergence.



that rescued hybrid viability was a *P*-element insertion in its 5′ region that resulted in a reduction in the amount of wild-type transcript. For *Hmr* to be considered a true 'speciation gene', it would be necessary to show that the *D. simulans* and *D. melanogaster* alleles are functionally divergent in their rescue effect of hybrid viability. A recent transgenic study indicates that this might indeed be the case (D. Barbash, personal communication).

Hybrid inviability in Drosophila species (Nup96). Complementation mapping (FIG. 2b) has been used to analyse hybrid inviability between *D. melanogaster* and *D. simulans*. High-resolution mapping has allowed a speciation gene, *Nup96*, to be cloned and characterized[22]. The *Nup96* allele from *D. simulans* causes inviability in the F_1 hybrids if the copy from *D. melanogaster* is absent. *Nup96*, which has homologues in yeast, worm and human genomes, encodes a subunit of a nuclear-pore complex, which transports macromolecules between the nucleus and cytoplasm[46] and is therefore essential for viability in flies.

An excess of non-synonymous substitutions in *Nup96* between *D. melanogaster* and *D. simulans* relative to non-synonymous polymorphisms within these species (calibrated against synonymous changes with the MCDONALD AND KREITMAN TEST) indicated that this gene is under positive selection. With the sequences from *D. mauritiana* and *D. yakuba*, it was possible to map putative adaptive changes onto an evolutionary tree. Presgraves *et al.*[22] concluded that the adaptive changes occurred in the distant past, a suggestion that is corroborated by the analysis of the extant polymorphisms in *D. melanogaster* and *D. simulans*. Had some adaptive changes occurred recently, a reduction in the amount of neutral polymorphism, which might also be accompanied by a skew towards very low- and/or very high-frequency variants, would have been expected. Neither was observed in *Nup96*.

Not only were Presgraves *et al.*[22] able to map the *Nup96* gene, but they were also able to locate the interacting locus in the DM model of hybrid incompatibility to the X chromosome. They did this by switching the source of the X chromosome in the hybrid males. One question to be addressed in the future is whether there are multiple loci on the X chromosome that interact with *Nup96*.

Ecological/behavioural races in Drosophila melanogaster (desat-2). The final example of a proven speciation gene (under our broad definition; see BOX 1) provides a glimpse of the molecular genetics of ecological, and possibly behavioural, isolation. *D. melanogaster* from central-southern Africa around Zimbabwe and those from the rest of the world (referred to as the Z and M types, respectively) have evolved to become different ecological/behavioural races. The females of African and cosmopolitan *D. melanogaster* carry different forms of a specific type of non-volatile CONTACT PHEROMONES. These two forms — the 5,9-heptacosadiene and 7,11-heptacosadiene forms of the 27-carbon cuticular hydrocarbons (CH)[47] — differ in the position of a double-bond in a long chain of saturated hydrocarbons. Two independent studies have identified the gene that controls the (5,9)/(7,11) difference to be a *desaturase* gene, *desat2* (REFS 48,49). Although CHs often act as contact pheromones between sexes, they have also been implicated in ecological adaptations, such as heat or starvation tolerance[50].

The *desat2* gene apparently diverts the synthesis of 7,11-heptacosadiene into the 5,9-type. The loss of the promoter in the *desat2* gene therefore results in the 7,11-type among the M flies. This observation raises the interesting possibility that loss of function of a gene has a role in this particular case of nascent speciation. The geographical distribution of the two *desat2* variants (predominantly $desat2^{+}$ in Africa and $desat2^{0}$ elsewhere) indicates that this strong differentiation must be maintained by differential selective pressure. An excess of high-frequency nucleotide mutations highlighted the influence of positive selection on the *desat2* polymorphism[49]. Greenberg *et al.*[17,50] were able to show, by gene knock-out, that the loss of the *desat2* gene (as in non-African M flies) results in an increase in cold tolerance and a decrease in starvation tolerance. It is plausible that, in the colder climate, a non-functional *desat2* would spread through the cosmopolitan populations. So, this seems to be a case of ecological adaptation and differentiation.

