

# Hybridization as an invasion of the genome

James Mallet

Galton Laboratory, University College London, Wolfson House, 4 Stephenson Way, London, UK, NW1 2HE

**Hybridization between species is commonplace in plants, but is often seen as unnatural and unusual in animals. Here, I survey studies of natural interspecific hybridization in plants and a variety of animals. At least 25% of plant species and 10% of animal species, mostly the youngest species, are involved in hybridization and potential introgression with other species. Species in nature are often incompletely isolated for millions of years after their formation. Therefore, much evolution of eventual reproductive isolation can occur while nascent species are in gene-flow contact, in sympatry or parapatry, long after divergence begins. Although the relative importance of geographic isolation and gene flow in the origin of species is still unknown, many key processes involved in speciation, such as 'reinforcement' of post-mating isolation by the evolution of assortative mating, will have ample opportunity to occur in the presence of continuing gene flow. Today, DNA sequence data and other molecular methods are beginning to show that limited invasions of the genome are widespread, with potentially important consequences in evolutionary biology, speciation, biodiversity, and conservation.**

## Introduction

Formless pods drift through space. They land on Earth. The pods germinate, and the developing embryos take on features of the individual humans that they will eventually replace. A chilling scenario emerges: an invasion has occurred and humanity is being taken over by an alien species with an agenda very different to our own. The 1956 cult movie 'Invasion of the Body Snatchers' achieves suspense by playing on our fear of infiltration and genetic usurpation.

Also around this time, the so-called 'biological' species concept, having been promoted widely for about a decade [1,2], became the prevailing view of species among evolutionary biologists. This viewpoint argued that species, unlike races or genera, were 'real' and had special, species-level qualities: 'isolating mechanisms', 'cohesion', and 'coadapted gene complexes'; species acted as vessels for the 'storage and protection of genetic variation.' Even today, many branches of biology continue to see species as discrete and fundamental units, rather than as poorly differentiated way-stations in a continuous hierarchy of biodiversity.

Hybridization (see Glossary) and introgression between species is the converse of reproductive isolation and

challenges the 'reality' of biological species. In the course of the development of the biological species concept, a sort of repugnance against hybridization prevailed, akin to the fear on which 'Invasion of the Body Snatchers' plays. Supporters of the biological species concept viewed hybridization as a 'breakdown of isolating mechanisms' [2]. When hybridization occurred, it was explained via species range changes and environmental disturbance, mostly as a result of human habitat alteration. F1 hybrids are generally less viable and fertile (even given some hybrid vigour). Backcrossed genotypes, if produced, are often inferior, so that introgression was assumed to be rare; if it did occur it was thought to lead only to deleterious effects [2]. These almost eugenic views about species were particularly prevalent among zoologists because of Ernst Mayr's influence. (By contrast, many botanists thought that introgression was common and important in adaptive evolution.) The same views led directly to the notorious hybrid policy of the US Endangered Species Act of 1973, by which 'hybrids' were deemed unworthy of conservation, whereas unsullied 'pure species' were apportioned higher status [3]. But today, tastes in biodiversity are changing, and the biological species concept is under attack: in 1990, the hybrid policy was rescinded [4].

Hybridization has been known at least since the time of Linnaeus, and has been discussed frequently by evolutionists [2,5,6] following Darwin's lead in the chapter 'Hybridism,' where he demonstrated the lack of a clear boundary between varieties and species [7]. Here, I collate and review lesser-known comparative data on natural hybridization rates and discuss results from newer, molecular methods for the detection of hybridization and introgression, rather than covering the subject *in toto*. Much of the best literature on natural hybridization is about plants, but I concentrate particularly on animal data because zoologists have traditionally been more skeptical about the importance of introgression.

I also attempt lay to rest our almost instinctive, common-sense view that hybridization is always unnatural or extremely rare (see also [5]). In zoology, we tend automatically to assume that hybridization is a 'reproductive mistake'. For example, from an excellent discussion on hybridization in birds of paradise, we read: 'We presume that a male will mate with whatever bird solicits his copulation – another *weakness* in the system that might allow hybridization' ([8], my emphasis). It is indeed probable that mating systems have often evolved to prevent this 'weakness', but we can no longer take it for

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

## Glossary

**Allopatric:** two groups of populations or species are allopatric if their ranges are not connected; there is a geographic barrier between them that prevents gene flow. Allopatric speciation is therefore speciation in which the whole process takes place under conditions of geographic isolation.

**Assortative mating:** the tendency for 'like to mate with like'. Species or morphs within a species are said to mate assortatively if mating is biased towards other individuals with a similar phenotype or genotype.

**Biological species concept:** the species concept in which species are considered to be reproductively isolated populations [2,6]. Conversely, in the 'recognition concept' [13], individuals are members of the same species if they share common fertilization or recognition mechanisms. These two concepts are almost mirror images: in the biological species concept, species are defined by incompatibilities with individuals outside the species; in the recognition concept, species are defined by compatibility with other individuals inside the species.

**Dominance theory:** a theory to explain **Haldane's Rule**. Hybrids between species can suffer deleterious incompatibilities because alleles from two or more genes do not interact well together in hybrids. If these genes are on sex chromosomes and act recessively, the heterogametic sex (i.e. chromosomally XY, as in the male in mammals and *Drosophila*, and the female in birds and Lepidoptera) is liable to suffer most because it has only a single copy of the deleterious sex-linked allele, which is therefore expressed. In the homogametic sex (i.e. chromosomally XX, as in the female in mammals and *Drosophila*, and the male in birds or Lepidoptera), the compatible dominant allele from the other species masks the deleterious effect. Dominance theory not only explains why it is the heterogametic sex that usually suffers most in a wide variety of animals, but also why species with large sex chromosomes tend suffer more greatly Haldane's Rule than species with small chromosomes [31].

**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 originally by JBS Haldane in the 1920s, who showed that hybrid incompatibilities were usually greater in the heterogametic sex, regardless of whether the heterogametic sex was male or female. For explanation, see **dominance theory**.

**Horizontal gene transfer:** transfer of genes between species via some route other than sexual contact and hybridization (e.g. transduction by a virus, transformation by free DNA, or transfer by means of a plasmid).

**Hybridization:** crossing of genetically distinguishable groups or taxa, leading to the production of viable hybrids. In this article, the term is used mainly to refer to hybridization in the wild, as opposed to experimental crosses in the laboratory.

