How reticulated are species?

James Mallet[1)2)]*, Nora Besansky[3)] and Matthew W. Hahn[4)]

Many groups of closely related species have reticulate phylogenies. Recent genomic analyses are showing this in many insects and vertebrates, as well as in microbes and plants. In microbes, lateral gene transfer is the dominant process that spoils strictly tree-like phylogenies, but in multicellular eukaryotes hybridization and introgression among related species is probably more important. Because many species, including the ancestors of ancient major lineages, seem to evolve rapidly in adaptive radiations, some sexual compatibility may exist among them. Introgression and reticulation can thereby affect all parts of the tree of life, not just the recent species at the tips. Our understanding of adaptive evolution, speciation, phylogenetics, and comparative biology must adapt to these mostly recent findings. Introgression has important practical implications as well, not least for the management of genetically modified organisms in pest and disease control.

Introduction

Not so long ago, analysis of microbial 16S ribosomal RNA sequences led to a revolutionary new "Universal Tree of Life," consisting of three monophyletic domains, here referred to as the Bacteria, the Archaea, and the Eukarya or eukaryotes [1, 2]. Yet almost as soon as the new system was established, this

[1)] Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA, USA
[2)] Department of Genetics, Evolution and Environment, University College London, London, UK
[3)] Department of Biological Sciences and Eck Institute for Global Health, University of Notre Dame, Notre Dame, IN, USA
[4)] Department of Biology and School of Informatics and Computing, Indiana University, Bloomington, IN, USA

*Corresponding author:
James Mallet
E-mail: jmallet@oeb.harvard.edu

tidy tree picture was threatened: sequencing of more microbial genes and then whole genomes quickly led to an understanding of the importance of horizontal or lateral gene transfer, the incorporation of foreign genes into the genome. Some of the major transitions in evolution were clearly due to lateral transfer: the eukaryotes were formed by endosymbiosis of α-proteobacteria with Archaea to form the eukaryotes. Later, endosymbiosis of cyanobacteria with eukaryotes led to green algae and plants. Many other gene transfers together with multiple other endosymbioses have been inferred. Microbiologists began to argue that the "tree" of life was more like a web or network than a tree [3–5].

Today, whole genome sequencing is providing unprecedented phylogenetic information about whole groups of eukaryotes [6–14]. Here we review genomic evidence suggesting that reticulate evolution may have considerable impact in multicellular eukaryotes as

well as microbes. Reproductively isolated species and bifurcating phylogenies have become an important basis for our understanding of evolution; now this bedrock seems threatened. As an ideal, species are often taken to be evolutionarily independent populations that are reproductively isolated from other such species, for example in the "biological species concept," although it was always known that hybridization does occur [15]. Reticulate evolution in plants has long been recognized [16], but recent genomic evidence from animals suggest that reticulation might be much more common than anticipated [17, 18]. Given abundant new data, it is time to enquire whether a major shift in our understanding of species, speciation, and phylogenetics is taking place.

Prokaryotes: Is there a universal tree of life?

Tree-like relationships among species arise because the genome evolves within cells. When a cell divides, copies of the same genome are found in each daughter cell. Ultimately, after populations of organisms diverge or "speciate," evolution along each branch will leave genomic signals of that branching event in daughter lineages. Sex and recombination can obscure this picture, but in both Bacteria and Archaea sex (in the eukaryote-like sense of homologous gene exchange) is mostly a transaction between closely related individuals, mostly within the same populations or "species" [19–22]. Eukaryotes are similar [23, 24]. Lateral transfer involving non-homologous exchange, on the

other hand, will lead to more wide-ranging phylogenetic discordance. In prokaryotes, both sex and lateral transfer involve relatively few genes at a time or even if more extensive, usually much less than 50% of the genome. Nevertheless, multiple exchanges may take place, and very large fractions of the genome might eventually be exchanged with other lineages or species over long periods. If so, it is possible that the signals of the organismic genealogy (the original "tree of cells") in the genome will be obliterated by multiple phylogenetic signals from sex and lateral transfer.

Before assessing new genomic evidence for phylogenetic discordance in multicellular eukaryotes, it is worth reviewing the controversy raging about the microbial "Tree of Life" over the last few decades. Carl Woese [25] argued that in spite of considerable lateral transfer, there is "a genealogy-defining core of genes whose common history dates back to the root of the universal tree." Woese suggested that the acquisition of sufficient co-adaptation among these key genes caused life to reach a "Darwinian threshold," which permitted divergence into separate species and allowed us to trace the organismal history, even while lateral transfer obscures the universal tree for many other genes. According to Woese, before the Darwinian threshold was reached, divergence and speciation could not take place, and no tree of genes would allow us to trace the organismal history.

It quickly became apparent that lateral transfer does indeed swamp the signal of the Universal Tree in microbial genomes: in fact no other genes support Woese's original 16S RNA tree [26]. Many microbiologists now deny a tree-like phylogeny of microbial evolution; instead the phylogeny of life looks more like a web or a ring [3, 27–29]. By excluding all genes that disagree with the Universal Tree, one can select 20–30 largely informational genes that more or less rescue the ribosomal RNA Tree [29–31]. But this almost seems like cheating, and is itself obtained only by pruning out a number of clear cases of lateral transfer in even these genes. As this anyway only applies to a tiny fraction of the genome, these recent incarnations of the Universal Tree have been

derided as "the tree of one per cent" [32]. Around 80% of eukaryotic proteins are actually more closely related to homologs in the Bacteria than in the Archaea; the Universal Tree's closer archaeal-eukaryote affinity is reflected in only about 15% of eukaryote proteins [28, 32], including those used by Ciccarelli et al. [30]. Because of concerns such as these, the existence of species and of the Universal Tree in microbes has been dismissed as a "myth" in the prokaryote literature [33]. Whether species or the Universal Tree exist in prokaryotes has become almost a philosophical rather than a biological issue [29], but it does seem clear that most of the original Universal Tree, whether identifiable or not, is located on the far side of what Woese originally intended by the Darwinian Threshold.

What causes phylogenetic incongruence in eukaryotes?

Findings of promiscuous gene exchange among prokaryotes have usually been contrasted with supposedly well-behaved trees in eukaryotes [33, 34]. Eukaryote genomes originated when an archaeal cell acquired many bacterial genes, in part but certainly not only associated with the bacterial endosymbiotic origins of mitochondria and chloroplasts [35]. Eukaryotes also invented meiosis, which allows recombination of whole genomes. In multicellular eukaryotes, reproduction itself often involves meiosis. This innovation effectively destroys the tree-like signal in an organismal ("tree of cells") phylogeny. In every meiosis recombinant haploid genomes from two successful, independent cells are thrown together to form diploid zygotes, before the sum of the genetic material is haphazardly and approximately equally recombined into haploid daughter cells. A "tree of cells" justification for the eukaryote Tree of Life is no longer possible.

While tree-like patterns are readily discernible in eukaryote phylogenies, we here highlight recent evidence suggesting that a number of regions of the eukaryotic tree show similar

pathologies to those found in prokaryotes. This raises doubt about the eukaryotic Tree of Life as a whole. Apart from phylogenetic estimation error and homoplasy, there are three main causes of phylogenetic incongruence: lateral gene transfer, incomplete lineage sorting, and introgression.