An interesting aspect of the Z–M differentiation is the unidirectional sexual isolation between these forms[51]. Zimbabwe females, in the presence of Z and M males, do not mate with M males. (Note that the observation by itself does not indicate male or female choice.) We know that at least seven or eight genes control female or male mating behaviour, respectively[13,14]. So, the question is whether *desat2* is one of the loci that governs Z females' mating characteristics (for example, reduced attractiveness to M males). CH differences have been known to govern females' attractiveness in interspecific crosses[52]. However, it was widely thought that *desat2* was not involved in female attractiveness in the Z–M system because Caribbean flies, which carry the African *desat2* allele, behave like M flies. Nevertheless, recent observations have shown that, within three African populations, the presence of the African *desat2* allele correlates nearly perfectly with Z-femaleness[53]. One possible interpretation of this pattern is that *desat2* governs female attractiveness to M males and that the Caribbean population is an anomaly that results from recent admixture between African and North American flies. Although this interpretation seems to contradict the widely-accepted view that *D. melanogaster* males might not be discriminatory when choosing a mate[54], new work indicates that M males might not court Z females as ardently as they court M females, especially when the females are not highly receptive (C.-T.T. and C.-I W., unpublished observations). If this is the case, the *desat2* gene might be playing a double role in this nascent speciation through differentiation in ecological adaptation and, secondarily, through mating preference.

MCDONALD AND KREITMAN TEST
A test that contrasts interspecific divergence against intraspecific polymorphism. It is a powerful test to detect excess of non-synonymous substitutions between species.

CONTACT PHEROMONES
Chemical signals that are transmitted through the direct physical contact of two individuals. Contact pheromones in *Drosophila* are often sexual signals.

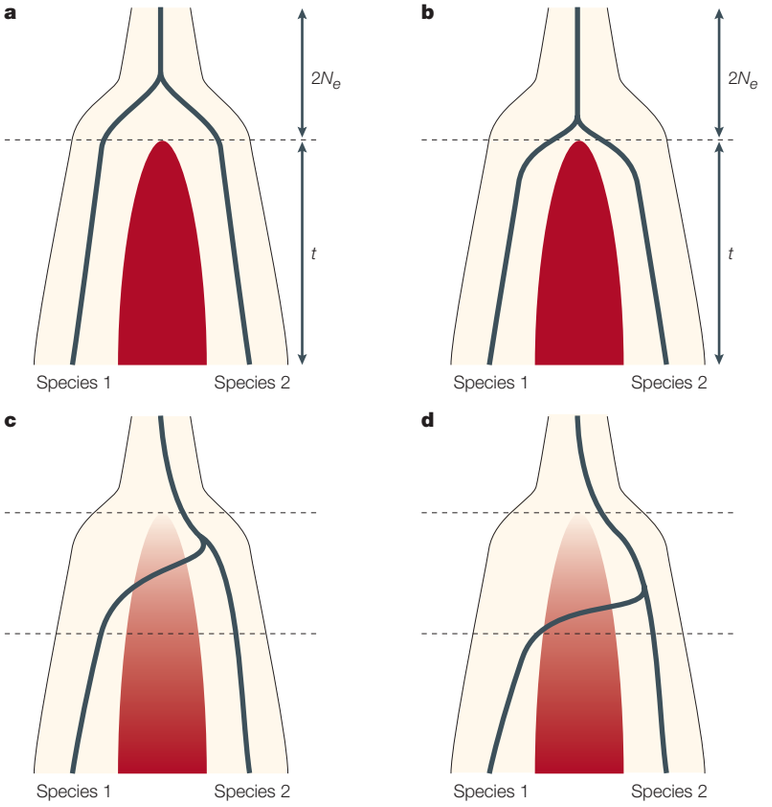


Figure 3 | **Geographical models of speciation.** The strength of the barrier to gene flow between nascent species is indicated by the intensity of the red shade between the two branches. The black line represents a gene tree as it diverges between the two species. The wider yellow-shaded areas represent the whole genome of the two diverging species and the parent species from which they derived. **a**,**b** | The allopatric model of speciation in which the barrier is complete at a fixed time for all genes. The difference in divergence time between the two loci that are depicted was present at the time of allopatric speciation (which is the minimum divergence time between any two homologues in the two species). **c**,**d** | The parapatric model of speciation in which there is an extended period of time during which some genes can permeate through the nascent species boundary whereas others cannot. N_e, effective population size; t, time since species divergence.