**Introgression:** invasion of foreign genetic material into a genome. The method of introduction is usually by sexual contact or hybridization; genetic transfer by other means is usually termed **horizontal gene transfer**.

**Parapatric:** two groups of populations or species are parapatric if they abut at their range boundaries. Gene flow is possible between parapatric populations or species, but is commonly limited to populations at the edge of each species range, and where the ranges of the two forms overlap. If hybridization occurs freely in the overlap, this results in a narrow band of hybridization, or hybrid zone.

**Phylogenetic species concept:** the species concept in which species are populations considered to differ by one or more fixed differences, or, in other formulations, in which species are considered to be the smallest phylogenetic group that can be defined by shared derived characters (i.e. monophyletic groups).

**Reinforcement:** if hybrids between a pair of sympatric species have low viability or sterility, these deleterious effects might 'reinforce' reproductive isolation by promoting the evolution of assortative mating, essentially to prevent gametic wastage.

**Sympatric:** populations are sympatric if individuals of each occur within 'cruising distance' of the other, so that gene flow is not only possible, but probable if there are no reproductive barriers between the forms. Sympatric speciation is speciation in which the whole process takes place while the diverging populations are in gene flow contact.

granted that hybridization is a 'weakness' *a priori*. For example, natural hybrids among Darwin's finches are fitter than purebreds, at least under current climatic conditions on Galapagos, very probably because the avoidance of deleterious homozygosity by outcrossing outweighs whatever benefits there might be of choosing mates from within the species [9]. Selection will regulate deleterious hybridization, but the selection pressure on the parents against hybridization can never be greater than the fraction of the population hybridizing; as soon as

hybridization becomes rare, selection might not be strong enough to deliver complete assortative mating. Greatly improved genetic data show clearly that horizontal gene transfer, hybridization, and introgression between species are ongoing and regular, if not always common processes in nature [5]. These discoveries have important implications in conservation [4] as well as in studies of the safety of transgenic crops that might hybridize with wild relatives [10]. Genomic invasions have been occurring at all levels of the Tree of Life since the dawn of evolution, and have contributed considerably to the adaptive radiation and diversification of early life ([11,12] and see below).

## The frequency of natural hybridization

Hybridization between species is always rare on a per-individual basis, but this statement is tautological because we would not be able to distinguish species if hybridization were common. Under biological [2,6] or recognition [13] species concepts, in which species are internally compatible and externally reproductively isolated, hybridization between individuals of separate species is rare by definition. Under phylogenetic or genealogical species concepts, species are defined by fixed, and/or uniquely derived genetic traits; such fixed differences similarly preclude an exchange of genes *a priori*. More pragmatic cluster definitions of species [6,14] also require that hybridization is rare, or it would swamp the process of natural selection against hybrids that maintain the differences between sympatric species.

It is not generally realized that in spite of rarity on a per individual basis, natural hybridization is usually common on a per species basis. Although hybrids are rare in populations, a few hybrids can provide a bridge to allow a trickle of alleles to pass between species; thus, if species that hybridize are common, even low rates of hybridization per individual can have important evolutionary consequences in a high fraction of species. In Table 1, and as discussed in more detail in the Online Supplementary Material Appendix 1, I collate hybridization rates on a per species basis for a variety of eukaryotes. Only hybridization in the wild is surveyed, and hybridization with non-native species is excluded.

It has often been said that vascular plant species are more likely to hybridize than animals (e.g. [2]). Comparing the studies on UK flora (25% of species hybridize) with world birds, European butterflies and European mammals overall (6–12% of species hybridize, maybe 10% as a rounded average; see Table 1 and Online Supplementary Material Appendix 1), this would certainly seem to be the case. However, animals display phylogenetic hotspots of natural hybridization that greatly exceed the 25% rate per species for vascular plants as a whole. For example, an astonishing three-quarters of British duck species, and all four British game birds are known to hybridize with at least one other species; meanwhile passion flower butterflies, birds of paradise, American warblers and tits all have as high a probability of hybridization, per species, as the average vascular plant.

It could be argued that the fraction of animal species hybridizing could be much lower than 10%, if the studies

**Table 1. Hybridization in the wild, including rates of hybridization per species<sup>a</sup>**

Taxon	No. species sampled	No. species hybridizing with at least one other	% species hybridizing with at least one other	Approx. rate of hybridization per individual	Notes and references
UK vascular plants (Pteridophyta, Angiospermae)	539 (sampled from 2290 species known in the UK)	135	25.0%	Variable, but mostly <1%	Stace [75,76] treated 2290 species. To estimate the rate of hybridization per species, I sampled one whole family every ten families; subsequent adjacent families were added to each sample, in Stace's sequence [76], if the sampled family/ies was/were small, until a total $\geq 20$ species was achieved per sample
Flora of Concord, Massachusetts (Angiospermae)	838	>27	>3.2%	Unknown	In Concord Township, 9.2% species were 'problematic', mainly for reasons other than hybridization. Mayr [15] included 'only a sample' of 11 cases of occasional hybridization. In most cases, hybrids 'fail to backcross, or do so very infrequently.' There was no 'obliteration of...integrity of...parent species' (p. 233). Only one species pair was considered problematic for reasons of current natural hybridization. See also [77,78]
European butterflies (Lepidoptera: Rhopalocera)	379	47	12.4%	Mostly <0.1%	Hybridization rates per individual are more frequent in some parapatric zones of overlap [20,79]
Swallowtail butterflies (Lepidoptera: Papilionidae; <i>Papilio</i> )	216	14–32	6.5–14.8%	Some hybrids common in restricted areas	Upper bounds include 'possible' hybrids; lower bounds are confirmed hybrids. Author suggests underreporting as a result of temperate zone bias. Per species hybridization was much higher in the better-studied N. American fauna (47.1%) [45]
Passion flower butterflies (Lepidoptera: Heliconiina)	73	19	26.0% (34.8% in genus <i>Heliconius</i> )	Mostly <0.1%	Hybrids between the parapatric species <i>Heliconius erato</i> × <i>Heliconius himera</i> , <i>Heliconius charithonia</i> × <i>Heliconius peruvianus</i> occur at a much higher rate of ~5% in narrow overlap zones [18]; see also <a href="http://www.ucl.ac.uk/taxome/hyb/">http://www.ucl.ac.uk/taxome/hyb/</a>
<i>Drosophila</i> of the world (Diptera: Drosophilidae)	~1750	18	1.0%	~0.01%	The per individual hybridization rate cited is for <i>Drosophila pseudoobscura</i> × <i>Drosophila persimilis</i> [36,37,80,81]
World birds (Aves)	9672	895	9.3%	<0.002%	19.5% of interspecific hybridizations are across genera [82]. Mayr [2] considered hybridization in birds unimportant because hybrids form <1 in 50 000 specimens
Native grouse of Britain (Tetraonidae)	4	4	100.0%	Mostly <0.1%	[83]
British ducks (Anatinae)	21	16	76.2%	Mostly <0.1%	Authors [84–86] include information on captive hybrids; only wild hybrids were counted here
Birds of paradise (Paradisaeidae)	42	18	42.9%	~0.03%	Most of the > 100 hybrid specimens involve 39 species of Paradisaeinae (46.2%); many intergeneric as well as interspecific hybrids [1,8]
American warblers (Parulidae)	116	28	24.1%	Mostly <0.1%	Usually monogamous, with both parents caring for young [87]
Tits of the world (Paridae)	70	20	28.6%	Mostly <0.1%	Nine of 20 hybridizing species hybridize outside their 'superspecies' [88]
Warblers of the western Palaearctic (Sylviidae)	62	0[20*]	0.0%[32.3%]	–	Hybrids are suspected, but not confirmed between Bonelli's and wood warblers ( <i>Phylloscopus bonelli</i> × <i>Phylloscopus sibilatrix</i> ) [89] [errata: *?unreliable – <a href="http://www.bird-hybrids.com">www.bird-hybrids.com</a> ]
European mammals	200	12	6.0%	Mostly <0.1%	[90–92]