Lateral transfer

In Eukaryotes, lateral or horizontal gene transfer is widespread, but is usually thought to be rare compared to that in prokaryotes [8, 36, 37]. It seems to be associated mainly with single-celled eukaryotes (the "protists"), especially those that engulf their food, or in multicellular organisms with parasites in close cellular contact with their hosts. Eukaryotes clearly seem to have acquired important genes via lateral transfer from both mitochondrial and chloroplast endosymbionts, but transfers also originate from other endosymbionts, parasites, and close associates [35]. Lateral processes in eukaryotes, in contrast to other possible causes of reticulation, may transfer genes between distantly related species, but typically involve relatively few genes at a time, as in prokaryotes. Lateral transfer is common in some multicellular groups [36], such as bdelloid rotifers, which, interestingly, lack meiotic sex [38, 39]. Horizontal gene transfer in the mitochondrial genomes of plants and yeasts is also widespread [40]. However, horizontal transfer is probably not an overriding factor in the evolution of the nuclear protein coding genes of most multicellular eukaryotes, unlike those of prokaryotes.

In contrast to the genes, eukaryotic genomes often consist largely of non-coding DNA, and 30–60% of this consists of recognizable mobile elements [41, 42]. Intergenic and intronic DNA is thought to originate largely via active or inactivated mobile genetic elements [43–45], most of which are thought to enter lineages via lateral transfer [46]. Mobile elements are particularly likely to be important in the evolution and spread of regulatory elements. Nonetheless, the introduction of new mobile elements via lateral transfer is rare, and the lifespan of

active proliferation via transposition is cut short by relatively rapid loss, inactivation and sequence degradation in the host genome [46].

Incomplete lineage sorting

The two main causes of gene tree – species tree discordance, at least for protein-coding genes in closely related groups of eukaryotes, are incomplete lineage sorting and introgression. Incomplete lineage sorting occurs when polymorphisms persist between speciation events, so that the actual (true) genealogical relationship of a gene or genome region differs from the true species branching pattern. As an example of incomplete lineage sorting, around 15% of human genes are more closely related to homologs in gorillas than to those in our true sister lineage, the chimpanzees, while another 15% of genes group gorilla and chimpanzee. This is expected from what we know about the ancestral effective population sizes of these species and the short time between human-gorilla and human-chimpanzee speciation events [47, 48].

In some cases, incomplete lineage sorting occurs as a result of balancing selection maintaining polymorphisms: when speciation occurs, both daughter species may maintain the same "trans-specific" polymorphisms, even though with recombination, the signal of ancestral origin may erode over time [49]. Good examples of shared polymorphisms between humans and apes are MHC [50] and ABO blood group loci [51], among other genes. In the species complex including the major mosquito vector, *Anopheles gambiae*, a very large chromosomal inversion, 2La (22 Mb in length, 8.5% of the total genome size) is maintained as a balanced polymorphism that has persisted across several speciation events [18].

Unlike lateral transfer and introgression, however, discordance created by incomplete lineage sorting does not imply phylogenetic reticulation at the level of species. It merely muddles the genomic signal of what might be a truly bifurcating phylogeny. In some trees with four or more taxa and rapid successive speciation events (the "anomaly zone" of phylogenetics), the species tree estimated from the gene trees has been shown to converge on an incorrect but highly significant solution [52, 53]. In spite of this "tyranny of the majority" in phylogenetic analysis, a coalescent-based analysis should nonetheless be able to retrieve the true bifurcation signal in spite of the confused gene tree signal [54, 55].

Introgression and reticulated evolution

The third source of phylogenetic incongruence, introgression, occurs when hybrids backcross and transfer genetic material between species. Hybridization may occur without strongly affecting the genomes of recipient populations if strongly resisted by selection, but genomic admixture results if the introgressed alleles are established.

Hybridization between related eukaryote species does occur reasonably frequently in nature; it is known to affect around 25% of the species of flowering plants and about 10% of animals [56–58]. The fraction of hybrids in natural populations, nevertheless, is usually low: natural interspecific hybridization rates in animals are typically 0.1% or less per generation in any species [57, 59]. Per generation hybridization rates can be much higher in some populations of plants and animals, where it reaches several per cent, for example in some oaks (*Quercus*), Darwin's finches, and some cases in *Heliconius* butterflies [60–63]; but these are probably exceptional. While some hybrids are sterile, a substantial fraction of such hybrids are at least partly fertile, leading to observed cases of backcrossing and introgression. It is important to realize that hybridization and introgression may occur among non-sister species as well as between sister species, especially during rapid adaptive radiations.

Closely related species hybridize more readily than more distant species [64]. The decline of natural hybridization rates with genetic distance, while noisy, may be very roughly approximated as exponential [59], mirroring the noisy decline of compatibility in meta-analyses of transformation experiments in prokaryotes and laboratory crosses in animals and plants [19–23, 64–67]. Thus, introgression tends to generate phylogenetic discordance mainly among closely related groups of species, unlike lateral gene transfer. This is a major difference between reticulate evolution in prokaryotes and eukaryotes: while lateral gene transfer weaves lineages together across disparate parts of a tree, introgression merely results in tangled knots on a local scale. Nonetheless, introgression has potentially important effects throughout the tree of life by obscuring relationships among lineages that diversified rapidly at any time, not just in those that did so recently.

Introgression was well known before the advent of genetic markers or genomics, and was long believed an important catalyst for adaptive evolutionary change in plants [16]. Introgression was thus familiar by the 1960s, but ideas of "coadapted gene complexes," and "the unity of the genotype" associated with the biological species concept led to a belief that hybridization had little importance in animals, at least. When hybridization did occur, it was often assumed to be unnatural and was attributed to environmental changes wrought by humans [68]. Because hybrids are mostly unfit, it was assumed that introgression among animal species very rarely had any long-term evolutionary impact [15].

With the potential for introgression, not only will individual gene trees tell different stories, but the actual organismal branching pattern between species will be reticulate, rather than strictly bifurcating. The true phylogeny may be approximately tree-like if introgression is rare and affects only a very small fraction of the genome, but will not be tree-like if introgression is common. However, the importance of introgression is only now becoming apparent with rapid genome sequencing.

Gene transfer is important in eukaryote genomes

The extent of introgression across the eukaryote tree

As we have seen, meiotic fertility has an increasing tendency to fail with genetic distance, but failure is often not complete in the closest hybrids. For this reason, introgression, which requires some fertile hybrid offspring, is most likely to occur among closely related

species. Hybridization between sister species will not usually affect the species tree topology, but will make the apparent divergence time between the species appear more recent [69]. However, if two widely distributed species interact in populations where they overlap, it may be possible that individual populations become on average more closely related locally to a sister species than to more distant conspecific populations. In contrast, hybridization and introgression among non-sister lineages can readily distort the species tree topology. If introgression between non-sister lineages is widespread across the genome, it may be very hard to retrieve the true bifurcation history of the species. This is because a unitary history of the genome may not exist; if inferred from multiple loci or whole genomes, this species tree may be meaningless or misleading. Here we discuss several recent examples from multicellular eukaryotes where this may have occurred. Interestingly, most of these examples come from rapid species radiations; these are exactly the cases in which closely related but non-sister species may be hybridizing.

