Gene flow and species delimitation

Rémy J. Petit[1,2] and Laurent Excoffier[3]

[1] INRA, UMR 1202 BIOGECO, 33610 Cestas, France
[2] Université de Bordeaux, UMR 1202 BIOGECO, 33610 Cestas, France
[3] Computational and Molecular Population Genetics Lab, Institute of Ecology and Evolution, University of Bern, 3012 Bern, Switzerland

A defining feature of species is that their constituting populations are connected by gene flow. However, interspecific gene flow (introgression) can affect species integrity. If some genome components were less prone to introgression than others, they should be particularly suitable to delimitate species. Recent simulation studies have predicted a negative correlation between intra- and interspecific gene flow, suggesting that markers associated with the most dispersing sex should better delimitate species. A review of studies of introgression in species with sex-biased dispersal largely confirms this prediction. Hence, species delimitation should be more effective with markers experiencing high levels of gene flow, a simple but not widely appreciated prediction.

Introduction

Species delimitation, the methodological problem of identifying the boundaries between a set of species and therefore inferring the number of extant species, is an old issue that has recently attracted renewed attention [1,2]. Species are a cornerstone of biology, ecology and conservation, so their correct delimitation is essential, as illustrated by the surge of interest in barcoding approaches [3–6]. When selecting genetic markers for delimiting species, researchers have mostly focussed on their variability (e.g. [7]). A neglected criterion is the level of gene flow experienced by the markers. This is surprising because gene flow is a defining feature of species (see Box 1), and its role in species cohesion has been recently reiterated [8,9]. Moreover, due to the prevalence of sex-biased dispersal [10], different DNA regions are subject to varying levels of gene flow as a consequence of their mode of inheritance (biparental, maternal or paternal). It seems therefore logical to consider the effect of gene flow and not only mutation rate or effective population size when selecting markers for species identification and delimitation. In this paper, we suggest and attempt to test the idea that genome components experiencing high rates of gene flow should be better suited for species delimitation than those experiencing low rates of gene flow, in part because high rates of intraspecific gene flow can prevent introgression.

High gene flow markers to delimitate species?

The challenge in molecular taxonomy is to distinguish species that have low levels of genetic divergence, either because speciation is recent or because the species continue to exchange genes. Although it is legitimate to try to

detect introgressed individuals using appropriate markers (see Ref. [11] for a comparison of existing methods to detect hybrids), identifying reproductive entities using genome components that introgress less readily than the rest of the genome (and hence might not be appropriate to measure admixture) is equally relevant. At first view, if there were a positive correlation between intraspecific and interspecific gene flow, the use of markers experiencing low rates of intraspecific gene flow would seem appropriate for species delimitation. However, recent demo-genetic studies investigating the joint dynamics of genetic and demographic

Glossary

Allopatric: refers to the geographic separation of different populations.
Assignment method: any of several related statistical methods using genetic information to ascertain population membership of individuals.
Barcoding: the use of short, standardised DNA sequences to quickly and easily identify and discover species.
Cellular automata: discrete models consisting of a regular grid of cells, each in one of a finite number of states; the state of a cell is a function of the states of neighbouring cells in the previous generation, and every cell has the same updating rule.
Diagnostic: in taxonomy, a marker or trait is said to be diagnostic if it presents different character states in different taxa.
Effective population size: the number of breeding individuals in an idealised population that would show the same amount of genetic drift or the same amount of inbreeding as the population under consideration.
Essentialism: in philosophy, the view that, for any specific kind of entity, there is a set of permanent and unalterable characteristics or properties that any entity of that kind must possess.
Gene circulation: movement of genes among populations due to the combined effects of dispersal and selective processes.
Gene flow: movement of genes among populations due to dispersal processes.
Hitchhiking: the process by which a neutral allele increases in frequency because it is linked to a beneficial allele under positive selection.
Introgression: the permanent incorporation of genes from one set of differentiated populations into another.
Neutral model: a model in which genetic drift predominates and where selective effects can be ignored, generally because selection coefficients are small relative to population size.
Parapatric: refers to populations living in nearby geographic areas.
Philopatry: the tendency for an animal to breed in the vicinity of its birthplace.
Polyphyly: when a group does not include the most recent common ancestor of all its members.
Positive selection: darwinian selection resulting in the increased frequency of a beneficial mutation (i.e. a mutation characterised by a positive selection coefficient).
Selective sweep: rapid increase in frequency of a genomic region under positive selection, resulting in its spread across the species range.
Sex-biased dispersal: refers to an imbalance in the dispersal behaviour of the males and females of a species, leading to a markedly higher dispersal of one of the two sexes (i.e. male-biased or female-biased dispersal).
Speciation: evolution of reproductive barriers as well as phenotypic, behavioural and genetic differences between populations, eventually leading to distinct species.
Surfing of alleles: refers to the spread of an allele by the wave of advance of a range expansion.
Sympatric: refers to populations living in the same geographic areas.

Corresponding author: Petit, R.J. (petit@pierroton.inra.fr).