<sup>a</sup>For a detailed explanation, see Online Supplementary Material Appendix 1.

reviewed here (Table 1) represent a biased sample, and hybridization has been recorded mainly in those species particularly prone to hybridize. Although some of the smaller groups (ducks, birds of paradise, heliconiine butterflies) might be argued to be in this category, the averages over larger groups (world birds, European butterflies, European mammals) are unbiased, and are simply those that have been collated because of good systematic knowledge. If anything, it seems probable that per species hybridization rates overall are higher than in Table 1 because many hybrids could have gone unrecognized in morphologically uniform subgroups such as European warblers, small mammals, and skipper butterflies (Hesperiidae) (Online Supplementary Material Appendix 1).

Another common belief, found especially in the work of Ernst Mayr (e.g. [2,15]), is that hybridization almost never results in gene flow and introgression. However, the first (F1) hybrid is often the most difficult to produce. The rarity of such hybrids in animals in the wild is largely a result of the very strong assortative mating shown by closely related species living in sympatry [16,17]. Once F1 hybrids are produced, and if they are viable and fertile, backcrossing to one of the parentals is much more straightforward: assuming additive inheritance of both mate choice and the phenotypic traits on which mate choice acts, the genetic gulf to be traversed will be halved. To my knowledge, few comparative data on this topic exist, but it is usually the case in butterflies that it is easier to cross F1 hybrids once they have been produced than to

obtain the F1 hybrid itself [18–20]. Provided that even some backcross hybrids are fertile, limited flow of genes between species is not only possible, but also probable. Many backcross hybrids are almost impossible to distinguish morphologically from parental species (e.g. [19]), so the frequency of backcrossing in the wild is almost certainly underestimated.

A similar story is found in birds: a typical young pair of sister species might be only 1–2 million years (MY) old [21], yet bird hybrids are fertile until the species are ~7–17 MY old and complete hybrid inviability only sets in around 11–55 MY [16]. In well-studied groups, about 10% of animal species hybridize with at least one related species (Table 1). However, many more must have gone through a phase of hybridization because species currently hybridizing are mostly the youngest. There is typically a long period after speciation when hybridization can lead to introgression. Most of the reproductive discontinuity between sister species is a result of assortative mating, and, during introgression, important speciation processes, such as reinforcement, can occur [17,22]. Whether introgression actually does occur, of course, will depend on the nature of selection keeping the species apart and can be detected via appropriate genetic analyses.

### Is hybridization natural, or are humans to blame?

The frequent examples of hybridization in nature (e.g. Table 1) are often attributed to environmental degradation [2,6]: if hybridization is assumed to be unnatural, its presence must indicate some failure of the ‘balance of nature’. Anderson [23], for example, argued that much hybridization and introgression was a result of ‘hybridization of the habitat’ of separately adapted species, mainly caused by human disturbance. (It is undoubtedly true that some examples of hybridization today, such as the European white-headed duck with the American ruddy duck or of introduced trout with wild relatives [4], are a result of human introductions and habitat disturbances. However, the data in Table 1 exclude all such introduced-species hybridization.)

The topic deserves more study, but here I use two examples in which environmental disturbance is unlikely to be the main cause of natural hybridization. It is well established that North American wolves might contain introgressed mitochondrial DNA (mtDNA) haplotypes from coyotes, and it is usually suggested that this is a result of deforestation and habitat clearance leading to a spread of the coyote into areas occupied by wolves followed by hybridization [24]. Maybe hybridization was indeed less prevalent in the past, but it is most unlikely that the coyote never hybridized with wolves at prairie–forest ecotones before the appearance of humans. Similar arguments can be made for *Heliconius* butterfly hybridization in the wild [18]: many of the known hybrid specimens were collected long before the advent of the chainsaw and modern neotropical forest fragmentation that brings species from open and closed canopy forest habitats into close contact. To suggest that these now-sympatric species never came into contact across savannah–forest boundaries or along river edges before human habitat disturbance occurred seems unparsimonious. In

conclusion, we should be much more suspicious than we have been of uncritical arguments that all or most hybridization surveyed in Table 1 is ‘unnatural’ and a result of environmental disturbance.