The group of eight African mosquito species known as the *A. gambiae* complex radiated within the last 2 million years [18]. Species distributions overlap extensively, and in areas of sympatry hybrids have been recorded at rates of ~0.02–0.75% [70, 71]. Despite F1 hybrid male sterility in most cases, introgression is plausible through the backcrossing of vigorous and fertile F1 hybrid females. When genomes of multiple members of the *A. gambiae* complex were sequenced and compared, the inferred species tree was evident in only 2% of the genome, mainly on the X chromosome, whereas the majority tree in the rest of the genome yields a completely contradictory tree [18]. While some of these differences are due to incomplete lineage sorting, much of this discordance is due to introgression between two non-sister species (Fig. 1). This is particularly clear for the 2La inversion mentioned above, which is inferred to have been polymorphic in the ancestor of the complex, but is affected by three losses of 2L+ and one of 2La, as well as one fairly recent (1 Mya) introgression of 2La
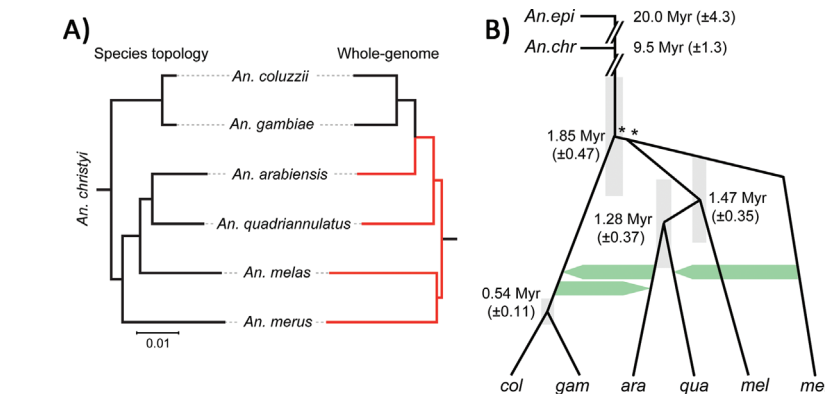
Figure 1. A: "Whole genome" versus "species" tree topologies of the *Anopheles gambiae* complex in Africa. **B:** The tree based on the X chromosome only, showing introgression events and estimated node divergence times. The average phylogeny of the whole genome is distorted by autosomal introgression between *A. gambiae* + *coluzzii* and *A. arabiensis*, but this was prevented on the X chromosome by X-linked hybrid incompatibilities and multiple overlapping inversions that prevent recombination. Modified and reprinted from [18] with permission from AAAS.

from *A. gambiae* to *A. arabiensis* [18]. Introgression is on-going, and is an excellent explanation for the phylogenetic discordance, because wild hybrids and backcrosses between the latter two species are ~0.22% of the individuals captured in sympatry [71, 72]. In deciding between conflicting topologies, the species tree was inferred from regions of the genome with the deepest coalescence times between species [18]. If this information had not been available, or if introgression had been even more complex, it would have been hard to infer the species tree at all.

In *Heliconius* butterflies, the "*mel-pomene*-silvaniform" clade consists of around 15 species. Most of these are "good" species that co-occur over large sympatric regions, and are somewhat interfertile with other members of the clade. However, rare hybrids and backcrosses are known from the wild and in captivity across this whole group, suggesting the possibility that a slow trickle of introgression is constantly occurring among the largely sympatric species in the group [59]. This suggestion has now been confirmed: because of introgression, a local population of *H. melpomene* can be more closely related to the locally overlapping population of its sister *H. cydno* than it is to conspecifics at over 40% of genome [17, 73].

Rapid radiations such as these tend to produce many closely related species that may be partially interfertile. For

example, per generation hybridization rates among closely related species of Darwin's finches can be as much as 6%, with high fertility of hybrids. The Darwin's finches began to diversify on the Cocos and Galapagos Islands less than 1 million years ago, and there is strong genomic evidence for past and continuing introgression across almost the entire group [74]. Other vertebrate groups such as African lake cichlids, *Xiphophorus* fishes, horses, and even hominins show similar phylogenetic discordance inferred to be due to introgression [75–78].

Much deeper evidence of reticulate evolutionary patterns also exists. For example, there is considerable phylogenetic discordance at the base of the Neoaves, or modern birds [79–81]. In fact *none* of the thousands of individual gene trees support the various conflicting estimates of the species tree [79, 81]. Trees built from indels and stable mobile element insertions (which are less prone to homoplasy than nucleotide or amino acid substitutions) show similar conflict, suggesting that the gene tree discordance is real, rather than due to phylogenetic error [79]. The authors of these papers argued that the tangle at the base of this ancient radiation was due to incomplete lineage sorting, but did not address the possibility of introgression. Yet introgression seems a likely additional cause: around 9% of today's bird species are known to hybridize in the wild [56], and birds retain some hybrid compatibility with congeners for ~10 My after

speciation [65]. After the demise of the dinosaurs, the early Neoaves had few competitors, and it is not unlikely that the first species in today's lineages were able to hybridize with one another during their global ecological diversification, much as the Darwin's finches do today on the Galapagos Islands [74]. An explanation for the strong signals of discordance at the base of the Neo-aves [79, 81] may therefore lie partly in gene flow among the lineages after they diversified. Given that other major groups, such as the placental mam-mals [13], or the animals as a whole [82], appear also to have evolved in rapid radiations, it seems likely that our persistent problems with estimation of trees for the deepest branches of these radiations is due to historical introgres-sion as well as incomplete lineage sorting during their initial diversification.

Is introgression adaptive?

Phylogenetic or genealogical studies of the extent of introgression across the genome do not, however, reveal whether the process is largely neutral or whether it is aided by a selective advantage on the new genomic back-ground. The relative importance of selection in introgression across the genome is still not known, and is an area of active research [83], but many introgression events are now known to have involved adaptation. A number of transfers of mimicry-determining loci have been documented in *Heliconius* butterflies (Fig. 2A and B), and in *Anopheles* the many cases of insecticide resistance alleles crossing species boundaries (see below) and the exis-tence of balancing selection at the 2La inversion make it rather hard to believe that selection is only rarely involved in introgression.

Adaptive introgression may also introduce adaptive combinations that lead to new species, or hybrid specia-tion [84, 85]. Plant examples have long been known [16, 85], but animal exam-ples are no longer rare. For example, the *Heliconius pardalinus*-like ancestor of *H. elevatus* seems to have recently acquired the majority of its defensive color pattern mimicry from *H. melpom-ene* (Fig. 2C), subsequently proving able to coexist in sympatry with both

parents [17]. That case remains to be fully worked out, but similar cases have been put forward for cichlid fish, monkeyflowers, and other hybridizing adaptive radiations [86, 87]. In one case, the beginnings of the process have been observed in "real time": a new hybrid finch species that breeds strictly endo-gamously has now been followed on a Galapagos island for seven generations since its formation via initial hybridiza-tion events in the early 1980s [88].

Introgression challenges notions of species and phylogeny

The meaning of species and speciation

We are thus confronted by extraordi-nary levels of introgression found in the genomes of rapidly radiating species (such as *Anopheles*, *Heliconius*, and Darwin's finches). Yet these taxa are currently readily identifiable to species using morphology or genetics: none of us doubt that the species is a useful rank, at least in multicellular eukar-yotes. We recognize these taxa as species not because of reproductive isolation per se, nor because they represent phylogenetic branching events, but because of the simpler observation that hybrids and intermedi-ates between the clusters we call species [89] are rare. While most of the introgression that has resulted in reticulate relationships occurred in the past – and may or may not be ongoing – these results suggest that species are like the Ship of Theseus in philosophy, which can progressively but almost completely be rebuilt with new wood, and yet remain the same ship. We do not yet know how common these effects are among genomes of other eukaryotes, but the recent discoveries in mammals, birds, fish, insects, plants, and fungi suggest that they may be widespread throughout the eukaryotic Tree of Life.

The "true phylogeny" versus the species tree

In introgressing species, different gene trees vary in the story they tell about

their genealogical history. The true phylogeny will trace the disparate histories of every gene, and cannot readily be represented on a page, certainly not as a single tree. Yet we propose that there may still be a true bifurcating tree of species out there (Fig. 3), in spite of the back-ground chaos of gene trees. Only if species fuse either wholly or in some geographic region to become a single cluster (e.g. in sticklebacks [90] or in hybrid speciation), does the species phylogeny itself become reticulate under this view.