Box 1. Gene flow and the biological species concept

According to Mayr's biological species concept, species are defined as 'groups of actually or potentially interbreeding natural populations, which are reproductively isolated from other such groups' ([81], p. 120). Gene flow is not explicitly mentioned in this definition but it is nevertheless of central importance, as made clear later by Mayr [82], pp. 521–522): '[t]he steady and high genetic input caused by gene flow is the main factor responsible for genetic cohesion among the populations of a species.' Since its formulation, the biological species concept has had to face two opposite criticisms: that there is too little intraspecific gene flow to hold all populations together [83,84] and too much interspecific gene flow to keep each species apart [40]. In addition, Beurton [42] argued that gene flow is not only insufficient to keep species together but also conceptually impotent in bringing cohesion. However, he suggested that a combination of positive selection and gene flow, which he named 'gene circulation,' could result in cohesion through time and space as a consequence of new positively selected alleles successively spreading across the species range. This is supported by simulations showing that even when gene flow is minimal across species ranges, collective evolution is possible for traits under positive selection [8]. Another recent simulation study based on cellular automata further shows that species can emerge spontaneously in space and persist side by side in a homogeneous environment. This self-organising process is caused by a form of disruptive selection: On one hand, outbreeding depression (the decreased fitness of the progeny of a cross between divergent parents) results in a selective pressure to establish barriers to gene flow. On the other hand, there is selection for cross-compatibility within species, to avoid the deleterious consequences of outbreeding depression ([85]; see Figure I). Species individuality is therefore achieved dynamically, thereby supporting the verbal model of Beurton [42]. This view of species can be seen as an answer to the critiques that the biological species concept results in a new kind of essentialism [86] and that gene flow cannot account for species cohesion [87].
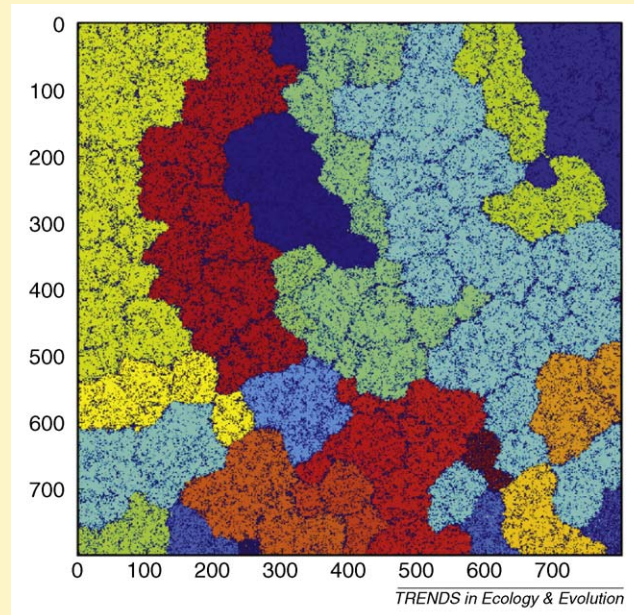

Figure I. Parapatric speciation in a homogeneous environment caused by gene flow limitation. This figure presents the genetic structure obtained following a simulation based on cellular automata, starting with uniform conditions, as described in Ref. [85]. The existence of genetically homogeneous geographic entities (corresponding to the different colours) within which individuals are interfertile indicates that cohesion is actively preserved within the newly emerged species. Individuals are located on an 800 × 800 cell grid, with 20% empty spaces located at random. They mate with each other to produce the next generation. All genomes start off identical, and genetic diversity arises through mutation. Dispersal is restricted, resulting in isolation by distance; moreover, there is a cost to mating between divergent genomes (outbreeding depression). False colours are used to depict genetic clustering on the simulated grid. The scattered cells in dark blue represent unoccupied cells or cells from which genomes were sampled that were not connected to any cluster. Each other colour represents a set of individuals with genome sequences that are identical at more than 40% of their nucleotide sites. Using this threshold helps identify different species with different colours. Courtesy of Guy Hoelzer and Rich Drewes.

processes have shown exactly the contrary: the rate of introgression should be most often *negatively* correlated with the rate of intraspecific gene flow [12]. The rationale is the following. Most species ranges are dynamic. If a species expands its range and meets a closely related species with which reproductive barriers are still incomplete, asymmetric introgression will take place from the local species to the colonising species [13]. However, if there is sufficient intraspecific gene flow among populations of the colonising species, genetic drift will be reduced and introgressed alleles will be less likely to increase in frequency by chance in the colonising species ([12]; see Box 2). Hence, the more intraspecific gene flow there is, the less interspecific gene flow is expected. This rather counterintuitive finding supports the idea that markers experiencing high rates of intraspecific gene flow should be preferred for species delimitation.

Testing the prediction

To evaluate the relevance of this prediction, we compiled data on rates of introgression for at least two types of markers having different modes of inheritance (i.e. mitochondrial DNA [mtDNA], sex chromosomes or autosomes). We focussed on wild species for which sex-biased dispersal

had been established or could be inferred. Three animal classes were considered, mammals (where many species have male-biased dispersal [10,14]), birds (where many species have female-biased dispersal [10,15]) and insects (where both cases can be found). Only those cases where the different markers provided discordant results were retained (i.e. when some markers were more introgressed than others). We considered examples involving named species but also subspecies or even chromosomal races, as long as some partial sterility barrier was shown to exist. The results of our literature search, which is not meant to be exhaustive, are summarised in Table 1. We report 16 'female-biased' and 21 'male-biased' examples, involving studies of 11 bird, 14 mammal and 12 insect species. We predicted that markers associated with the least-dispersing sex should be more introgressed and therefore less diagnostic for species, subspecies or race identification.

The results strongly support our expectation (Table 1; Box 3). In all 16 studies of species with female-biased dispersal, involving most birds, some insects and one mammal, nuclear markers are more introgressed than mtDNA markers and are thus less diagnostic. By contrast, in 17 out of 19 studies of species characterised by male-biased dispersal (mostly studies of mammals but also some

Box 2. Role of intraspecific gene flow in modulating introgression

Factors that affect the frequency of hybridisation and subsequent introgression include disturbance (reviewed in Ref. [88]) and colonisation (e.g. [89]). Both factors seem to result in the disruption of normal patterns of matings, perhaps owing to the rarity of conspecific mates [90]. Colonisation in particular can result in massive introgression of genes of the local species into the genome of the colonising species [12]. However, no prediction has been formulated so far regarding which genes are more likely to introgress when two closely related species meet. It is important to realise that two species that come into contact might greatly differ demographically, for instance when an immigrating species, which experiences rapid demographic growth, meets with an already established species at demographic equilibrium. During the expansion of a species, genetic drift occurs at the wavefront, where populations are still at low density. Alleles of the colonising populations can 'surf' on the wave of advance, reaching high frequencies and occupying large areas, even if they have no selective advantage [91]. Because any allele can surf, there is a competition among all alleles present at the edge of the wavefront for surfing. If the colonising species can interbreed with a local species, alleles originating from the local species will also be able to surf and reach high frequencies in the colonising species. However, if gene flow is high among populations of the colonising species, introgressed genes will compete for surfing with genes migrating from the interior of the wavefront, making it less likely that introgressed genes will surf and reach high frequencies. As a consequence, introgression will remain limited, suggesting that high rates of intraspecific gene flow can efficiently mitigate introgression (Figure I).
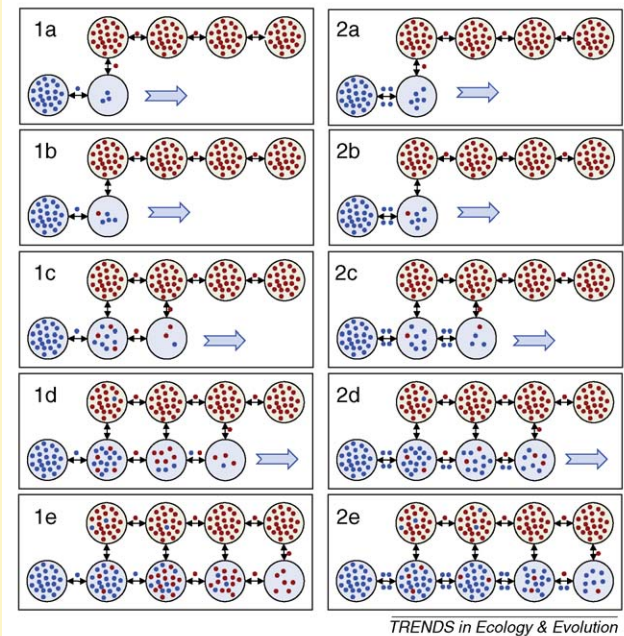