### Factors affecting the frequency of hybridization

#### ‘Isolating mechanisms’

The low frequency of hybrids between most species, on a per individual basis, is largely explained by prezygotic and postzygotic effects known as ‘isolating mechanisms.’ This term is associated with the biological species concept, and seems to imply that the ‘mechanisms’ have been designed by natural selection to prevent hybridization. Today, a more neutral term such as ‘reproductive barriers’ seems better because much reproductive isolation is ‘unintended’ by natural selection and is instead a by-product of other evolutionary changes [6]. For example, assortative mating as a result of partial spatial separation is often produced as an incidental effect of adaptation to a new environment [25–27]. Similarly, hybrid inviability or sterility will almost always arise as a pleiotropic result of some other evolutionary change [2,6,7]. However, when hybrids are less fit than the parental species, assortative mating can evolve adaptively (‘reinforcement’) [22], and, on the basis of comparative studies, this seems likely to have played an important role in speciation in *Drosophila* [17]. The wealth of behavioural, ecological, physiological, and genetic causes of reproductive isolation are well-treated elsewhere (e.g. [2,5,6]). Here, I ask how and why the strength of isolation might vary from group to group (Table 1).

#### Time since divergence

Incompatibilities between populations are strongly affected by selection, and so are not expected to evolve in a regular, clock-like manner. Nonetheless, both pre- and postmating compatibility in laboratory hybridization experiments are strongly negatively correlated with genetic distance, and therefore with time since separation of taxa [5,16,17,28]. Although not well studied, the frequency of natural hybrids almost certainly obeys the same law. For instance, in *Heliconius*, abundantly hybridizing geographic races within species almost all have mtDNA sequence differences of less than 2%; species that hybridize occasionally in the wild are mostly 2–6% divergent; whereas no hybrids are found between species >10% divergent [18]. These results suggest that reproductive isolation between pairs of overlapping species evolves gradually and progressively, albeit somewhat irregularly.

#### Taxon-specific differences

Some groups of organisms seem to hybridize much more readily than others, even accounting for time of divergence. For instance, vascular plants generally hybridize at around 2–3 times the per species rate found in animals (Table 1). It has been suggested that ‘the open, less integrative, and plastic patterns of plant morphogenesis’ are more permissive to large-effect genetic changes (which would include those resulting from hybridization and polyploidy) than those of animals [29]. However, many

animal groups have as high a frequency of hybridization per species as British vascular plants as a whole (Table 1). Amphibian and bird species remain compatible enough to produce viable laboratory hybrids for a longer evolutionary time than mammals (measured using a DNA–DNA hybridization-based molecular clock). It has been argued that faster rates of mammalian developmental change and regulatory gene evolution cause a more rapid accumulation of hybrid incompatibilities [30]. However, the comparison between mammals of Europe and birds of the world suggests that the fraction of species producing viable hybrids in nature is not as dissimilar (6% versus 9% of species, respectively; Table 1) as had been suggested by the limited laboratory compatibility data [30]. More recently, additional genetic explanations have been shown to explain variation in incompatibility between species. Incompatibilities affecting sterility of the heterogametic sex are more intense in species with larger X or Z chromosomes than those with smaller sex chromosomes, as expected from the dominance theory of Haldane's rule [31].

Developmental and genetic explanations for variability in hybridization rates ignore the importance of behavioural decisions to mate, which play the major role in preventing hybridization between sympatric animals (e.g. [16,17]). Mayr suggested that the birds most likely to hybridize belong to polygamous species in which the male plays no role in rearing nestlings [2]. Males of such species might be keen to mate with every potential female because they have few investment costs per mating. This explanation would apply to many ducks, gamebirds and birds of paradise (Table 1). But, this view preceded the discovery that extra-pair copulation is widespread: many apparently monogamous species are now known to exhibit cryptic cuckoldry and polygamy. In addition, high rates of hybridization in tits (Paridae) and American warblers (Parulidae), and the contrast between the high per species rate of hybridization in American warblers and the low rate in their European ecological equivalents the Eurasian warblers (Sylviidae) cannot be explained this way (Table 1); in all these groups, males are monogamous and help at the nest.

Another probable reason for the variability of per species hybridization rates among groups of birds is ascertainment bias. Hybrids are easier to identify, and are generally recorded more commonly, in brightly coloured species (e.g. ducks, gamebirds, birds of paradise, as well as parulid warblers, tits, and butterfly groups such as *Heliconius* and other Nymphalidae). Fewer hybrids are known in drab, uniform groups, such as sylviid warblers and numerous butterflies such as skippers included in the European butterflies as a whole (Table 1), probably because they cannot be recognized easily. Hybrids between morphologically uniform species will be cryptic: experts can distinguish specimens of willow warbler *Phylloscopus trochilus* and chiff-chaff *P. collybita* sibling species (Sylviidae) using morphological characters alone, but a hypothetical hybrid between these two very similar species would almost certainly be lumped with one or other parent species. By contrast, hybrids between more brightly and diversely coloured American warblers (Parulidae) or tits (Paridae) are much more easily detected.

In conclusion, the loss of a tendency to hybridize is reasonably, but relatively coarsely clock-like: sister taxa become increasingly incompatible, but at somewhat variable rates. This must be a result of a mixture of behavioural, genetic, and developmental peculiarities of taxa, as well as a result of the variation in rate of cladogenesis, and hence numbers of sympatric, closely related species, in each group. The probable ascertainment bias suggests that most rates of hybridization are underestimates, especially in poorly studied or uniform groups. However, it is now clear that most rapidly radiating groups undergo interspecific hybridization in the wild, often at high rates per species.

### Genetic evidence for hybridization and introgression

Hybridization and introgression are well known in plants. For example, genetic studies confirming hybridization and introgression exist in many taxa long known to produce hybrids, such as *Eucalyptus*, oaks, and willows [32–34], and good surveys are available of genetic studies on plants [5]. Hybridization in animals is more controversial, and, as a zoologist, I therefore concentrate on animal data.

### Identification of hybrids

Most data on hybridization in the wild (Table 1) come from analyses of museum specimens or sightings in natural populations. However, morphological identification of hybrids is difficult because ancestral polymorphism or mutations at a few colour pattern genes can give the appearance of hybrids. Molecular genetics enables more rigorous tests for hybridization or introgression. In general, such studies have confirmed and amplified the extent of hybridization inferred from morphology.

For example, in a recent study of a pair of spadefoot toads [35], F1 hybrids were identified by inspection of multilocus allozyme genotypes to demonstrate a decline of hybridization rates over the past few decades. The authors suggested that this was a result of rapid evolution of reinforcement. One problem with the use of allozymes to identify hybrids is that variants of the same alleles on a gel could be derived independently by different mutations leading to a change in electric charge of the protein. Chromosomal variants, by contrast, provide much stronger evidence for hybridization under the assumption that the same breakpoints rarely evolve twice. Proofs of existence of natural hybrids between related species has often been via cytology, such as in the flies *Drosophila*, *Simulium* and *Anopheles* [36–39].