The population genetics of speciation

Elucidating the molecular basis of speciation through identifying and studying speciation genes attacks the problem of speciation from the bottom up. At the same time, genomic approaches are also making top-down studies much more powerful. In particular, now that speciation studies are entering the molecular era, the debate on the pervasiveness of allopatric speciation seems resolvable.

Testing the allopatric model with genomic data. The allopatric model makes a strong prediction — that all genes between the same two species have identical species-divergence time (*t* in FIG. 3a,b). These allopatric genealogies are not identical in all cases because of differences that were previously present when the species split ($2N_e$ in FIG. 3a,b)[55–59]. However, they are distinct from the genealogical histories of two genes that became non-introgressable at different times (FIG. 3c,d; see also FIG. 1). As discussed earlier under a neo-Darwinian scheme — in which extrinsic (for example, geographical) barriers to gene flow are necessary before the onset of speciation — such a scheme, as depicted in FIG. 3c,d, is improbable.

One approach to addressing the question of whether gene flow can occur during speciation (that is, parapatric speciation) was elegantly modelled by J. Hey, J. Wakeley and colleagues[60–63]. This approach considers both divergence between species and polymorphism within species. If there has been continual gene flow during speciation, the variance in the numbers of shared polymorphisms and fixed differences would increase above that predicted for speciation that has occurred in strict allopatry (see FIG. 1a,b). Machado *et al.*[63] used this approach to show that gene flow occurred between *D. pseudoobscura* and *D. simulans* during speciation (see also REF. 61).

Although an analysis that incorporates within-species polymorphisms is useful for understanding the population genetics of speciation in *Drosophila* (for example, Kliman *et al.*[62]), it might not always be so useful in other groups. For example, most polymorphisms in humans can actually be traced back to common ancestral alleles that can be dated to no more than a million years ago. Such polymorphisms can tell us nothing about the human/chimpanzee speciation event that took place more than five million years ago. Furthermore, data on within-species polymorphisms in closely-related species are usually not available for many loci. An approach that circumvents these problems focuses on one genome from each species, but looks at data from hundreds of loci, and asks the simple question: is the *t* value constant across a large number of genes (see FIG. 3)?

In principle, *t* should be reflected in the amount of neutral divergence between the species. Therefore, we would expect greater variation in K_s (number of synonymous substitutions per synonymous site) when speciation occurs in parapatry (FIG. 3c,d) than when it occurs in allopatry (FIG. 3a,b). An outgroup can be used to calibrate the divergence between the species and so account for any variation in K_s owing to mutation or selection. The residual variation in K_s across loci should then be a function of *t* and $2N_e$ in FIG. 3 (REFS 55,57–59,64). The larger the variation in K_s, the larger N_e has to be to account for the observation when *t* is assumed to be a constant. Given the variation in K_s between primates, it has been necessary to invoke a large N_e for the ancestral human/chimpanzee population — often much larger than the N_e of the extant human populations[65–67].

An alternative explanation for the large among-gene variation in K_s is that there is large variation in *t*. In a new study, the null hypothesis of a constant ratio of $t/2N_e$ can be rejected for 347 coding versus 143 intergenic regions compared between humans and chimpanzees (N. Osada and C.-I W., unpublished observations). This indicates that intergenic regions might have remained introgressable across the nascent species boundary between human and chimpanzee long after many coding regions have ceased to be able to introgress. We should note that the null hypothesis that was rejected in these cases was based on the

simplest kind of allopatric speciation. There is a class of models that overlay allopatry on species with a deep population structure. Such models, which are intermediate between allopatric and parapatric ones, deserve further study.

Another recent study also addressed the question of parapatric speciation from a different angle[68]. The authors observed that the K_A/K_S RATIOS between human and chimpanzee are higher for genes on rearranged chromosomes than on collinear ones, in agreement with the parapatric model[68]. However, a separate analysis showed that the K_a/K_s pattern is observable, with nearly identical values, between human and orangutan or human and macaque[69]. So, although the observations in this study are meaningful in other respects, given the positive result from the 'negative control', the interpretation of parapatric speciation cannot be supported[68,70].