Figure I. Schematic representation of the role of intraspecific gene flow in mitigating introgression between a colonising and a local species. In this scenario, the blue species progressively colonises the territory of a local species (in red). The course of the colonisation is displayed from top to bottom. Interbreeding (vertical arrows) can occur locally between species, and populations from the same species can exchange migrants (horizontal double arrows). The left column (1a–1e) corresponds to a case with low gene flow resulting (1e) in massive introgression of local red alleles in the lastly colonised regions due to recurrent introgression events and the surfing of introgressed alleles. The right column (2a–2e) corresponds to the high gene flow case, where introgression of red alleles is limited by intraspecific gene flow among populations of the colonising species (persistence of the colonising blue alleles through time).

of insects and one of birds), nuclear markers are less introgressed and hence more diagnostic than mtDNA markers. This agreement with our prediction is remarkable given that mtDNA and nuclear DNA genomes differ not only in the rate of gene flow they experience (correlated to the extent of sex-biased dispersal) but also in their effective population size [16]. If we consider only uniparentally inherited markers with contrasted modes of transmission, where the situation is symmetric in this respect, the trend remains the same: markers on the paternally inherited Y chromosome are more differentiated between species than maternally inherited mtDNA markers in eight out of nine species with male-biased dispersal (all mammals). Unfortunately, we did not find cases of species with female-biased dispersal where maternally inherited and paternally inherited markers had been studied to test the opposite prediction.

In plants, seed dispersal is often considerably less effective than pollen dispersal [17]. Therefore, maternally inherited organelle markers (chloroplast DNA [cpDNA] or mtDNA) that are only dispersed by seeds should be more frequently introgressed and hence of more limited taxonomic value than nuclear markers. This is actually a well-established result [18]. In fact, when seed dispersal is particularly limited, patterns of cpDNA variation are often more influenced by geography than by taxonomy (e.g. [19–21]). Conifers are of special interest because their chloroplast and mitochondrial genomes have opposite

modes of inheritance: mtDNA is typically maternally inherited and cpDNA paternally inherited [22]. This situation allows for a direct comparison between two organelle genomes experiencing contrasted rates of gene flow. In keeping with our prediction that introgression is negatively correlated with intraspecific gene flow, species limits in conifers are more congruent with patterns of cpDNA variation than with patterns of mtDNA variation in all cases reported to date [23].

Other interpretations of marker-specific rates of introgression

Haldane's rule [24], which states that 'when in the offspring of two different animal races one sex is absent, rare, or sterile, that sex is the heterogametic sex,' is the most frequent explanation for the heterogeneity of introgression rates at markers with contrasted modes of inheritance, because it implies that markers transmitted only by the heterogametic sex will be less introgressed. Our survey included 17 cases where the female is the heterogametic sex (ZW, birds and Lepidoptera) and 20 cases where the male is heterogametic (XY or XO) or haploid. In 7 out of these 37 cases (1 bird, 1 mammal and 5 insect species pairs), the heterogametic sex was not the sex dispersing most. In all 7 cases, interspecific differentiation was better predicted by dispersal than by chromosomes. Moreover, among the 30 remaining cases, there were some examples

Table 1. Relative rates of introgression at different markers in species with sex-biased dispersal