If species are hybridizing across a narrow hybrid zone, geographic discordance between clines of genetic markers provides excellent evidence for hybridization and introgression. Often, it is suspected that foreign mtDNA haplotypes might introgress preferentially into one of the species (e.g. in the mouse *Mus*, wolves and the fish *Barbus* [24,40,41]). In other cases, nuclear loci are suspected of permeating across the hybrid zone (e.g. in *Chorthippus* grasshoppers and birds such as warblers and flycatchers [42–44]). Haldane's rule predicts that the heterogametic sex will be preferentially affected by hybrid incompatibilities, preventing mitochondrial transmission between species when the heterogametic sex is female.

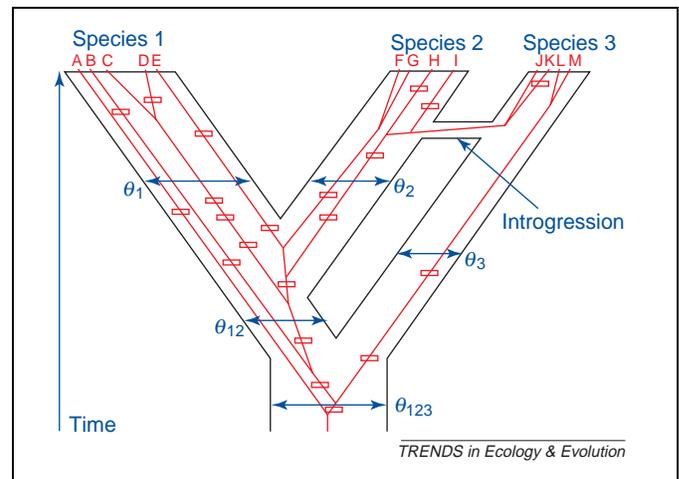
This probably explains why mitochondrial introgression is less common in birds and butterflies, where the heterogametic sex is the female, than in mammals and flies, where it is the male [45].

Often, genetic differences between closely related species are not fixed, and populations are sympatric. Modern statistical techniques to detect hybridization or introgression (Online Supplementary Material Appendix 2) become increasingly important in these cases. For example, suspected sympatric hybridizations between the butterflies *Papilio machaon* and *Papilio hospiton* [46,47] and between host races of the moth *Zeiraphera diniana* [48] were confirmed by analyzing multilocus haplotype data using newer Bayesian methods (e.g. [49,50]; see also Online Supplementary Material Appendix 2).

### Introgression

When recently diverged species hybridize in sympatry, an interesting genomic pattern can result. If relatively few loci are under divergent selection in the two species, these loci and nearby genomic regions are likely to remain distinct, whereas unlinked or distantly linked chromosomal regions should be able to flow relatively freely between species. We therefore expect chromosomal segments containing genes experiencing divergent selection to display strongly differentiated loci (many of them neutral); segments lacking divergent adaptations will remain similar as a result of introgression and frequent recombination. Probable examples of such genomic patterns have been found in periwinkles *Littorina* and larch budmoths *Zeiraphera* [48,51]. In a hybridizing pair of *Drosophila* species, divergent adaptations are associated particularly with paracentric inversions, which strongly inhibit recombination. It seems probable that the existence of these inversions enhances the persistence of each species in the presence of gene flow [52]. In such cases, only part of the genome is effectively reproductively isolated and determines the differences between the species: reproductive isolation is a term not easily applied to the whole genome.

Sympatric hybridization has often been studied using mitochondrial sequence data, although, on its own, mtDNA provides very little information about hybridization. However, mtDNA coupled with information from other loci or geographic data has often provided convincing evidence of introgression between sympatric species such as mule and white-tailed deer *Odocoileus*, minnows *Gila*, and mussels *Mytilus* [53–55]. Nuclear-genotype or DNA-sequence data can give more powerful evidence of introgression, although gene flow is hard to separate from ancestral polymorphism (Figure 1; Online Supplementary Material Appendix 2). A good example is in Darwin's finches *Geospiza*, where hybridization is relatively frequent, and little genetic differentiation was found between species at any of the marker loci studied [9,56]. Similar patterns are found in other birds such as the herring gull group *Larus* [57]. Among insects, the tephritid flies *Bactrocera* and *Heliconius* butterflies give similar multilocus sequence evidence for introgression [58,59], as do malaria-carrying species of the *Anopheles gambiae* group



**Figure 1.** The genealogy of a single gene in three species. The figure shows the true phylogeny of three closely related species in black outline. Time is shown flowing upwards with the present at the top. The lengths of the branches in the phylogeny represent the time since divergence. The true genealogy of the haplotypes sampled at a single gene is shown in red, with substitutions shown as small red oblong boxes. The widths of the branches of the phylogeny are proportional to the effective population size of each species. Assuming that all substitutions in the genealogy are neutral, the population size can be measured by  $\theta = 4N_e\mu$ , where  $N_e$  is the effective population size of the species and  $\mu$  is the mutation rate per base pair per unit time. Although this seems an odd way of measuring population size, it is convenient because  $\theta$  is the expected number of differences between any pair of haplotypes within a species per base pair. The current species population sizes ( $\theta_1$ ,  $\theta_2$  and  $\theta_3$ ) and the ancestral species population sizes ( $\theta_{12}$  and  $\theta_{123}$ ), the times since divergence, and the probability of introgression will all affect the probability of allelic sharing between species. Genealogical methods use multilocus haplotype data to attempt to estimate the true gene genealogies, times since divergence, ancestral population sizes, and the level or probability of introgression between species. This is an extremely challenging statistical task, and there has been only limited progress towards this objective (see Online Supplementary Material Appendix 2).

[60,61]. Perhaps the largest samples of haplotypes and loci have been studied in the *Drosophila pseudoobscura*–*persimilis* species pair, which convincingly show that some loci (such as *Adh* and the mitochondrial genome) flow relatively freely between species, whereas other loci show strong evidence of isolation [62–64].

In conclusion, in most cases where hybridization and introgression is suspected on morphological and behavioural grounds, genetic evidence confirms it. Introgression can be highly selective, affecting only some parts of the genome, whereas other genomic regions strongly affected by divergent selection remain virtually isolated. Introgression is certainly not restricted to plants: animals within many studied animal groups, whether vertebrate or invertebrate, hybridize and introgress. This includes particularly many birds and mammals in which introgression was thought previously to be relatively unimportant.