Strictly speaking, the discussion in this section is relevant to animal studies only. Plant literature is replete with references to hybridization and introgression during speciation. 'Hybridization speciation', in which a third species is formed by mixing the genetic materials from two parental species[71], is another important demonstration. What might be the basis of these discrepant views between animal and plant literature? It is possible that plant genomes are more modular, such that mixing components from different sources can still make a well-fit plant. If that is true, 'allopatric genealogy' (FIG. 3a,b) in plants should often be rejected using the type of analysis described above.

Implications and perspectives

Speciation is not an easy subject. However, the myth that speciation is 'the mystery of mysteries' or that it is both 'unknown and unknowable' has not helped us to understand the subject. The perceived difficulties with the topic stem largely from a lack of knowledge at the fundamental genic level. For example, how could we hope to understand post-mating isolation when the phenotype that defines the class of genes that we are interested in does not even contain a hint of the original function for which each of these genes evolved? The five cases discussed above represent all those we could find that fulfil the criteria of a 'speciation gene'. Nevertheless, even from this limited set, which is based mainly on one taxonomic group (four of them are from *Drosophila*), the range of the molecular identity, as well as the underlying principle, is broad. Three of the five cases are related to transcriptional regulation (*Xmrk*, *OdsH* and *Hmr*), supporting the common postulate that species divergence is regulatory in nature. Moreover, in *desat2,* which is not a regulatory gene, the change during INCIPIENT SPECIATION is nevertheless in the regulatory region.

The nature of speciation genes is also likely to be a function of the age of the speciation event. The range of genes underlying RI between highly differentiated species, such as *D. melanogaster* and *D. simulans*, might be different from that between incipient species. As the divergence increases, more genes might contribute to RI and the range might become broader and more evenly reflective of the entire genome. So, it is the range of genes that are involved during incipient speciation that holds the greatest interest. In all five cases, there is substantial divergence in DNA sequences and, among several of them, in expression as well. In at least four of the five cases, positive selection has driven the divergence. The results are relevant to the debate on whether RI might have evolved neutrally in the absence of adaptive forces.

We propose that, for a gene to diverge in function, it needs to be released from its old functional niche. The release would then allow the diverging species to use the gene differently. We shall refer to this hypothesis as the genetic 'niche-release' hypothesis. There might be two different ways for a gene to experience niche release — environmental change and gene duplication. Divergence in *desat2* is an example of gene divergence that has occurred as a result of environmental change that took place when the flies migrated out of Africa, whereas *OdsH* and *Xmrk* are both the result of gene duplication. (Note that this present hypothesis is different from a previous model on gene duplication and RI[72].) Whichever mechanism is involved, a gene under niche release should be more likely to be functionally dispensable. Operationally, functional dispensability means that the deletion of the gene would not lead to lethality, sterility or other forms of severe fitness reduction. Such dispensable genes might be prone to diverge in function, often becoming non-functional. Under this definition, both *OdsH* and *desat2* are dispensable.

In the classical DM model, the emphasis has always been on the incompatibility interaction (the green double-headed arrow in BOX 2), but we might ask a deeper question — whether the process of divergence (the black double-headed arrows) and the resultant interaction are related and, if so, how they are related. Such a linkage seems obvious for *Xmrk-2* and *Nup96*. For example, in the former, the dark spots, when unregulated, become melanomas. On the other hand, it might not be surprising to find some RI genes for which the normal function and the RI phenotype are only weakly coupled, or even completely unrelated. For example, the deletion of *OdsH* has a subtle effect on male fertility but, in the appropriate genetic background, the presence or absence of the allele from *D. mauritiana* determines full fertility versus (nearly) normal fertility. Similarly, the normal function of *desat2* might be cold tolerance but a correlated response is the change in CH, or the 'perfume' on the females. In either case, the RI phenotype is out of the range of what might have been predicted on the basis of the normal function/phenotype of the speciation gene.