Species[a]		Family	Heterogametic sex	Introgression rate[b,c]	Refs
With female-biased dispersal					
Imperial eagles	*Aquila adalberti-A. heliaca*	Accipitridae	Female	nucl>mt	[47]
Spotted eagles	*Aquila clanga-A. pomarina*	Accipitridae	Female	nucl>mt	[48]
Flycatchers	*Ficedula hypoleuca-F.albicollis*	Muscicapidae	Female	nucl>mt	[49]
Hippolais warblers	*Hippolais icterina-H. polyglotta*	Sylviidae	Female	nucl>mt	[50]
Large white-headed gulls	*Larus argentatus* spp.	Laridae	Female	nucl>mt	[51]
Manakins	*Manacus candei-M. vitellinus*	Pipridae	Female	nucl>mt	[52]
Lazuli and indigo buntings	*Passerina amoena-P. cyanea*	Cardinalidae	Female	nucl>mt	[53]
Chiffchaffs	*Phylloscopus brehmi-P. collybita*	Sylviidae	Female	nucl>mt	[54]
Willow warbler and chiffchaff	*Phylloscopus trochilus-P. collybita*	Sylviidae	Female	nucl>mt	[55]
Golden- and blue-winged warblers	*Vermivora chrysoptera-V. pinus*	Parulidae	Female	nucl>mt	[56]
Chimpanzee	*Pan t. troglodytes-P. verus*	Hominidae	Male	nucl>mt=Y	[57]
Heliconius butterflies	*Heliconius cydno-H. melpomene*	Nymphalidae	Female	nucl>mt	[58]
Tree weta	*Hemideina thoracica* chromos. races	Anostostomatidae	Male XO	nucl>mt	[59]
Admiral butterflies	*Limenitis a. arthemis-L. a. astyanax*	Nymphalidae	Female	nucl>mt	[60]
Water striders	*Limnoporus dissortis-L. notabilis*	Gerridae	Male	nucl>mt	[61]
Swallowtails	*Papilio machaon-P. hospiton*	Papilionidae	Female	nucl>mt	[62]
With male-biased dispersal					
Gadwall and falcated ducks	*Anas strepera-A. falcate*	Anatidae	Female	mt>nucl	[63]
Goats and wild relatives	*Capra sp.*	Bovidae	Male	mt>Y	[64]
Red-backed voles	*Eothenomys andersoni-E. smithii*	Cricetidae	Male	mt>nucl>Y	[65]
Hares	*Lepus timidus-L. europaeus-L. granatus*	Leporidae	Male	mt>nucl	[66]
African elephants	*Loxodonta africana-L. cyclotis*	Proboscidea	Male	mt>Y=nucl	[67]
Long-tailed and Rhesus macaques	*Macaca fascicularis-M. mulatta*	Cercopithecidae	Male	Y>nucl>mt	[31]
Sulawesi macaques	*Macaca nemestrina ssp.*	Cercopithecidae	Male	mt>nucl	[68]
House mouse	*Mus m. musculus-M. m. domesticus*	Muridae	Male	mt>nucl>Y=X	[26]
Mouse-eared bats	*Myotis myotis-M. blythii*	Vespertilionidae	Male	mt>nucl	[38]
Mule and white-tailed deers	*Odocoileus hemione-O. virginianus*	Cervidae	Male	mt>Y=nucl	[25]
European rabbit	*Oryctolagus c. cuniculus-O. c. algirus*	Leporidae	Male	nucl>mt>Y	[69]
Common shrew	*Sorex araneus* chromos. races	Soricidae	Male	mt>Y	[70]
Ground squirrels	*Spermophilus sp.*	Sciuridae	Male	mt>Y=nucl	[71,72]
Chipmunks	*Tamias ruficaudus-T. amoenus*	Sciuridae	Male	mt>nucl	[73]
Bark beetles	*Aphanarthrum glabrum-A. subglabrum*	Curculionidae	Male	mt>nucl	[74]
Carabid beetles	*Carabus sp.*	Carabidae	Male	mt>nucl	[75]
Common yellow butterfly	*Eurema hecabe* colour types	Pieridae	Female	mt>nucl	[76]
Tiger moths	*Grammia sp.*	Noctuidae	Female	mt>nucl	[77]
Blue butterfly	*Lycaeides m. melissa-L. m. samuelis*	Lycaenidae	Female	mt>nucl	[78]
Damselflies	*Mnais costalis-M. pruinosa*	Calopterygidae	Male XO	mt>nucl	[79]
Sawflies	*Neodiprion leconti* group	Diprionidae	Male haploid	mt>nucl	[80]

[a]Birds are shown in yellow, mammals in green and insects in blue.
[b]Abbreviations: mt, mtDNA markers; nucl, nuclear markers; Y, Y chromosome markers.
[c]Orange indicates that introgression at mtDNA is lower than at nuclear or Y chromosome markers, and grey indicates that introgression at mtDNA markers is larger than at nuclear or Y chromosome markers.

where Haldane's rule could not explain patterns of introgression (e.g. [25]; see Box 3). Hence, the level of intraspecific gene flow seems a better predictor of rates of introgression than Haldane's rule, even if asymmetric sterility barriers could sometimes play a role in generating differences in introgression rates across markers.

It has also been argued that organelle markers should introgress more easily than nuclear markers because they are less likely to hitchhike with a region under selection that prevents introgression. If true, this mechanism should hold for any type of organism, but mtDNA was found to introgress less easily than nuclear DNA markers in most birds and in some insect species (Table 1). Nevertheless, this explanation could account for some of the discrepancies between observations and predictions of the intraspecific gene flow model, for instance in the house mouse, where selection on the X and Y chromosomes has been inferred [26].

Poor taxonomic resolution can be caused not only by introgression but also by the retention of ancestral polymorphisms in descendant species [27]. In that case as well, markers experiencing less gene flow should be less

diagnostic, for two reasons [23,28]. First, the effective size of a subdivided population is larger than that of a random mating population with the same census size [29]. Therefore, ancestral lineages will be lost by drift less rapidly at markers experiencing low levels of gene flow. Second, limited gene flow precludes selective sweeps within species, which can be important in establishing or maintaining differences between species (see Box 1). In contrast with the model of introgression during range expansion, which predicts higher introgression in the colonising species [13], models of retention of ancestral polymorphism and of successful spread of variants within species do not account for asymmetries in the distribution of variation between the colonising and the local species, except by chance.

A higher propensity for genome components subject to low gene flow to introgress is only expected when species have colonised regions where related species already existed ([12]; see Box 2). This will be the case if the species have evolved in allopatry and are now at least partly sympatric or if they form moving hybrid zones [30]. Although these situations should have frequently occurred, there might

Box 3. Contrasted introgression of mtDNA and nuclear DNA in bats *Myotis* and in birds *Hippolais*

We illustrate here two case studies (taken from Table 1), one on bats and one on birds, where researchers have detected diametrically opposite patterns of introgression at mtDNA and nuclear markers. We attribute this difference to the opposite patterns of dispersal of the two sexes in the corresponding species.

Widespread mtDNA introgression in *Myotis* bats
The two sibling mouse-eared bat species *Myotis myotis* and *M. blythii* (Figure Ia) occur in sympatry over wide areas of Southern and Central Europe, with *M. blythii*, originating from Asia, representing the latest coloniser. Morphological and ecological evidence show that the two species constitute differentiated groups. Gene flow is highly male biased in these bats as in most mammal species [10]. Berthier *et al.* [38] studied mtDNA and nuclear diversity in sympatric and allopatric populations of both species. European *M. blythii* share multiple identical mtDNA haplotypes with *M. myotis*, whereas allopatric Asian *M. blythii* individuals have very divergent mtDNA sequences, suggesting that the mitochondrial genome of the European *M. blythii* had been replaced by that of *M. myotis*. By contrast, both species remain well differentiated at nuclear markers in mixed nursery

colonies. Hence, the results fit well with our expectation that in species with male-biased dispersal, mtDNA markers should introgress more readily than biparentally inherited nuclear markers.