### Introgression, adaptation, speciation, and biodiversity

As already mentioned, hybridization between introduced and native species can cause problems in conservation [4,65] and has become an important topic in the debate about the release of transgenic crops [10]. In Europe, hybridization between native white-headed and American ruddy ducks *Oxyura* is a high-profile example, and introductions of trout for fisheries have been a particular problem for native trout species [4]. Invasion of the genome can be a problem.

But when framing conservation policy, we should remember that invasions of the genome are sometimes natural, and, although often deleterious, can ultimately contribute to adaptability and diversification. This is now so well known that I give only a few examples here (see also [5]). Many microbes thrive on gene exchange with distantly related taxa, and horizontal gene transfer and recombination can often provide adaptive traits. For example, baculoviruses are said to ‘sample their genomic environment... for beneficial genes’ [66], and bacteria often acquire resistance to antibiotics via interspecific transfer (e.g. in *Streptococcus pneumoniae* [67]). There is no reason why such benefits should not also accrue via introgressive hybridization among animals or plants. Useful crops and domestic animals often have hybrid origins, as do many pests, diseases and other organisms that cause problems for our species [68]. For example, gene flow between species has been blamed for transfer of insecticide resistance genes in malaria mosquitoes and blackflies [69,70]. In Darwin’s finches, high heritability and variability of adaptive traits within species can often be to the result of enhanced genetic variance obtained via introgression [56]. In *Papilio* and *Heliconius* butterflies, adaptive colour pattern diversity is strongly suspected to have been enhanced by interspecific gene flow [45,71]. It is undoubtedly true that interspecific hybridization is usually strongly deleterious, but these data suggest that introgression can occasionally allow adaptive combinations to evolve at a higher rate than in the absence of an input of variation from hybridization.

Allopolyploid speciation (i.e. via chromosome doubling of F1 hybrids) is well known in plants. Around 2–4% of cases of speciation in flowering plants are associated with polyploidization events, whereas 7% of fern species originate in this way [72]; many of these cases are thought to be allopolyploids that originate by hybridization. In plants, rare cases of homoploid hybrid speciation, in which chromosome numbers do not double, are also known [73]. However, homoploid hybrid speciation is more difficult to verify than polyploidy, and seems likely to have been underestimated. Indeed, it is difficult to distinguish homoploid hybrid speciation from hybridization and introgression. Polyploidy is rare in animals [72], which are mostly dioecious and cannot reproduce clonally; however, homoploid hybrid speciation has no such barriers for animals. It has even been suggested that large radiations might receive a boost when their ancestors obtain genetic variation from hybridization [74].

An extreme view of the evolutionary significance of reproductively isolated species arose in the middle of the last century: ‘Without speciation, there would be no diversification of the organic world, no adaptive radiation, and very little evolutionary progress’ ([2], p. 661). Whether reproductive isolation is really necessary for adaptive diversification, or is instead merely a by-product of divergent evolution, now seems a much harder question to answer. But, even if genetically isolated species play a role in diversification, today we know that evolutionary progress can continue while species undergo genomic invasions from other species. Not only that, variation introduced via introgression regularly contributes to adaptation and diversification throughout the Tree of Life.

## Acknowledgements

The article was greatly improved thanks to useful suggestions and criticisms from Michael Hochberg, Mohamed Noor, Sally Otto, Clive Stace, Ziheng Yang, and an anonymous reviewer.

## Supplementary data

Supplementary data associated with this article can be found at [doi:10.1016/j.tree.2005.02.010](https://doi.org/10.1016/j.tree.2005.02.010)

## References

- 1 Mayr, E. (1942) *Systematics and Origin of Species*, Columbia University Press
- 2 Mayr, E. (1963) *Animal Species and Evolution*, Harvard University Press
- 3 O’Brien, S.J. and Mayr, E. (1991) Bureaucratic mischief: recognizing endangered species and subspecies. *Science* 251, 1187–1188
- 4 Allendorf, F.W. *et al.* (2001) The problems with hybrids: setting conservation guidelines. *Trends Ecol. Evol.* 16, 613–622
- 5 Arnold, M.L. (1997) *Natural Hybridization and Evolution*, Oxford University Press
- 6 Coyne, J.A. and Orr, H.A. (2004) *Speciation*, Sinauer Associates
- 7 Darwin, C. (1859) *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*, John Murray
- 8 Frith, C.B. and Beehler, B.M. (1998) *The Birds of Paradise*, Oxford University Press
- 9 Grant, P.R. *et al.* (2003) Inbreeding and interbreeding in Darwin’s finches. *Evolution* 57, 2911–2916
- 10 Hails, R.S. and Morley, K. (2005) Genes invading new populations: a risk assessment perspective. *Trends Ecol. Evol.* doi: 10.1016/j.tree.2005.02.006
- 11 Doolittle, W.F. (1999) Phylogenetic classification and the universal tree. *Science* 284, 2124–2129
- 12 Rivera, M.C. and Lake, J.A. (2004) The ring of life provides evidence for a genome fusion origin of eukaryotes. *Nature* 431, 152–155
- 13 Paterson, H.E.H. (1985) The recognition concept of species. In *Species and Speciation* (Vrba, E.S., ed.), pp. 21–29, Transvaal Museum
- 14 Mallet, J. (1995) A species definition for the Modern Synthesis. *Trends Ecol. Evol.* 10, 294–299
- 15 Mayr, E. (1992) A local flora and the biological species concept. *Am. J. Bot.* 79, 222–238
- 16 Price, T.D. and Bouvier, M.M. (2002) The evolution of F1 postzygotic incompatibilities in birds. *Evolution* 56, 2083–2089
- 17 Coyne, J.A. and Orr, H.A. (1997) Patterns of speciation in *Drosophila* revisited. *Evolution* 51, 295–303
- 18 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
- 19 Naisbit, R.E. *et al.* (2003) Mimicry: developmental genes that contribute to speciation. *Evol. Dev.* 5, 269–280
- 20 Descimon, H. and Mallet, J. Bad species. In *Ecology and Evolution of European Butterflies* (Shreeve, T.G. and Dennis, R.H.L., eds), Oxford University Press (in press)
- 21 Weir, J.T. and Schluter, D. (2004) Ice sheets promote speciation in boreal birds. *Proc. R. Soc. London B Biol. Sci.* 271, 1881–1887
- 22 Servedio, M.R. and Noor, M.A.F. (2003) The role of reinforcement in speciation: theory and data. *Annu. Rev. Ecol. Syst.* 34, 339–364
- 23 Anderson, E. (1948) Hybridization of the habitat. *Evolution* 2, 1–9
- 24 Lehman, N. *et al.* (1991) Introgression of coyote mitochondrial DNA into sympatric North American gray wolf populations. *Evolution* 45, 104–119
- 25 Emelianov, I. *et al.* (2001) Host-induced assortative mating in host races of the larch budmoth. *Evolution* 55, 2002–2010
- 26 Berlocher, S.H. and Feder, J.L. (2002) Sympatric speciation in phytophagous insects: moving beyond controversy? *Annu. Rev. Entomol.* 47, 773–815
- 27 Drès, M. and Mallet, J. (2002) Host races in plant-feeding insects and their importance in sympatric speciation. *Philos. Trans. R. Soc. London Ser. B* 357, 471–492
- 28 Edmands, S. (2002) Does parental divergence predict reproductive compatibility? *Trends Ecol. Evol.* 17, 520–527