Possible alternatives to the species tree is some consensus of gene trees, or perhaps the tree based on the "demo-cratic majority" of the genome [91]. Obtaining the maximum likelihood or most probable species tree from a series of genes is in fact the aim of many phylogenetic and phylogenomic stud-ies, at least among eukaryote system-atists [92, 93]. This program assumes that the true species tree is more likely to emerge via analysis of larger fractions of the genome. Under the viewpoint proposed here, this is not necessarily true if there is abundant introgression (Fig. 3). For example, as shown above, the single most common tree inferred from whole genomes of the *Anopheles gambiae* complex in Africa gives an incorrect rendering of the group's history [18] (Fig. 1).

Historical introgression events in taxa such as *Anopheles* have been inferred to affect the majority of the genome, even though natural hybrids are relatively rare among the contem-porary species (see above for rates of hybridization). Nonetheless, hybrid-ization can introduce variation at rates much higher than mutation, so that significant levels of genomic replace-ment may accrue over long periods, even at the low hybridization rates known in *Anopheles* today. Similar results also apply in some *Heliconius* species. If we wish the species tree to be determined by the democratic opinion of the genes, we are therefore forced to accept a peculiar species definition that perhaps applies only to terminal taxa, rather than the original bifurcating ancestors, because the branches of the tree change their species identity whenever accumula-tion of introgressed regions flips the
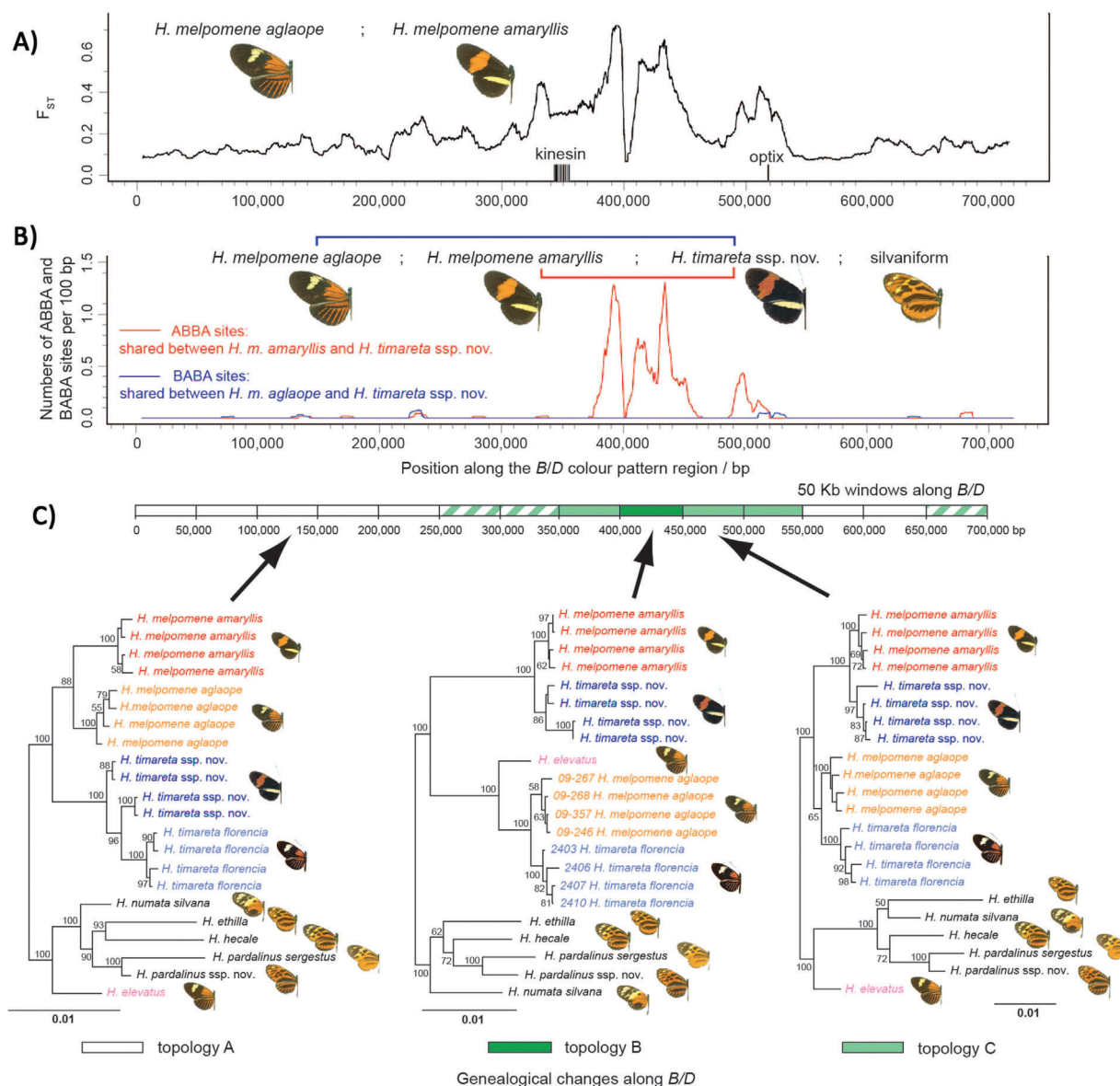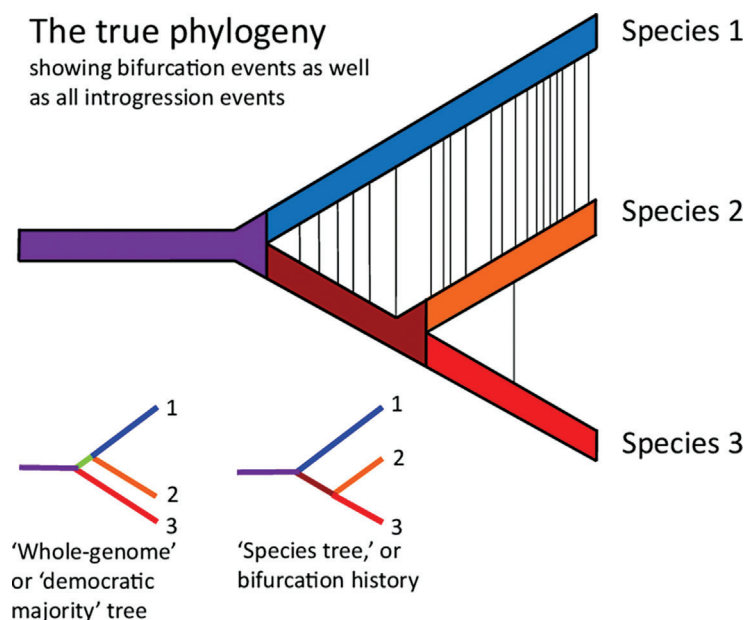
Figure 2. Phylogenetic discordance B/D mimicry region of *Heliconius* genomes. **A:** F_{ST} plot shows divergent *optix* regulatory region determining mimicry differences between geographic races within *H. melpomene*. Mimicry has been shown to have very strong adaptive value in *Heliconius*. **B:** The same region shows a strong excess of ABBA phylogenetic sites over BABA sites, implicating introgression between *H. melpomene* and *H. timareta*. **C:** Furthermore, the non-sister species *H. elevatus* shows a phylogenetic topology indicating introgression of the rayed mimicry pattern from the *melpomene-timareta* clade in the same genomic region. Modified and reprinted with permission from [17].

democratic majority of the genes to another topology. It is perhaps defensible to argue that the "democratic opinion" tree is more predictive of the origins of the genes, though it is marred by potential inferences of ancestral species (pale green) that never existed (Fig. 3). We instead favor the idea that the species tree is the bifurcation history (Fig. 3). This we would argue is closer to what we mean by the speciation history, in spite of the difficulty of its discovery, and acknowledging a lowered expectation of its predictiveness for the histories of its component genes.