Lack of mtDNA introgression in *Hippolais* birds
Hippolais icterina and *H. polyglotta* are two morphologically similar European warblers (Figure Ib). However, they are easily discriminated by wing characteristics and song. In this genus, males are highly territorial and females are the dispersing sex, as in most bird species [10]. In Western Europe, *H. polyglotta* has expanded north- and eastward at the expense of *H. icterina* over the past 40 years. Secondi *et al.* [50] have investigated the genetic structure of both species in this moving contact zone. They found no instances of cross-species transmission of mtDNA but detected clear signal of nuclear introgression, with higher levels of introgression in the expanding species, as predicted by previous simulation studies [12,13]. There was no evidence for Haldane's effect, as hybrids of both sexes were found. Altogether, the findings fit well with our prediction that in species with female-biased dispersal, nuclear markers should introgress more readily than maternally inherited mtDNA markers.
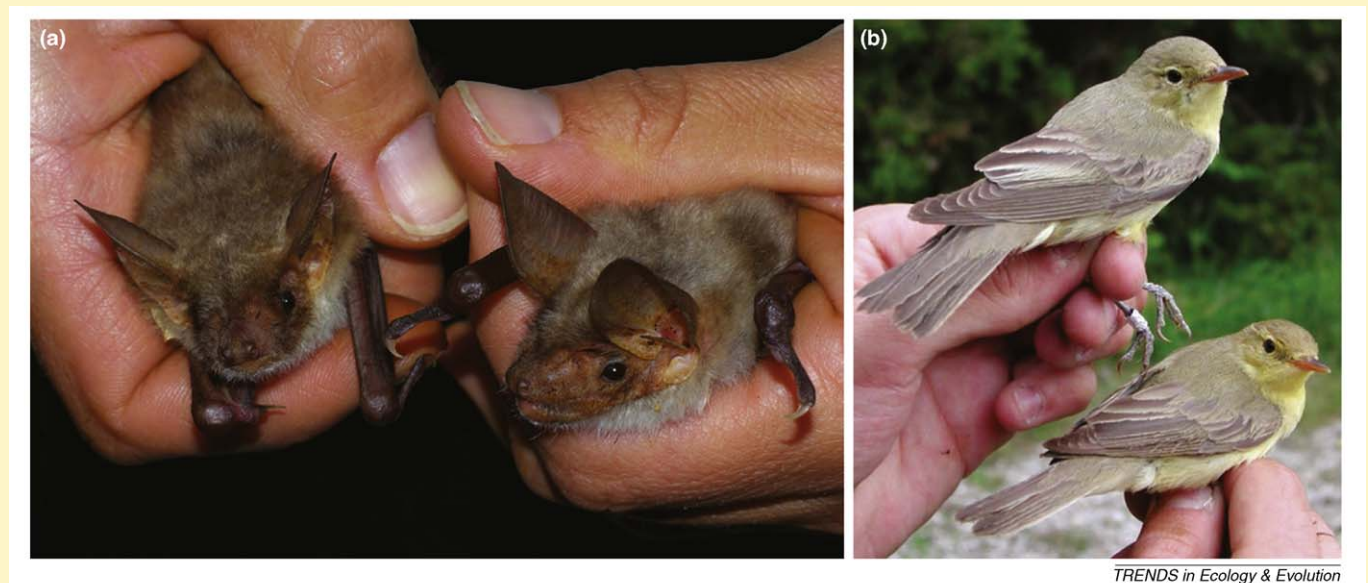

Figure I. Two pairs of sibling species showing opposite patterns of sex-biased dispersal and introgression. **(a)** Lesser mouse-eared bat (*Myotis blythii*) (left) and greater mouse-eared bat (*Myotis myotis*) (right). **(b)** Icterine warbler (*Hippolais icterina*) (above) with melodious warbler (*H. polyglotta*) (below). Photographs by Paolo Debernardi and Ferran López.

be situations where species have only recently established a stable hybrid zone. In such a case, the more intuitive prediction that intra- and interspecific gene flow must be positively correlated could hold. This appears to be the case in one of the two exceptions that we detected in our survey (Table 1). In a hybrid zone between Rhesus macaque and long-tailed macaque (two species characterised by male-biased dispersal), there is evidence for Y chromosome but not for mtDNA introgression [31]. It appears that only male Rhesus macaques move into the range of long-tailed macaques, thereby accounting for the massive introgression of Y chromosomes, which totally replaced local Y chromosomes without a large impact on the global phenotype of the long-tailed macaque. This unusual process might have been facilitated by the low density of the long-tailed macaque population [32] and the generally low male effective population sizes in macaques [33].

Perspectives for species delimitation
Our counterintuitive prediction that markers experiencing more intraspecific gene flow should have more difficulty in crossing species barriers seems to apply broadly. Alternative explanations for observed differences in introgression rates among markers have less explanatory power, even though they might be valid in specific cases. The model we propose does not rely on selective processes, and should therefore be considered as a null neutral model that could be useful for hypothesis testing. It should be confirmed by future experimental studies and cross-species comparisons. In particular, fish, amphibians and reptiles could represent valuable additional systems to validate the model, owing to the diversity of behaviours and of sex-determining mechanisms in these groups. A more systematic use of multiple markers with different modes of inheritance and a better knowledge of sex-specific dispersal [34]

should help increase the power of future comparative studies, but some conclusions regarding species delimitation can already be drawn.

Our findings should help explain controversial results of past taxonomic efforts aiming at delimitating species with different molecular markers. For instance, organelle DNA sequences have proved particularly useful in delimiting taxonomically problematic species in Lepidoptera [3] and birds [4,5], whereas being much less successful in flies [35], beetles [6], grasshoppers [36] or plants [37]. Taking into account the level of gene flow experienced by mtDNA or cpDNA markers in these different organisms could help explain these observations.