- 29 Gottlieb, L.D. (1984) Genetics and morphological evolution in plants. *Am. Nat.* 123, 681–709
- 30 Prager, E.M. and Wilson, A.C. (1975) Slow evolutionary loss of the potential for interspecific hybridization in birds: a manifestation of slow regulatory evolution. *Proc. Natl. Acad. Sci. U. S. A.* 72, 200–204
- 31 Turelli, M. and Begun, D.J. (1997) Haldane's rule and X-chromosome size in *Drosophila*. *Genetics* 147, 1799–1815
- 32 Griffin, A.R. *et al.* (1988) Patterns of natural and manipulated hybridization in the genus *Eucalyptus* L'Herit. – a review. *Austral. J. Bot.* 36, 41–66
- 33 Dumolin-Lapègue, S. *et al.* (1999) Are chloroplast and mitochondrial DNA variation species-independent in oaks? *Evolution* 53, 1406–1413
- 34 Hardig, T.M. *et al.* (2000) Morphological and molecular evidence for hybridization and introgression in a willow (*Salix*) hybrid zone. *Mol. Ecol.* 9, 9–24
- 35 Pfennig, K.S. (2003) A test of alternative hypotheses for the evolution of reproductive isolation between spadefoot toads: support for the reinforcement hypothesis. *Evolution* 57, 2842–2851
- 36 Noor, M.A.F. *et al.* (2000) Gene flow between *Drosophila pseudoobscura* and *D. persimilis*. *Evolution* 54, 2174–2175
- 37 Carson, H.L. *et al.* (1975) *Drosophila* hybrids in nature: proof of gene exchange between sympatric species. *Science* 189, 806–807
- 38 Post, R.J. (1984) Natural interspecific hybridisation of *Simulium sanctipauli* s.l. with *Simulium squamosum* and *Simulium yahense* (Diptera: Simuliidae). *Tropenmed. Parasitol.* 35, 58–60
- 39 della Torre, A. *et al.* (1997) Selective introgression of paracentric inversions between two sibling species of the *Anopheles gambiae* complex. *Genetics* 146, 239–244
- 40 Ferris, S.D. *et al.* (1983) Flow of mitochondrial DNA across a species boundary. *Proc. Natl. Acad. Sci. U. S. A.* 80, 2290–2294
- 41 Crespin, L. *et al.* (1999) Asymmetrical introgression in a freshwater fish hybrid zone as revealed by a morphological index of hybridization. *Biol. J. Linn. Soc.* 67, 57–72
- 42 Ferris, C. *et al.* (1993) One-way introgression of a subspecific sex-chromosome marker in a hybrid zone. *Heredity* 71, 119–129
- 43 Rohwer, S. *et al.* (2001) Plumage and mtDNA haplotype variation across a moving hybrid zone. *Evolution* 55, 405–422
- 44 Saetre, G.P. *et al.* (2001) Speciation, introgressive hybridization and nonlinear rate of molecular evolution in flycatchers. *Mol. Ecol.* 10, 737–749
- 45 Sperling, F.A.H. (1990) Natural hybrids of *Papilio* (Insecta: Lepidoptera): poor taxonomy or interesting evolutionary problem? *Can. J. Zool.* 68, 1790–1799
- 46 Aubert, J. *et al.* (1996) Population biology and conservation of the Corsican swallowtail butterfly *Papilio hospiton* Gén. *Biol. Conserv.* 78, 247–255
- 47 Cianchi, R. *et al.* (2003) Differential patterns of hybridization and introgression between the swallowtails *Papilio machaon* and *P. hospiton* from Sardinia and Corsica islands (Lepidoptera, Papilionidae). *Mol. Ecol.* 12, 1461–1471
- 48 Emelianov, I. *et al.* (2004) Genomic evidence for divergence with gene flow in host races of the larch budmoth. *Proc. R. Soc. London B Biol. Sci.* 271, 97–105
- 49 Anderson, E.C. and Thompson, E.A. (2002) A model-based method for identifying species hybrids using multilocus genetic data. *Genetics* 160, 1217–1229
- 50 Pritchard, J.K. *et al.* (2000) Inference of population structure using multilocus genotype data. *Genetics* 155, 945–959
- 51 Wilding, C.S. *et al.* (2001) Differential gene exchange between parapatric morphs of *Littorina saxatilis* detected using AFLP markers. *J. Evol. Biol.* 14, 611–619
- 52 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
- 53 Carr, S.M. *et al.* (1986) Mitochondrial DNA analysis of hybridization between sympatric white-tailed deer and mule deer in west Texas. *Proc. Natl. Acad. Sci. U. S. A.* 83, 9576–9580
- 54 Dowling, T.E. and DeMarais, B.D. (1993) Evolutionary significance of introgressive hybridization in cyprinid fishes. *Nature* 362, 444–446
- 55 Gardner, J.P.A. (1996) The *Mytilus edulis* species complex in south-west England: effects of hybridization and introgression upon interlocus associations and morphometric variation. *Mar. Biol.* 125, 385–399
- 56 Freeland, J.R. and Boag, P.T. (1999) The mitochondrial and nuclear genetic homogeneity of the phenotypically diverse Darwin's ground finches. *Evolution* 53, 1553–1563
- 57 Crochet, P.-A. *et al.* (2003) Genetic differentiation at nuclear and mitochondrial loci among large white-headed gulls: sex-biased interspecific gene flow? *Evolution* 57, 2865–2878
- 58 Morrow, J. *et al.* (2000) Close genetic similarity between two sympatric species of tephritid fruit fly reproductively isolated by mating time. *Evolution* 54, 899–910
- 59 Beltrán, M.S. *et al.* (2002) Phylogenetic discordance at the species boundary: comparative gene genealogies among rapidly radiating *Heliconius* butterflies. *Mol. Biol. Evol.* 19, 2176–2190
- 60 Gentile, G. *et al.* (2002) Genetic differentiation in the African malaria vector, *Anopheles gambiae* s.s., and the problem of taxonomic status. *Genetics* 161, 1561–1578
- 61 Besansky, N.J. *et al.* (2003) Semipermeable species boundaries between *Anopheles gambiae* and *Anopheles arabiensis*: evidence from multilocus DNA sequence variation. *Proc. Natl. Acad. Sci. U. S. A.* 100, 10818–10823
- 62 Wang, R.L. *et al.* (1997) Gene flow and natural selection in the origin of *Drosophila pseudoobscura* and close relatives. *Genetics* 147, 1091–1106
- 63 Machado, C.A. *et al.* (2002) Inferring the history of speciation from multilocus DNA sequence data: the case of *Drosophila pseudoobscura* and its close relatives. *Mol. Biol. Evol.* 19, 472–488
- 64 Hey, J. and Nielsen, R. (2004) Multilocus methods for estimating population sizes, migration rates and divergence times, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics* 167, 747–760
- 65 Rhymer, J.M. and Simberloff, D. (1996) Extinction by hybridization and introgression. *Annu. Rev. Ecol. Syst.* 27, 83–109
- 66 Herniou, E.A. *et al.* (2001) Use of whole genome sequence to infer baculovirus phylogeny. *J. Virol.* 75, 8117–8126
- 67 Dowson, C.G. *et al.* (1989) Horizontal transfer of penicillin-binding genes in penicillin-resistant clinical isolates of *Streptococcus pneumoniae*. *Proc. Natl. Acad. Sci. U. S. A.* 86, 8842–8846
- 68 Arnold, M.L. (2004) Natural hybridization and the evolution of domesticated, pest and disease organisms. *Mol. Ecol.* 13, 997–1007
- 69 Weill, M. *et al.* (2000) The kdr mutation occurs in the Mopti form of *Anopheles gambiae* s.s. through introgression. *Insect Mol. Biol.* 9, 451–455
- 70 Boakye, D.A. and Meredith, S.E.O. (1993) Introgression between members of the *Simulium damnosum* complex – larvicidal implications. *Med. Vet. Entomol.* 7, 393–397
- 71 Gilbert, L.E. (2003) Adaptive novelty through introgression in *Heliconius* wing patterns: evidence for shared genetic “tool box” from synthetic hybrid zones and a theory of diversification. In *Ecology and Evolution Taking Flight: Butterflies as Model Systems* (Boggs, C.L. *et al.*, eds), pp. 281–318, University of Chicago Press
- 72 Otto, S.P. and Whitton, J. (2000) Polyploid incidence and evolution. *Annu. Rev. Genet.* 34, 401–437
- 73 Rieseberg, L.H. (1997) Hybrid origins of plant species. *Annu. Rev. Ecol. Syst.* 28, 359–389
- 74 Seehausen, O. (2003) Hybridization and adaptive radiation. *Trends Ecol. Evol.* 19, 198–207
- 75 Stace, C.A. (1975) *Hybridization and the Flora of the British Isles*, Academic Press
- 76 Stace, C.A. (1997) *The New Flora of the British Isles*, (2nd edn), Cambridge University Press
- 77 Whittemore, A.T. (1993) Species concepts – reply. *Taxon* 42, 573–583
- 78 Mitchell, R. *et al.* (1993) Species concepts. *Nature* 364, 20
- 79 Guillaumin, M. and Descimon, H. (1976) La notion d'espèce chez les lépidoptères. In *Les Problèmes de l'Espèce dans le Règne Animal* (Volume 1) (Bocquet, C. *et al.*, eds), pp. 129–201, Société zoologique de France
- 80 Gupta, J.P. *et al.* (1980) Natural hybridization in *Drosophila*. *Experientia* 36, 290
- 81 Lachaise, D. *et al.* (2000) Evolutionary novelties on islands: *Drosophila santomea*, a new melanogaster sibling species from São Tomé. *Proc. R. Soc. London B Biol. Sci.* 267, 1487–1495
- 82 Grant, P.R. and Grant, B.R. (1992) Hybridization of bird species. *Science* 256, 193–197