Are species incompatible?

Another conclusion that arises from these findings is that large fractions of different species' genomes may in fact be compatible. The genomic distribution of "intrinsic" incompatibilities (such as "Dobzhansky-Muller incompatibilities" [94, 95]) is poorly known except in a few species [96]. In *Saccharomyces* yeasts, it is possible to replace whole chromosomes with little effect on viability, while in *Drosophila* many hybrid sterility loci seem scattered very widely across the genome [97–101]. It is possible that the situation in *Drosophila* is unusual, perhaps a result of "faster male" sexual selection that leads to

Think again

The true phylogeny
showing bifurcation events as well as all introgression events

Species 1

Species 2

Species 3

1
2
3
'Whole-genome' or 'democratic majority' tree

1
2
3
'Species tree,' or bifurcation history

Figure 3. A simple case where introgression can distort the history of species and speciation. By "the true phylogeny," in this paper, we mean the totality of true histories of every part of the genome. This is not readily depicted: our simplified cartoon of the true phylogeny network above indicates abundant introgression between species 1 and 2 after their bifurcation, but little between sister species 2 and 3. It does not, however, show which gene travels in which direction and when, all of which is surely important information about the "true phylogeny" as well. If introgression is extensive, the whole genome tree (bottom left) may indicate an incorrect bifurcation history, as well as ancestral species that never existed (such as the apparent ancestor of 1 and 2 in the diagram). The true bifurcation history of species is shown bottom right.

genome-wide effects on male hybrid sterility [102, 103]. Even though incompatibility loci have been mapped in crosses between *A. gambiae* and *A. arabiensis* [104], genomic evidence for very widespread homologous replacement between species in the autosomes of *Anopheles* and *Heliconius* [18, 73] suggests either that incompatibilities were not very common in those genomes, or that some introgressed alleles are advantageous enough to overcome initial incompatibility. Although autosomal genes introgress readily in both groups, the preponderance of "species tree" genealogies in the sex chromosome in the *Anopheles gambiae* complex [18] is likely due to multiple overlapping inversions that differ between *A. gambiae* + *coluzzii* and *A. arabiensis*. These inversions suppress recombination and so inhibit introgression of small chromosomal fragments on this chromosome. If adaptive alleles are widely available to introgress, determining the number and effect of incompatibilities will not be

adequate to assess the potential for introgression between species: we will also need to know the number and selective effects of these variants.

As far as is known, classical lateral transfer from distantly related species is not a major recent source of phylogenetic incongruence in multicellular eukaryotes, and most of the phylogenetic reticulation we observe is due to homologous exchange via hybridization. The selective advantages of sex within species remain contentious, but sex surely optimizes some balance between benefits and costs of recombination [105, 106]. Typically, hybridizing with another species is viewed as "the grossest blunder in sexual preference," and mate choice (reinforcement) is expected to evolve to limit hybridization among sympatric species [107]. However, given that hybridization does still occur, and sometimes leads to beneficial effects, we should now perhaps broaden our view of sex across the species boundary, where the same cost/benefit function is confronted by

individuals seeking sexual partners, albeit with different parameter values. If outcrossing within and between species is regulated by the same cost/benefit equation, a sexual selection process similar to reinforcement should apply to interactions within as well as between species.

Practical implications of introgression

The prevalence of laterally transferred antibiotic resistance genes among bacterial species is a well-known problem for human health [108, 109]. Similar problems might therefore be expected to result from introgression or lateral transfer among related eukaryotic pest and disease species. The African malaria-carrying mosquitoes provide some worrying examples. For example, rates of hybridization between *Anopheles gambiae* and *A. arabiensis* are only ~0.22% per generation [71]. However, because this introduces foreign alleles at a rate far higher than mutation, there are persistent concerns that insecticide resistance evolution in one species may lead to the rapid spread of that resistance to others via introgression [72]. Multiple cases of introgression of alleles encoding both organophosphate and pyrethroid insecticide resistance are certainly known between the sister species *A. gambiae* and *A. coluzzii* [110–113]; these two are known to hybridize and backcross much more frequently [63] than do *A. gambiae* with *A. arabiensis*. Similarly, large sibling species complexes of the black fly genus *Simulium* transmit river blindness in Africa and tropical America, and may also exchange genes. Among sympatric species of the African *S. damnosum* complex, hybridization rates may reach 0.1% per generation. Introgression is thought likely to explain the rapid spread of insecticide resistance among multiple *Simulium* species in Africa [114]. The same problem occurs even in vertebrate pests: a genomic region containing a rodenticide resistance allele spread via introgression between two partially interfertile mouse species in Western Europe [115].

Recent advances in genetics and genetic engineering are revolutionizing pest control, allowing for "designer

organisms" in agriculture and human health. Several major transgenic crops, especially those expressing herbicide or insect resistance, have been released in many countries. At the same time, new molecular marker and genomic analyses let us gather evidence on gene flow between crops and wild relatives for the first time. The results are clear: introgression does occur, and weedy relatives are acquiring novel genetic variation from crops, including transgenes that are liable to make these weeds more noxious [116].

The use of transgenic organisms is more advanced in agriculture than in human health. However, a variety of genetic control measures of vectors have been suggested and in some cases are being used to engineer disease vector populations [117]. For example, infection of *Aedes* mosquitoes by *Wolbachia* causes refractoriness to dengue virus proliferation [118], while *Wolbachia*-infected *Anopheles* mosquitoes show reduced *Plasmodium* infection [119]. In addition, there is the possibility of manipulating the genetics of mosquito innate immunity in order to reduce their efficiency as a vector [120]. Of these, probably the most successfully used cases so far are a number of releases of *Wolbachia*-infected *Aedes aegypti* to control dengue ([118] www.eliminatedengue.com). As with transgenic plants, because the transmission of genetic traits requires mating, these genetic traits may "leak" to related species via introgression. This may not have negative impacts, especially in comparison to the potentially positive benefits of the engineered trait on the target species. However, given genomic evidence for introgression of many other traits, its importance should not be underemphasized when seeking regulatory approval for release of genetically modified organisms (recognizing that *Wolbachia* infection does not technically qualify as a genetic modification to most regulatory bodies).

Conclusions and outlook

Our main conclusion is that many more species are likely exchanging genes than has been appreciated. It is not only sister species that hybridize and undergo genomic introgression: whole groups of rapidly radiating species may exchange adaptive as well as non-adaptive genomic regions, as in *Heliconius*, *Anopheles*, cichlids, *Xiphophorus*, Darwin's finches, horses, and hominins. In fact, because hybridization between sister species does not always affect the species tree – and because introgression between sister species is more likely – it may be that estimates of introgression rates from species tree topologies alone vastly underestimate the amount of gene flow occurring in nature. For many systems we may think we are able to infer a species tree signal, but we must recognize that this signal may only be represented by a small fraction of genes.

As well as causing problems for phylogenetics, abundant introgression and incomplete lineage sorting might greatly weaken inferences in comparative analysis. When we map character traits onto the tree of a rapidly radiating group, we should be cautious. For instance, the raptorial habit is thought to be ancestral to the entire core landbirds, but today it is present in several monophyletic groups, each more closely related to birds that have apparently lost the habit [79–81]. Alternatively, core landbirds may have been ancestrally non-raptorial, and a number of raptorial traits could have could have been shared at the base of these lineages by introgression among the early species of each lineage. This is not dissimilar to what we observe in mimicry patterns in *Heliconius* or in beak morphology of Darwin's finches, among species of radiations that we see hybridizing today [17, 74]. Similarly, inferences from phylogeography – such as geographic origins of rapidly radiating groups inferred from phylogenetic methods – should be affected as well. The origins of traits, and the genes that determine them can have very different histories from that of the species tree.