In the past five years of limited molecular analysis on speciation, we have learned the identities of several speciation genes. Their biological functions show the molecular bases of species differentiation. These studies should re-focus our attention to the genic basis of speciation. The tenet of speciation study, namely the concept of RI, is fundamentally a whole-genome concept and should be revisited after we have a more comprehensive understanding of genes and their roles in speciation.

K_A/K_S RATIOS
Ratios of non-synonymous substitutions to synonymous substitutions per site.

INCIPIENT SPECIATION
The initial stage of species formation during which reproductive isolation is only partial.

1. Dobzhansky, T. *Genetics of the Evolutionary Process* (Columbia Univ. Press, New York, 1970).
2. Fisher, R. A. *The Genetical Theory of Natural Selection* (Clarendon, Oxford, 1930).
3. Mayr, E. *Animal Species and Evolution* (The Belknap, Cambridge, Massachusetts, 1963).
4. Wilson, A. C., Maxon, L. R. & Sarich, V. M. Two types of molecular evolution. Evidence from studies of interspecific hybridization. *Proc. Natl Acad. Sci. USA* **71**, 2843–2847 (1974).
5. Wu, C.-I. The genic view of the process of speciation. *J. Evol. Biol.* **14**, 851–865 (2001).
In this genic view of speciation, RI develops gradually across the genome with some genes being unable to permeate through the nascent species boundary early on, whereas other genes continue to flow between them until complete RI evolves. The view suggests that parapatry is a common mode of speciation.
6. Mayr, E. Integration of genotypes: synthesis. *Cold Spring Harb. Symp. Quant. Biol.* **20**, 327–333 (1955).
7. Wu, C.-I. Gene and speciation. *J. Evol. Biol.* **14**, 889–891 (2001).
8. Howard, D. J. & Berlocher, S. H. *Endless Forms* (Oxford Univ. Press, New York, 1998).
9. Wu, C.-I. & Hollocher, H. in *Endless Forms: Species and Speciation* (ed. Berlocher, S.) 339–351 (Oxford Univ. Press, Oxford, 1998).
10. Wu, C.-I. & Palopoli, M. F. Genetics of postmating reproductive isolation in animals. *Ann. Rev. Genet.* **28**, 283–308 (1994).
11. Coyne, J. A. & Orr, H. A. The evolutionary genetics of speciation. *Philos. Trans. R. Soc. Lond. B* **28**, 287–305 (1998).
12. Zeng, Z.-B. *et al.* Genetic architecture of a morphological shape difference between two *Drosophila* species. *Genetics* **154**, 299–310 (2000).
13. Ting, C.-T., Takahashi, A. & Wu, C.-I. Incipient speciation by sexual isolation in *Drosophila*: concurrent evolution at multiple loci. *Proc. Natl Acad. Sci. USA* **98**, 6709–6713 (2001).
14. Takahashi, A. & Ting, C.-T. Genetic basis of sexual isolation in *Drosophila melanogaster*. *Genetia* (in the press).
15. Doi, M., Matsuda, M., Tomaru, M., Matsubayashi, H. & Oguma, Y. A locus for female mate discrimination causing sexual isolation in *Drosophila*. *Proc. Natl Acad. Sci. USA* **98**, 6714–6719 (2001).
16. Tao, Y., Zeng, Z.-B., Li, J., Hartle, D. L. & Laurie, C. C. Genetic dissection of hybrid incompatabilities between *Drosophila simulans* and *D. mauritiana*. II. Mapping hybrid sterility loci on the third chromosome. *Genetics* **164**, 1399–1418 (2003).
17. Wu, C.-I. & Davis, A. W. Evolution of postmating reproductive isolation: the composite nature of Haldane's rule and its genetic bases. *Am. Nat.* **142**, 187–212 (1993).
18. Wu, C.-I., Johnson, N. A. & Palopoli, M. F. Haldane's rule and its legacy: why are there so many sterile males? *Trends Ecol. Evol.* **11**, 281–284 (1996).
19. Presgraves, D. C. A fine-scale genetic analysis of hybrid incompatibilities in *Drosophila*. *Genetics* **163**, 955–972 (2003).