Clearly, the results presented here caution against the use of uniparentally inherited markers for species delimitation when they are inherited only from the least-dispersing sex. By contrast, markers inherited from the highest-dispersing sex were shown to better differentiate hybridising taxa. However, we do not advise the use of markers from a single uniparentally inherited genome for species delimitation, even if they experience high levels of gene flow. Despite their potential advantages, they still represent a single realisation of evolution, from which it would be risky to derive strong conclusions regarding species status. Multilocus assignment methods (e.g. [38,39]) have considerably more power and can also inform on individual admixture coefficients. Although all unlinked markers having the same mode of inheritance should experience similar rates of gene flow, selection can greatly modify the expected patterns of allelic frequencies within and between species. In particular, some markers might have experienced selective sweeps within species ([40]; see Box 1). Such markers, if selected against in the background of related hybridising species, should provide the most powerful tools to delimitate species (e.g. [41]). This expands our argument that high gene flow *sensu lato* (rapid gene circulation *sensu* Beurton [42]; see Box 1) is the key to effective species delimitation. As a consequence, we propose that multiple unlinked high gene flow markers, coupled with the use of model-based assignment methods, represent the best option to efficiently delimitate species. However, a single marker might represent a cost-effective approach for barcoding projects of taxonomically already well-described species [43], provided that it is inherited through the sex dispersing most.

Perspectives for studies of speciation

The ideas reported in this paper emerged while studying the demo-genetic processes taking place during species interactions in the context of range shifts and expansions. Our findings suggest that previously assumed relationships between intra- and interspecific gene flow were seriously misleading, because they did not take into account the spatial and dynamic aspects of species hybridisation. The process we have outlined here provides a mechanistic basis for the recent observation that hybridisation fails to cause taxonomic problems despite its frequent mention as the primary cause of 'fuzzy' species boundaries in plants [44]. The new paradigm proposed here should help explain why species integrity can be maintained in the face of extensive interspecific gene

flow (e.g. [45,46]), leading us to reconsider the role of gene flow in speciation. The negative correlation we observe between levels of intra- and interspecific gene flow and the realisation that high rates of gene flow within species promotes rapid monophyly at the species level indicate that enhanced gene flow is beneficial for the maintenance of species and their cohesive evolution. By contrast, reduced gene flow might represent a threat to species integrity. These results thus raise the question of the existence of selective processes actively promoting gene flow within species (see Box 1).

References

1 Sites, J.W. and Marshall, J.C. (2003) Delimiting species: a renaissance issue in systematic biology. *Trends Ecol. Evol.* 18, 462–470
2 de Queiroz, K. (2007) Species concepts and species delimitation. *Syst. Biol.* 56, 879–886
3 Hajibabaei, M. *et al.* (2006) DNA barcodes distinguish species of tropical Lepidoptera. *Proc. Natl. Acad. Sci. U. S. A.* 103, 968–971
4 Hebert, P.D.N. *et al.* (2004) Identification of birds through DNA barcodes. *PLoS Biol.* 2, e312
5 Tavares, E. and Baker, A. (2008) Single mitochondrial gene barcodes reliably identify sister-species in diverse clades of birds. *BMC Evol. Biol.* 8, 81
6 Monaghan, M.T. *et al.* (2005) DNA-based species delineation in tropical beetles using mitochondrial and nuclear markers. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 360, 1925–1933
7 Rubinoff, D. *et al.* (2006) A genomic perspective on the shortcomings of mitochondrial DNA for "barcoding" identification. *J. Hered.* 97, 581–594
8 Morjan, C.L. and Rieseberg, L.H. (2004) How species evolve collectively: implications of gene flow and selection for the spread of advantageous alleles. *Mol. Ecol.* 13, 1341–1356
9 Rieseberg, L.H. and Burke, J.M. (2001) The biological reality of species: gene flow, selection, and collective evolution. *Taxon* 50, 47–67
10 Greenwood, P.J. (1980) Mating systems, philopatry and dispersal in birds and mammals. *Anim. Behav.* 28, 1140–1162
11 Vähä, J-P. and Primmer, C.R. (2006) Efficiency of model-based Bayesian methods for detecting hybrid individuals under different hybridization scenarios and with different numbers of loci. *Mol. Ecol.* 15, 63–72
12 Currat, M. *et al.* (2008) The hidden side of invasions: massive introgression by local genes. *Evolution Int. J. Org. Evolution* 62, 1908–1920
13 Currat, M. and Excoffier, L. (2004) Modern humans did not admix with Neanderthals during their range expansion into Europe. *PLoS Biol.* 2, e421
14 Lawson Handley, L. and Perrin, N. (2007) Advances in our understanding of mammalian sex-biased dispersal. *Mol. Ecol.* 16, 1559–1578
15 Clarke, A.L. *et al.* (1997) Sex biases in avian dispersal: a reappraisal. *Oikos* 79, 429–438
16 Palumbi, S. *et al.* (2001) Predicting nuclear gene coalescence from mitochondrial DNA: the three-times rule. *Evolution Int. J. Org. Evolution* 55, 859–868
17 Petit, R.J. *et al.* (2005) Comparative organization of chloroplast, mitochondrial and nuclear diversity in plant populations. *Mol. Ecol.* 14, 689–701
18 Rieseberg, L.H. *et al.* (1996) Molecular marker incongruence in plant hybrid zones and phylogenetic trees. *Acta Bot. Neerl.* 45, 243–262

19 Dumolin-Lapègue, S. *et al.* (1999) Are chloroplast and mitochondrial DNA variation species independent in oaks? *Evolution Int. J. Org. Evolution* 53, 1406–1413

20 McKinnon, G.E. *et al.* (2001) Chloroplast sharing in the Tasmanian eucalypts. *Evolution Int. J. Org. Evolution* 55, 703–711

21 Wolf, P.G. and Sipes, S.D. (1997) Species-independent, geographical structuring of chloroplast DNA haplotypes in a montane herb *Ipomopsis* (Polemoniaceae). *Mol. Ecol.* 6, 283–291

22 Petit, R.J. and Vendramin, G.G. (2002) Plant phylogeography based on organelle genes: an introduction. In *Phylogeography of the Southern European Refugia* (Weiss, S. and Ferrand, N., eds), pp. 23–97, Springer

23 Du, F. *et al.* (2009) More introgression with less gene flow: chloroplast versus mitochondrial DNA in the *Picea asperata* complex in China, and comparison with other conifers. *Mol. Ecol.* 18, 1396–1407