Acknowledgments

References

1. **Woese CR**, **Kandler O**, **Wheelis ML.** 1990. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc Natl Acad Sci USA* **87**: 4576–9.
2. **Woese CR**, **Fox GE.** 1977. Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc Natl Acad Sci USA* **74**: 5088–90.
3. **Doolittle WF.** 1999. Phylogenetic classification and the universal tree. *Science* **284**: 2124–9.
4. **Hilario E**, **Gogarten JP.** 1993. Horizontal transfer of ATPase genes – the tree of life becomes a net of life. *BioSyst* **31**: 111–9.
5. **Martin W.** 1999. Mosaic bacterial chromosomes: a challenge en route to a tree of genomes. *BioEssays* **21**: 99–104.
6. **Clark AG**, **Eisen MB**, **Smith DR**, **Bergman CM**, et al. 2007. Evolution of genes and genomes on the *Drosophila* phylogeny. *Nature* **450**: 203–18.
7. **Lindblad-Toh K**, **Garber M**, **Zuk O**, **Lin MF**, et al. 2011. A high-resolution map of human evolutionary constraint using 29 mammals. *Nature* **478**: 476–82.
8. **Dujon B**, **Sherman D**, **Fischer G**, **Durrens P**, et al. 2004. Genome evolution in yeasts. *Nature* **430**: 35–44.
9. **Maia VH**, **Gitzendanner MA**, **Soltis PS**, **Wong GK-S**, et al. 2014. Angiosperm phylogeny based on 18S/26S rDNA sequence data: constructing a large data set using next-generation sequence data. *Int J Plant Sci* **175**: 613–50.
10. **Lee EK**, **Cibrian-Jaramillo A**, **Kolokotronis S-O**, **Katari MS**, et al. 2011. A functional pylogenomic view of the seed plants. *PLoS Genet* **7**: e1002411.
11. **Marcussen T**, **Sandve SR**, **Heier L**, **Spannagl M**, et al. 2014. Ancient hybridizations among the ancestral genomes of bread wheat. *Science* **345**: 1250092.
12. **Brandvain Y**, **Kenney AM**, **Flagel L**, **Coop G**, et al. 2014. Speciation and introgression between *Mimulus nasutus* and *Mimulus guttatus*. *PLoS Genet* **10**: e1004410.
13. **McCormack JE**, **Faircloth BC**, **Crawford NG**, **Gowaty PA**, et al. 2012. Ultraconserved elements are novel phylogenomic markers that resolve placental mammal phylogeny when combined with species-tree analysis. *Genome Res* **22**: 746–54.
14. **Yang Y**, **Moore MJ**, **Brockington SF**, **Soltis DE**, et al. 2015. Dissecting molecular evolution in the highly diverse plant clade Caryophyllales using transcriptome sequencing. *Mol Biol Evol* **32**: 2001–14.
15. **Mayr E.** 1963. *Animal Species and Evolution*. Cambridge, MA: Harvard University Press.
16. **Anderson E**, **Stebbins GL.** 1954. Hybridization as an evolutionary stimulus. *Evolution* **8**: 378–88.

17. **Heliconius Genome Consortium.** 2012. Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. *Nature* **487**: 94–8.

18. **Fontaine MC, Pease JB, Steele A, Waterhouse RM**, et al. 2015. Extensive introgression in a malaria vector species complex revealed by phylogenomics. *Science* **347**: 1258524.

19. **Zawadzki P, Roberts MS, Cohan FM.** 1995. The log-linear relationship between sexual isolation and sequence divergence in *Bacillus* is robust. *Genetics* **140**: 917–32.

20. **Majewski J, Zawadzki P, Pickerill P, Cohan FM**, et al. 2000. Barriers to genetic exchange between bacterial species: *Streptococcus pneumoniae* transformation. *J Bacteriol* **182**: 1016–23.

21. **Eppley JM, Tyson GW, Getz WM, Banfield JF.** 2007. Genetic exchange across a species boundary in the archaeal genus *Ferroplasma*. *Genetics* **177**: 407–16.

22. **Naor A, Lapierre P, Mevarech M, Papke RT**, et al. 2012. Low species barriers in halophilic Archaea and the formation of recombinant hybrids. *Curr Biol* **22**: 1444–8.

23. **Gourbière S, Mallet J.** 2010. Are species real? The shape of the species boundary with exponential failure, reinforcement, and the "missing snowball". *Evolution* **64**: 1–24.

24. **Morales L, Dujon B.** 2012. Evolutionary role of interspecies hybridization and genetic exchanges in yeasts. *Microbiol Mol Biol Rev* **76**: 721–39.

25. **Woese CR.** 2002. On the evolution of cells. *Proc Natl Acad Sci USA* **99**: 8742–7.

26. **Creevey CJ, Doerks T, Fitzpatrick DA, Raes J**, et al. 2011. Universally distributed single-copy genes indicate a constant rate of horizontal transfer. *PLoS ONE* **6**: e22099.

27. **Rivera MC, Lake JA.** 2004. The ring of life provides evidence for a genome fusion origin of eukaryotes. *Nature* **431**: 152–5.

28. **McInerney JO, O'Connell MJ, Pisani D.** 2014. The hybrid nature of the Eukaryota and a consilient view of life on Earth. *Nat Rev Microbiol* **12**: 449–55.

29. **Puigbò P, Wolf YI, Koonin EV.** 2009. Search for a 'Tree of Life' in the thicket of the phylogenetic forest. *J Biol* **8**: 1–17.

30. **Ciccarelli FD, Doerks T, von Mering C, Creevey CJ**, et al. 2006. Toward automatic reconstruction of a highly resolved tree of life. *Science* **311**: 1283–7.

31. **Lang JM, Darling AE, Eisen JA.** 2013. Phylogeny of bacterial and archaeal genomes using conserved genes: supertrees and supermatrices. *PLoS ONE* **8**: e62510.

32. **Dagan T, Martin W.** 2006. The tree of one percent. *Genome Biol* **7**: 118.

33. **Lawrence JG, Retchless AC.** 2010. The myth of bacterial species and speciation. *Biol Philos* **25**: 569–88.

34. **Mindell DP.** 2013. The Tree of Life: metaphor, model, and heuristic device. *Syst Biol* **62**: 479–89.

35. **Huang J.** 2013. Horizontal gene transfer in eukaryotes: the weak-link model. *BioEssays* **35**: 868–75.

36. **Keeling PJ, Palmer JD.** 2008. Horizontal gene transfer in eukaryotic evolution. *Nat Rev Genet* **9**: 605–18.

37. **Crisp A, Boschetti C, Perry M, Tunnacliffe A**, et al. 2015. Expression of multiple horizontally acquired genes is a hallmark of both vertebrate and invertebrate genomes. *Genome Biol* **16**: 50.

38. **Gladyshev EA, Meselson M, Arkhipova IR.** 2008. Massive horizontal gene transfer in bdelloid rotifers. *Science* **320**: 1210–3.

39. **Boschetti C, Carr A, Crisp A, Eyres I**, et al. 2012. Biochemical diversification through foreign gene expression in bdelloid rotifers. *PLoS Genet* **8**: e1003035.