20. Sawamura, K., Davis, A. W. & Wu, C.-I. Genetic analysis of speciation by means of introgression into *Drosophila melanogaster*. *Proc. Natl Acad. Sci. USA* **97**, 2652–2655 (2000).
21. Sawamura, K., Roote, J., Wu, C.-I. & Yamamoto, M. Reassessment of the two-locus Dobzhansky-Muller model of reproductive isolation: extreme genetic complexity underlying hybrid male sterility in *Drosophila*. *Genetics* (in the press) .
22. Presgraves, D. C., Balagopalan, L., Abmayr, S. M. & Orr, H. A. Adaptive evolution drives divergence of a hybrid inviability gene between two species of *Drosophila*. *Nature* **423**, 715–719 (2003).
One of the five 'speciation genes' that has been characterized so far is presented in this study. The influence of positive selection on its divergence between species is clearly shown (see main text for detail).
23. Rawson, P. D. & Burton, R. S. Functional coadaptation between cytochrome c and cytochrome c oxidase within allopatric populations of a marine copepod. *Proc. Natl Acad. Sci. USA* **99**, 12955–12858 (2002).
24. Cordon, M. The genetics of a viviparous top-minnow *Platpoecilus*: the inheritance of two kinds of melanophores. *Genetics* **12**, 253–2283 (1927).
25. Kosswig, C. Uber bastarde der teleostier *Platypoecilus* und *Xiphophorus*. *Zeitschrift fur induktive Abstammungs und Vererbungslehre* **44**, 253 (1927).
26. Anders, F. Contributions of the Gordon-Kosswig melanoma system to the present concept of neoplasia. *Pigment Cell Res.* **4**, 7–29 (1991).
27. Schartl, M. A sex chromosomal restriction-fragment-length marker linked to melanoma-determining *Tu* loci in *Xiphophorus*. *Genetics* **119**, 679–685 (1988).
28. Wittbrodt, J. *et al.* Novel putative receptor tyrosine kinase encoded by the melanoma-inducing *Tu* locus in *Xiphophorus*. *Nature* **341**, 415–421 (1989).
***Tu* is perhaps one of the first genes that was identified to be incompatible between closely related species. The mechanism of fitness reduction in the hybrids is shown to be tumorigenesis (see main text for detail).**
29. Zechel, C., Schleenbecker, U., Anders, A. & Anders, F. *v-erbB* related sequences in *Xiphophorus* that map to melanoma determining Mendelian loci and overexpress in a melanoma cell line. *Oncogene* **3**, 605–617 (1988).
30. Wittbrodt, J., Lammers, R., Malitschek, B., Ullrich, A. & Schartl, M. The Xmrk receptor tyrosine kinase is activated in *Xiphophorus* malignant melanoma. *EMBO J.* **11**, 4239–4246 (1992).
31. Adam, D., Dimitrijevic, N. & Schartl, M. Tumor suppression in *Xiphophorus* by an accidentally acquired promoter. *Science* **259**, 816–819 (1995).
32. Gomez, A., Wellbrock, C., Gutbrod, H., Dimitrijevic, N. & Schartl, M. Ligand-independent dimerization and activation of the oncogenic Xmrk receptor by two mutations in the extracellular domain. *J. Biol. Chem.* **276**, 3333–3340 (2001).
33. Wellbrock, C. *et al.* Signalling by the oncogenic receptor tyrosine kinase Xmrk leads to activation of STAT5 in *Xiphophorus* melanoma. *Oncogene* **16**, 3047–3056 (1998).
34. Ting, C.-T., Tsaur, S. C., Wu, M.-L. & Wu, C.-I. A rapidly evolving homeobox at the site of a hybrid sterility gene. *Science* **282**, 1501–1504 (1998).
This study reported the cloning of the *OdsH* gene of hybrid male sterility. *OdsH* is an example of a homeobox gene, a family of genes that are generally highly conservative and are therefore unlikely candidates for speciation genes.
35. Sun, S. *Functional Analysis of the Hybrid Male Sterility Gene* Odysseus *in* Drosophila. (Univ. Chicago Press, Chicago, 2003).
36. Ashburner, M. Drosophila: *A Laboratory Handbook* (Cold Spring Harbor Laboratory Press, New York, 1989).