24 Haldane, J. (1922) Sex ratio and unisexual sterility in hybrid animals. *J. Genet.* 12, 101–109

25 Cathey, J.C. *et al.* (1998) Introgressive hybridization and nonconcordant evolutionary history of maternal and paternal lineages in North American deer. *Evolution Int. J. Org. Evolution* 52, 1224–1229

26 Dod, B. *et al.* (1993) Counterselection on sex-chromosomes in the *Mus musculus* European hybrid zone. *J. Evol. Biol.* 6, 529–546

27 Funk, D.J. and Omland, K.E. (2003) Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. *Annu. Rev. Ecol. Evol. Syst.* 34, 397–423

28 Hoelzer, G.A. (1997) Inferring phylogenies from mtDNA variation: mitochondrial-gene trees versus nuclear-gene trees revisited. *Evolution Int. J. Org. Evolution* 51, 622–626

29 Wright, S. (1943) Isolation by distance. *Genetics* 28, 114–138

30 Buggs, R.J.A. (2007) Empirical study of hybrid zone movement. *Heredity* 99, 301–312

31 Bonhomme, M. *et al.* (2008) Assessing natural introgression in two biomedical model species, the Rhesus macaque (*Macaca mulatta*) and the long-tailed macaque (*Macaca fascicularis*). *J. Hered.* 100, 158–169

32 Hamada, Y. *et al.* (2008) Morphological and body color variation in Thai *Macaca fascicularis fascicularis* north and south of the Isthmus of Kra. *Int. J. Primatol.* 29, 1271–1294

33 Tosi, A. *et al.* (2003) Paternal, maternal, and biparental molecular markers provide unique windows onto the evolutionary history of macaque monkeys. *Evolution Int. J. Org. Evolution* 57, 1419–1435

34 Mossman, C. and Waser, P. (1999) Genetic detection of sex-biased dispersal. *Mol. Ecol.* 8, 1063–1067

35 Meier, R. *et al.* (2006) DNA barcoding and taxonomy in Diptera: a tale of high intraspecific variability and low identification success. *Syst. Biol.* 55, 715–728

36 Trewick, S. (2008) DNA barcoding is not enough: mismatch of taxonomy and genealogy in New Zealand grasshoppers (Orthoptera: Acrididae). *Cladistics* 24, 240–254

37 Little, D. and Stevenson, D. (2007) A comparison of algorithms for the identification of specimens using DNA barcodes: examples from gymnosperms. *Cladistics* 23, 1–21

38 Berthier, P. *et al.* (2006) Recurrent replacement of mtDNA and cryptic hybridization between two sibling bat species *Myotis myotis* and *Myotis blythii*. *Proc. R. Soc. Lond. B Biol. Sci.* 273, 3101–3109

39 Duminil, J. *et al.* (2006) Blind population genetics survey of tropical rainforest trees. *Mol. Ecol.* 15, 3505–3513

40 Rieseberg, L.H. *et al.* (2004) Integration of populations and differentiation of species. *New Phytol.* 161, 59–69

41 Lexer, C. *et al.* (2006) Shared alleles in sympatric oaks: recurrent gene flow is a more parsimonious explanation than ancestral polymorphism. *Mol. Ecol.* 15, 2007–2012

42 Beurton, P. (1995) How is a species kept together? *Biol. Philos.* 10, 181–196

43 Rubinoff, D. (2006) Utility of mitochondrial DNA barcodes in species conservation. *Conserv. Biol.* 20, 1026–1033

44 Rieseberg, L.H. *et al.* (2006) The nature of plant species. *Nature* 440, 524–527

45 Fitzpatrick, B.M. *et al.* (2008) Distinctiveness in the face of gene flow: hybridization between specialist and generalist gartersnakes. *Mol. Ecol.* 17, 4107–4117

46 Strasburg, J. and Rieseberg, L. (2008) Molecular demographic history of the annual sunflowers *Helianthus annuus* and *H. petiolaris*—large

47 Martínez-Cruz, B. and Godoy, J.A. (2007) Genetic evidence for a recent divergence and subsequent gene flow between Spanish and Eastern imperial eagles. *BMC Evol. Biol.* 7, 170

48 Helbig, A.J. *et al.* (2005) Genetic differentiation and hybridization between greater and lesser spotted eagles (Accipitriformes: *Aquila clanga*. A. pomarina). *J. Ornithol.* 146, 226–234

49 Sætre, G-P. *et al.* (2001) Speciation, introgressive hybridization and nonlinear rate of molecular evolution in flycatchers. *Mol. Ecol.* 10, 737–749

50 Secondi, J. *et al.* (2006) Spreading introgression in the wake of a moving contact zone. *Mol. Ecol.* 15, 2463–2475

51 Crochet, P.A. *et al.* (2003) Genetic differentiation at nuclear and mitochondrial loci among large white-headed gulls: sex-biased interspecific gene flow? *Evolution Int. J. Org. Evolution* 57, 2865–2878

52 Brumfield, R.T. *et al.* (2001) Evolutionary implications of divergent clines in an avian (*Manacus*: Aves) hybrid zone. *Evolution Int. J. Org. Evolution* 55, 2070–2087

53 Carling, M.D. and Brumfield, R.T. (2008) Haldane's rule in an avian system: using cline theory and divergence population genetics to test for differential introgression of mitochondrial, autosomal, and sex-linked loci across the *Passerina* bunting hybrid zone. *Evolution Int. J. Org. Evolution* 62, 2600–2615

54 Helbig, A.J. *et al.* (2001) Male-biased gene flow across an avian hybrid zone: evidence from mitochondrial and microsatellite DNA. *J. Evol. Biol.* 14, 277–287

55 Bensch, S. *et al.* (2006) Conflicting patterns of mitochondrial and nuclear DNA diversity in *Phylloscopus* warblers. *Mol. Ecol.* 15, 161–171

56 Vallender, R. *et al.* (2007) Complex hybridization dynamics between golden-winged and blue-winged warblers (*Vermivora chrysoptera* and *Vermivora pinus*) revealed by AFLP, microsatellite, intron and mtDNA markers. *Mol. Ecol.* 16, 2017–2029

57 Won, Y.J. and Hey, J. (2005) Divergence population genetics of chimpanzees. *Mol. Biol. Evol.* 22, 297–307

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

59 Morgan-Richards, M. and Wallis, G.P. (2003) A comparison of five hybrid zones of the weta *Hemideina thoracica* (Orthoptera: Anostostomatidae): degree of cytogenetic differentiation fails to predict zone width. *Evolution Int. J. Org. Evolution* 57, 849–861