40. **Wu B, Buljic A, Hao W.** 2015. Extensive horizontal transfer and homologous recombination generate highly chimeric mitochondrial genomes in yeast. *Mol Biol Evol* **32**: 2559–70.

41. **Sela N, Kim E, Ast G.** 2010. The role of transposable elements in the evolution of non-mammalian vertebrates and invertebrates. *Genome Biol* **11**: R59.

42. **Kidwell MG.** 2002. Transposable elements and the evolution of genome size in eukaryotes. *Genetica* **115**: 49–63.

43. **Lynch M.** 2007. *The Origins of Genome Architecture.* Sunderland, MA: Sinauer Associates.

44. **Lynch M.** 2006. The origins of eukaryotic gene structure. *Mol Biol Evol* **23**: 450–68.

45. **Rogozin IB, Carmel L, Csuros M, Koonin EV.** 2012. Origin and evolution of spliceosomal introns. *Biol Direct* **7**: 11.

46. **Schaack S, Gilbert C, Feschotte C.** 2010. Promiscuous DNA: horizontal transfer of transposable elements and why it matters for eukaryotic evolution. *Trends Ecol Evol* **25**: 537–46.

47. **Burgess R, Yang Z.** 2008. Estimation of hominoid ancestral population sizes under Bayesian coalescent models incorporating mutation rate variation and sequencing errors. *Mol Biol Evol* **25**: 1979–94.

48. **Dutheil JY, Munch K, Nam K, Mailund T**, et al. 2015. Strong selective sweeps on the X chromosome in the human-chimpanzee ancestor explain its low divergence. *PLoS Genet* **11**: e1005451.

49. **Wiuf C, Zhao K, Innan H, Nordborg M.** 2004. The probability and chromosomal extent of trans-specific polymorphism. *Genetics* **168**: 2363–72.

50. **Teixeira JC, de Filippo C, Weihmann A, Meneu JR**, et al. 2015. Long-term balancing selection in *LAD1* maintains a missense trans-species polymorphism in humans, chimpanzees, and bonobos. *Mol Biol Evol* **32**: 1186–96.

51. **Saitou N, Yamamoto F.** 1997. Evolution of primate ABO blood group genes and their homologous genes. *Mol Biol Evol* **14**: 399–411.

52. **Degnan JH, Rosenberg NA.** 2006. Discordance of species trees with their most likely gene trees. *PLoS Genet* **2**: e68.

53. **Yang Z.** 2014. *Molecular Evolution. A Statistical Approach.* Oxford: Oxford University Press.

54. **Yang Z, Rannala B.** 2014. Unguided species delimitation using DNA sequence data from multiple loci. *Mol Biol Evol* **31**: 3125–35.

55. **Degnan JH, Rosenberg NA.** 2009. Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends Ecol Evol* **24**: 332–40.

56. **Grant PR, Grant BR.** 1992. Hybridization of bird species. *Science* **256**: 193–7.

57. **Mallet J.** 2005. Hybridization as an invasion of the genome. *Trends Ecol Evol* **20**: 229–37.

58. **Arnold ML, Sapir Y, Martin NH.** 2008. Genetic exchange and the origin of adaptations: prokaryotes to primates. *Philos Trans R Soc Lond B Biol Sci* **363**: 2813–20.

59. **Mallet J, Beltrán M, Neukirchen W, Linares M.** 2007. Natural hybridization in heliconiine butterflies: the species boundary as a continuum. *BMC Evol Biol* **7**: 28.

60. **Bacilieri R, Ducousso A, Petit RJ, Kremer A.** 1996. Mating system and asymmetric hybridization in a mixed stand of European oaks. *Evolution* **50**: 900–8.

61. **Grant PR.** 1993. Hybridization of Darwin's finches on Isla Daphne Major, Galápagos. *Philos Trans R Soc Lond B Biol Sci* **340**: 127–39.

62. **Mallet J, McMillan WO, Jiggins CD.** 1998. Estimating the mating behavior of a pair of hybridizing *Heliconius* species in the wild. *Evolution* **52**: 503–10.

63. **Lee Y, Marsden CD, Norris LC, Collier TC**, et al. 2013. Spatiotemporal dynamics of gene flow and hybrid fitness between the M and S forms of the malaria mosquito, *Anopheles gambiae. Proc Natl Acad Sci USA* **110**: 19854–9.

64. **Coyne JA, Orr HA.** 1997. "Patterns of speciation in *Drosophila*" revisited. *Evolution* **51**: 295–303.

65. **Price TD, Bouvier MM.** 2002. The evolution of F1 postzygotic incompatibilities in birds. *Evolution* **56**: 2083–9.

66. **Moyle LC, Nakazato T.** 2010. Hybrid incompatibility "snowballs" between *Solanum* species. *Science* **329**: 1521–3.

67. **Moyle LC, Olson MS, Tiffin P, Baum D.** 2004. Patterns of reproductive isolation in three angiosperm genera. *Evolution* **58**: 1195–208.

68. **Anderson E.** 1948. Hybridization of the habitat. *Evolution* **2**: 1–9.

69. **Leaché AD, Harris RB, Rannala B, Yang Z.** 2013. The influence of gene flow on species tree estimation: a simulation study. *Syst Biol* **63**: 17–30.

70. **Coluzzi M, Sabatini A, Petrarca V, Di Deco MA.** 1979. Chromosomal differentiation and adaptation to human environments in the *Anopheles gambiae* complex. *Trans R Soc Trop Med Hyg* **73**: 483–97.

71. **Weetman D, Steen K, Rippon EJ, Mawejje HD**, et al. 2014. Contemporary gene flow between wild *An. gambiae s.s.* and *An. arabiensis. Parasites Vectors* **7**: 345.

72. **Mawejje HD, Wilding CS, Rippon EJ, Hughes A**, et al. 2013. Insecticide resistance monitoring of field-collected *Anopheles gambiae s.l.* populations from Jinja, eastern Uganda, identifies high levels of pyrethroid resistance. *Med Vet Entomol* **27**: 276–83.

73. **Martin SH, Dasmahapatra KK, Nadeau NJ, Salazar C**, et al. 2013. Genome-wide evidence for speciation with gene flow in *Heliconius* butterflies. *Genome Res* **23**: 1817–28.

74. **Lamichhaney S, Berglund J, Almen MS, Maqbool K**, et al. 2015. Evolution of Darwin's finches and their beaks revealed by genome sequencing. *Nature* **518**: 371–5.

75. **Brawand D, Wagner CE, Li YI, Malinsky M**, et al. 2014. The genomic substrate for adaptive radiation in African cichlid fish. *Nature* **513**: 375–81.

76. **Jónsson H, Schubert M, Seguin-Orlando A, Ginolhac A**, et al. 2014. Speciation with gene flow in equids despite extensive

chromosomal plasticity. *Proc Natl Acad Sci USA* **111**: 18655–60.

77. **Cui R**, **Schumer M**, **Kruesi K**, **Walter R**, et al. 2013. Phylogenomics reveals extensive reticulate evolution in *Xiphophorus* fishes. *Evolution* **67**: 2166–79.

78. **Green RE**, **Krause J**, **Briggs AW**, **Maricic T**, et al. 2010. A draft sequence of the Neandertal genome. *Science* **328**: 710–22.

79. **Jarvis ED**, **Mirarab S**, **Aberer AJ**, **Li B**, et al. 2014. Whole-genome analyses resolve early branches in the tree of life of modern birds. *Science* **346**: 1320–31.

80. **Suh A**, **Smeds L**, **Ellegren H.** 2015. The dynamics of incomplete lineage sorting across the ancient adaptive radiation of neoavian birds. *PLoS Biol* **13**: e1002224.