37. Sawamura, K., Yamamoto, M.-T. & Watanabe, T. K. Hybrid lethal systems in the *Drosophila melanogaster* species complex. II. The *zygotic hybrid rescue (zhr)* gene of *Drosophila simulans*. *Genetics* **133**, 307–313 (1993).
38. Hutter, P. Genetics of hybrid inviability in *Drosophila*. *Adv. Genet.* **36**, 157–185 (1997).
39. Watanabe, T. K. A gene that rescues the lethal hybrids between *Drosophila melanogaster* and *Drosophila simulans*. *Jpn. J. Genet.* **54**, 325–331 (1979).
40. Hutter, P. & Ashburner, M. Genetic rescue of inviable hybrids between *Drosophila melanogaster* and its sibling species. *Nature* **327**, 331–333 (1987).
41. Hutter, P., Roote, J. & Ashburner, M. A genetic basis for the inviability of hybrids between sibling species of *Drosophila*. *Genetics* **124**, 909–920 (1990).
42. Sawamura, K., Watanabe, T. K. & Yamamoto, M.-T. Hybrid lethal systems in the *Drosophila melanogaster* species complex. *Genetica* **88**, 175–185 (1993).
A detailed description of the various mutations in *D. melanogaster* that suppress hybird inviabilities.
43. Sawamura, K., Taira, T. & Watanabe, T. K. Hybrid lethal systems in the *Drosophila melanogaster* species complex. I. The *maternal hybrid rescue* (*mhr*) gene of *Drosophila simulans*. *Genetics* **133**, 299–305 (1993).
44. Barbash, D. A., Siino, D. F., Tarone, A. M. & Roote, J. A rapidly evolving MYB-related protein causes species isolation in *Drosophila*. *Proc. Natl Acad. Sci. USA* **100**, 5302–5307 (2003).
Molecular cloning of one of the suppressor mutations of hybrid inviabilities (see main text for detail).
45. Barbash, D. A., Roote, J. & Ashburner, M. The *Drosophila melanogaster hybrid male rescue* gene causes inviability in male and female species hybrids. *Genetics* **154**, 1747–1771 (2003).
46. Belgareh, N. *et al.* An evolutionarily conserved NPC subcomplex, which redistributes in part to kinetochores in mammalian cells. *J. Cell Biol.* **154**, 1147–1160 (2001).
47. Ferveur, J. F., Cobb, M., Boukella, H. & Jallon, J.-M. World-wide variation in *Drosophila melanogaster* sex pheromone: behavioural effects, genetic bases and potential evolutionary consequences. *Genetica* **97**, 73–80 (1996).
48. Dallerac, R. *et al.* A *delta 9 desaturase* gene with a different substrate specificity is responsible for the cuticular diene hydrocarbon polymorphism in *Drosophila melanogaster*. *Proc. Natl Acad. Sci. USA* **97**, 9449–9454 (2000).
49. Takahashi, A., Tsaur, S. C., Coyne, J. A. & Wu, C.-I. The nucleotide changes governing cuticular hydrocarbon variation and their evolution in *Drosophila melanogaster*. *Proc. Natl Acad. Sci. USA* **98**, 3920–3925 (2001).
50. Greenberg, A. J., Moran, J. R., Coyne, J. A. & Wu, C.-I. Ecological adaptation during incipient speciation revealed by precise gene replacement. *Science* **320**, 1754–1757 (2003).
The first application of gene replacement between nascent species shows the ecological consequence of desaturase divergence, cold tolerance, starvation intolerance and pheromonal changes. Mate choice might be affected by these pheromones.
51. Wu, C.-I. *et al.* Sexual isolation in *Drosophila melanogaster*: a possible case of incipient speciation. *Proc. Natl Acad. Sci. USA* **92**, 2519–2523 (1995).
52. Coyne, J. A., Crittenden, A. P. & Mah, K. Genetics of a pheromonal difference contributing to reproductive isolation in *Drosophila*. *Science* **265**, 1461–1464 (1994).
53. Fang, S., Takahashi, A. & Wu, C.-I. A mutation in the promoter of *desaturase 2* is correlated with sexual isolation between *Drosophila* behavioral races. *Genetics* **162**, 781–784 (2003).
54. Wood, D. & Ringo, J. M. Male mating discrimination in *Drosophila melanogaster*, *D. simulans* and their hybrids. *Evolution* **34**, 320–329 (1980).