60 Mullen, S.P. *et al.* (2008) Hybrid zone origins, species boundaries, and the evolution of wing-pattern diversity in a polytypic species complex of North American admiral butterflies (Nymphalidae: *Limenitis*). *Evolution Int. J. Org. Evolution* 62, 1400–1417

61 Abe, T.A. *et al.* (2005) Mitochondrial introgression is restricted relative to nuclear markers in a water strider (Hemiptera: Gerridae) hybrid zone. *Can. J. Zool.* 83, 432–444

62 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

63 Peters, J.L. *et al.* (2007) Nuclear loci and coalescent methods support ancient hybridization as cause of mitochondrial paraphyly between gadwall and falcated duck (*Anas* spp.). *Evolution Int. J. Org. Evolution* 61, 1992–2006

64 Pidancier, N. *et al.* (2006) Evolutionary history of the genus *Capra* (Mammalia, Artiodactyla): discordance between mitochondrial DNA and Y-chromosome phylogenies. *Mol. Phylogenet. Evol.* 40, 739–749

65 Iwasa, M.A. and Suzuki, H. (2003) Intra- and interspecific genetic complexities of two *Eothenomys* species in Honshu. *Japan. Zoolog. Sci.* 20, 1305–1313

66 Alves, P.C. *et al.* (2003) Ancient introgression of *Lepus timidus* mtDNA into *L. granatensis* and *L. europaeus* in the Iberian Peninsula. *Mol. Phylogenet. Evol.* 27, 70–80

67 Roca, A.L. *et al.* (2005) Cytonuclear genomic dissociation in African elephant species. *Nat. Genet.* 37, 96–100

68 Evans, B.J. *et al.* (2003) Diversification of Sulawesi macaque monkeys: decoupled evolution of mitochondrial and autosomal DNA. *Evolution Int. J. Org. Evolution* 57, 1931–1946

69 Geraldes, A. *et al.* (2008) Reduced introgression of the Y chromosome between subspecies of the European rabbit (*Oryctolagus cuniculus*) in the Iberian Peninsula. *Mol. Ecol.* 17, 4489–4499

effective population sizes and rates of long-term gene flow. *Evolution Int. J. Org. Evolution* 62, 1936–1950

70 Balloux, F. *et al.* (2000) Microsatellites can be misleading: an empirical and simulation study. *Evolution Int. J. Org. Evolution* 54, 1414–1422

71 Ermakov, O.A. *et al.* (2006) A search for Y-chromosomal species-specific markers and their use for hybridization analysis in ground squirrels (*Spermophilus*: Rodentia, Sciuridae). *Russ. J. Genet.* 42, 429–438

72 Ermakov, O.A. *et al.* (2002) A molecular genetic study of hybridization in four species of ground squirrels (*Spermophilus*: Rodentia, Sciuridae). *Russ. J. Genet.* 38, 796–809

73 Good, J.M. *et al.* (2008) Ancient hybridization and mitochondrial capture between two species of chipmunks. *Mol. Ecol.* 17, 1313–1327

74 Jordal, B.H. *et al.* (2006) Apparent 'sympatric' speciation in ecologically similar herbivorous beetles facilitated by multiple colonizations of an island. *Mol. Ecol.* 15, 2935–2947

75 Sota, T. (2002) Radiation and reticulation: extensive introgressive hybridization in the carabid beetles *Ohomopterus* inferred from mitochondrial gene genealogy. *Popul. Ecol.* 44, 145–156

76 Narita, S. *et al.* (2006) Genetic structure of sibling butterfly species affected by Wolbachia infection sweep: evolutionary and biogeographical implications. *Mol. Ecol.* 15, 1095–1108

77 Schmidt, B.C. and Sperling, F.A.H. (2008) Widespread decoupling of mtDNA variation and species integrity in *Grammia* tiger moths (Lepidoptera: Noctuidae). *Syst. Entomol.* 33, 613–634

78 Gompert, Z. *et al.* (2006) Identifying units for conservation using molecular systematics: the cautionary tale of the Karner blue butterfly. *Mol. Ecol.* 15, 1759–1768

79 Hayashi, F. *et al.* (2005) Disturbed population genetics: suspected introgressive hybridization between two *Mnais* damselfly species (Odonata). *Zoolog. Sci.* 22, 869–881

80 Linnen, C.R. and Farrell, B.D. (2007) Mitonuclear discordance is caused by rampant mitochondrial introgression in *Neodiprion* (Hymenoptera: Diprionidae) sawflies. *Evolution Int. J. Org. Evolution* 61, 1417–1438

81 Mayr, E. (1942) *Systematics and the Origin of Species.* Columbia University Press

82 Mayr, E. (1963) *Animal Species and Evolution.* Harvard University Press

83 Ehrlich, P.R. and Raven, P.H. (1969) Differentiation of populations. *Science* 165, 1228–1232

84 Levin, D. and Kerster, H. (1974) Gene flow in seed plants. *Evol. Biol.* 7, 138–220

85 Hoelzer, G.A. *et al.* (2008) Isolation-by-distance and outbreeding depression are sufficient to drive parapatric speciation in the absence of environmental influences. *PLoS Comp. Biol.* 4, e1000126

86 Mallet, J. (2001) Species, concepts of. In *Encyclopedia of Biodiversity* (Levin, S., ed.), pp. 427–440, Academic Press

87 Barker, M.J. (2007) The empirical inadequacy of species cohesion by gene flow. *Philos. Sci.* 74, 654–665

88 Stebbins, G. (1950) *Variation and Evolution in Plants.* Columbia University Press

89 Hewitt, G. (1993) Postglacial distribution and species substructure: lessons from pollen, insects and hybrid zones. In *Evolutionary Patterns and Processes* (Lees, D. and Edwards, D., eds), pp. 97–123, Academic Press

90 Hubbs, C.L. (1955) Hybridization between fish species in nature. *Syst. Zool.* 4, 1–20

91 Klopfstein, S. *et al.* (2006) The fate of mutations surfing on the wave of a range expansion. *Mol. Biol. Evol.* 23, 482–490