81. **Prum RO**, **Berv JS**, **Dornburg A**, **Field DJ**, et al. 2015. A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA sequencing. *Nature* **526**: 569–73.

82. **Whelan NV**, **Kocot KM**, **Moroz LL**, **Halanych KM.** 2015. Error, signal, and the placement of Ctenophora sister to all other animals. *Proc Natl Acad Sci USA* **112**: 5773–8.

83. **Hedrick PW.** 2013. Adaptive introgression in animals: examples and comparison to new mutation and standing variation as sources of adaptive variation. *Mol Ecol* **22**: 4606–18.

84. **Mallet J.** 2007. Hybrid speciation. *Nature* **446**: 279–83.

85. **Rieseberg LH.** 1997. Hybrid origins of plant species. *Annu Rev Ecol Syst* **28**: 359–89.

86. **Seehausen O.** 2004. Hybridization and adaptive radiation. *Trends Ecol Evol* **19**: 198–207.

87. **Stankowski S**, **Streisfeld MA.** 2015. Introgressive hybridization facilitates adaptive divergence in a recent radiation of monkeyflowers. *Proc Biol Sci* **282**: 20151666.

88. **Grant PR**, **Grant BR.** 2014. Synergism of natural selection and introgression in the origin of a new species. *Am Naturalist* **183**: 671–81.

89. **Mallet J.** 2013. Species, concepts of. In Levin SA, ed; *Encyclopedia of Biodiversity*, Second Edition, Volume 6. Waltham, MA: Academic Press. p. 679–91.

90. **Behm JE**, **Ives AR**, **Boughman JW.** 2010. Breakdown in postmating isolation and the collapse of a species pair through hybridization. *Am Naturalist* **175**: 11–26.

91. **Maddison WP.** 1997. Gene trees in species trees. *Syst Biol* **46**: 523–36.

92. **Liu L**, **Yu L**, **Kubatko L**, **Pearl DK**, et al. 2009. Coalescent methods for estimating phylogenetic trees. *Mol Phylogen Evol* **53**: 320–8.

93. **Edwards SV.** 2009. Is a new and general theory of molecular systematics emerging? *Evolution* **63**: 1–19.

94. **Orr HA**, **Turelli M.** 2001. The evolution of postzygotic isolation: accumulating Dobzhansky-Muller incompatibilities. *Evolution* **55**: 1085–94.

95. **Coyne JA**, **Orr HA.** 2004. *Speciation*. Sunderland, MA: Sinauer Associates.

96. **Orr HA**, **Madden LD**, **Coyne JA**, **Goodwin R**, et al. 1997. The developmental genetics of hybrid inviability: a mitotic defect in *Drosophila* hybrids. *Genetics* **145**: 1031–40.

97. **Greig D.** 2007. A screen for recessive speciation genes expressed in the gametes of F1 hybrid yeast. *PLoS Genet* **3**: 281–6.

98. **Greig D.** 2008. Reproductive isolation in *Saccharomyces*. *Heredity* **102**: 39–44.

99. **True JR**, **Weir BS**, **Laurie CC.** 1996. A genome-wide survey of hybrid incompatibility factors by the introgression of marked segments of *Drosophila mauritiana* chromosomes into *Drosophila simulans*. *Genetics* **142**: 819–37.

100. **Hollocher H**, **Wu CI.** 1996. The genetics of reproductive isolation in the *Drosophila simulans* clade: X vs. autosomal effects and male vs. female effects. *Genetics* **143**: 1243–55.

101. **Masly JP**, **Presgraves DC.** 2007. High-resolution genome-wide dissection of the two rules of speciation in *Drosophila*. *PLoS Biol* **5**: e243.

102. **Wu CI**, **Palopoli MF**, **Johnson NA.** 1996. Haldane's rule and its legacy: why are there so many sterile males? *Trends Ecol Evol* **11**: 281–4.

103. **Chapman T**, **Partridge L.** 1996. Sexual conflict as fuel for evolution. *Nature* **381**: 189–90.

104. **Slotman M**, **della Torre A**, **Powell JR.** 2004. The genetics of inviability and male sterility in hybrids between *Anopheles gambiae* and *An. arabiensis*. *Genetics* **167**: 275–87.

105. **Maynard Smith J.** 1978. *The Evolution of Sex*. Cambridge: Cambridge University Press.

106. **Otto SP.** 2009. The evolutionary enigma of sex. *Am Naturalist* **174**: S1–S14.

107. **Fisher RA.** 1930. *The Genetical Theory of Natural Selection*. Oxford: Clarendon Press.

108. **Davies J**, **Davies D.** 2010. Origins and evolution of antibiotic resistance. *Microbiol Mol Biol Rev* **74**: 417–33.

109. **Finley RL**, **Collignon P**, **Larsson DGJ**, **McEwen SA**, et al. 2013. The scourge of antibiotic resistance: the important role of the environment. *Clin Infect Dis* **57**: 704–10.

110. **Djogbénou L**, **Chandre F**, **Berthomieu A**, **Dabiré R**, et al. 2008. Evidence of introgression of the *ace-1ᴿ* mutation and of the *ace-1* duplication in West African *Anopheles gambiae* s.s. *PLoS ONE* **3**: e2172.

111. **Weill M**, **Chandre F**, **Brengues C**, **Manguin S**, et al. 2000. The *kdr* mutation occurs in the Mopti form of *Anopheles gambiae s.s.* through introgression. *Insect Mol Biol* **9**: 451–5.

112. **Norris LC**, **Main BJ**, **Lee Y**, **Collier TC**, et al. 2015. Adaptive introgression in an African malaria mosquito coincident with the increased usage of insecticide-treated bed nets. *Proc Natl Acad Sci USA* **112**: 815–20.

113. **Clarkson CS**, **Weetman D**, **Essandoh J**, **Yawson AE**, et al. 2014. Adaptive introgression between *Anopheles* sibling species eliminates a major genomic island but not reproductive isolation. *Nat Commun* **5**: 4248.

114. **Adler PH**, **Cheke RA**, **Post RJ.** 2010. Evolution, epidemiology, and population genetics of black flies (Diptera: Simuliidae). *Infect Genet Evol* **10**: 846–65.

115. **Song Y**, **Endepols S**, **Klemann N**, **Richter D**, et al. 2011. Adaptive introgression of anticoagulant rodent poison resistance by hybridization between Old World mice. *Curr Biol* **21**: 1296–301.

116. **Ellstrand NC**, **Meirmans P**, **Rong J**, **Bartsch D**, et al. 2013. Introgression of crop alleles into wild or weedy populations. *Annu Rev Ecol Evol Syst* **44**: 325–45.

117. **Alphey L.** 2014. Genetic control of mosquitoes. *Annu Rev Entomol* **59**: 205–24.

118. **Hoffmann AA**, **Montgomery BL**, **Popovici J**, **Iturbe-Ormaetxe I**, et al. 2011. Successful establishment of *Wolbachia* in *Aedes* populations to suppress dengue transmission. *Nature* **476**: 454–7.

119. **Bian G**, **Joshi D**, **Dong Y**, **Lu P**, et al. 2013. *Wolbachia* invades *Anopheles stephensi* populations and induces refractoriness to *Plasmodium* infection. *Science* **340**: 748–51.

120. **Pike AD**, **Cirimotich CM**, **Dimopoulos G.** 2013. Impact of transgenic immune deployment on mosquito fitness. In Koenraadt S, Takken W, eds; *Ecology and Control of Vector-Borne Diseases, Volume 3, Ecology of Parasite-Vector Interactions*. Wageningen: Academic Publishers. p. 19–33.

Think again