55. Watterson, G. A. On the number of segregating sites in genetical models without recombination. *Theoret. Pop. Biol.* **7**, 256–276 (1975).
56. Li, W. H. Distribution of nucleotide differences between two randomly chosen cistrons in a finite population. *Genetics* **85**, 331–337 (1977).
57. Gillespie, J. H. & Langley, C. H. Are evolutionary rates really variable? *J. Mol. Evol.* **13**, 27–34 (1979).
58. Hudson, R. R. Properties of a neutral allele model with intragenic recombination. *Theoret. Pop. Biol.* **23**, 183–201 (1983).
59. Takahata, N. Gene diversity in finite populations. *Genet. Res.* **46**, 107–113 (1985).
60. Wakeley, J. & Hey, J. Estimating ancestral population parameter. *Genetics* **145**, 847–855 (1997).
61. Wang, R. L., Wakeley, J. & Hey, J. Gene flow and natural selection in the origin of *Drosophila pseudoobscura* and close relatives. *Genetics* **147**, 1091–1106 (1997).
62. Kliman, R. M. *et al.* The population genetics of the origin and divergence of the *Drosophila simulans* complex species. *Genetics* **156**, 1913–1931 (2000).
63. Machado, C. A., Kliman, R. M., Markert, J. A. & Hey, J. Inferring the history of speciation from multilocus DNA sequence data: the case of *Drosophila pseudoobscura* and close relatives. *Mol. Biol. Evol.* **19**, 472–488 (2002).
64. Li, W. H. Distribution of nucleotide differences between two randomly chosen cistrons in a finite population. *Genetics* **85**, 331–337 (1977).
65. Takahata, N., Satta, Y. & Klein, J. Divergence time and population size in the lineage leading to modern humans. *Theoret. Pop. Biol.* **48**, 198–221 (1995).
66. Takahata, N. & Satta, Y. Evolution of the primate lineage leading to modern humans: phylogenetic and demographic inference from DNA sequences. *Proc. Natl Acad. Sci. USA* **94**, 4811–4815 (1997).
67. Wall, J. D. Estimating ancestral population sizes and divergence times. *Genetics* **163**, 395–404 (2003).
68. Navarro, A. & Barton, N. H. Chromosomal speciation and molecular divergence-accelerated evolution in rearranged chromosomes. *Science* **300**, 321–324. (2003).
69. Lu, J., Li, W. H. & Wu, C.-I. Comment on 'chromosomal speciation and molecular divergence-accelerated evolution in rearranged chromosomes'. *Science* **302**, 988 (2003).
70. Navarro, A., Marquès-Bonet, T. & Barton, N. H. Response to comment on 'chromosomal speciation and molecular divergence-accelerated evolution in rearranged chromosomes'. *Science* **302**, 988 (2003).
71. Rieseberg, L. H. *et al.* Major ecological transitions in wild sunflowers facilitated by hybridization. *Science* **301**, 1211–1216 (2003).
This study recreated new species in sunflowers by hybridization between two good species. The synthetic new species mimicked the naturally-occurring ancient hybridizations. Certain combinations of the parental genomes are favoured by natural selection. Hybridization speciation shows the 'modular' nature of the genome (see reference 5).
72. Lynch, M. & Conery, J. S. The evolutionary fate and consequences of duplicate genes. *Science* **290**, 1151–1155 (2000).
73. Mayr, E. Wu's genic view of speciation. *J. Evol. Biol.* **14**, 866–867 (2001).

Acknowledgements

We thank S. Shi and N. Osada for technical help and two anonymous referees for critical comments. Grants from the National Institutes of Health (USA) to C.-I Wu and from the National Science Council (ROC) to C.-T. Ting are gratefully acknowledged.

Competing interests statement

The authors declare that they have no competing financial interests.

Online links

DATABASES

The following terms in this article are linked online to:

FlyBase: http://www.flybase.org

desat2 | *Hmr* | *Nup96* | *OdsH* | *unc-4*

Access to this interactive links box is free online.