- 83 Millais, J.G. (1892) *Game Birds and Shooting – Sketches Illustrating the Habits, Modes of Capture, Stages of Plumage and the Hybrids and Varieties which Occur amongst them*, Henry Southeran & Co
- 84 Millais, J.G. (1902) *The Natural History of the British Surface-Feeding Ducks*, Longmans, Green & Co
- 85 Millais, J.G. (1913) *British Diving Ducks*, Longmans, Green & Co
- 86 Gillham, E. and Gillham, B.L. (1996) *Hybrid Ducks. A Contribution towards an Inventory*, Published by the authors
- 87 Curson, J. *et al.* (1994) *New World Warblers*, Christopher Helm
- 88 Harrap, S. and Quinn, D. (1996) *Tits, Nuthatches and Treecreepers*, Christopher Helm
- 89 Parmenter, T. and Byers, C. (1991) *A Guide to the Warblers of the Western Palaearctic*, Bruce Coleman
- 90 Gray, A.P. (1972) *Mammalian Hybrids. A Check-List with Bibliography* (2nd edn), Commonwealth Agricultural Bureau
- 91 MacDonald, D. and Barrett, P. (1995) *European Mammals. Evolution and Behaviour*, Harper Collins
- 92 Arnason, U. *et al.* (1991) Molecular identification of hybrids between the two largest whale species, the blue whale (*Balaenoptera musculus*) and the fin whale (*B. physalus*). *Hereditas* 115, 183–189