Alternative views of biological species: reproductively isolated units or genotypic clusters?

Abstract

Wang et al. have recently outlined their thoughts on two ideas about species: the biological species concept (BSC), and the genic view of speciation. Although both are based on reproductive isolation, the genic view, they argue, is likely preferred due to the possibility of allowing considerable gene flow among species, which is arguably what recent genomic data shows. These data, however, mostly failed to distinguish between the BSC and the genic view according to Wang et al., because it could not be ruled out that the observed introgression occurred early in speciation, when both models allow gene flow. I propose that the lack of resolution in the authors' debate is chiefly due to the difficulty of deciding when speciation is "complete" under both views tested. I agree with Wang et al. that the study of reproductive isolation is worthwhile in order to understand speciation, but I prefer to use a simpler, third criterion for speciation: the acquisition of genetic differences that allow persistence of distinguishable populations in spite of geographic overlap and the potential for continued gene flow. Under this multilocus "genotypic cluster" view, gene flow may take place at any time after species are recognized, and we do not have to decide whether gene flow is early or late in the speciation process. I detail recent genomic evidence from *Anopheles* mosquitoes and *Heliconius* butterflies showing that such "leaky" species seem to be able to coexist in spite of massive levels of introgression, often among non-sister species that show hybrid sterility in one sex.

Introduction

Wang et al. in this journal argue that it may be time to abandon a classic idea about species, the biological species concept (BSC), given recent findings with genomic data on closely related taxa (1). Their paper also proposes a set of tests on genetic or genomic data that might lead to acceptance or rejection of the biological species concept.

Wang et al. are not the first to critique the biological species concept (BSC), which has survived an onslaught of attacks from many fronts. What is the biological species concept? And what is the new critique about? In this commentary, I briefly cover recent history of thoughts about species, discuss current work that depends on rich, genome-scale sequence data, before attempting to resolve the issues.

The biological species concept

The term "biological species concept" (BSC) was coined originally by Ernst Mayr, building on Dobzhansky's ideas of reproductive isolation (2-4) during the later phases of the "Modern Synthesis" between Darwinian evolution and Mendelian genetics. Dobzhansky and Mayr argued that species are best viewed as reproductively isolated units: "Species are groups of actually or potentially interbreeding natural populations, which are reproductively isolated from other such groups" (3, 4). Reproductive isolation was viewed as a set of "mechanisms" that prevent species from merging: "each species is a delicately integrated genetic system ... Hybridization would lead to a breakdown of this system and would result in the production of disharmonious types. It is the function of the isolating mechanisms to prevent such a breakdown, and to protect the integrity of the genetic system of species." (ref. 4, p. 109).

Associated with these views there exists a related set of ideas that species are more "real" in some way than lower taxa such as geographic races, or higher taxa such as genera: "The evolutionary significance of species is now quite clear. ... The species are the real units of evolution, as the temporary incarnation of harmonious, well-integrated gene complexes" (ref. 4, p. 621). Under this new, mid-20th Century view of species, speciation was a special kind of evolution distinct from other modes of evolution at the population level, and Darwin was criticized as having not really dealt with speciation at all in his book "On the Origin of Species..." (5). According to Ernst Mayr, "Darwin's failure... resulted to a large extent from a misunderstanding of the true nature of species." (ref. 4, p. 14). Many people today still largely agree with these criticisms of Darwin's ideas about speciation (6, 7).

Although the resultant BSC is a very successful and useful-seeming meme, the idea that species are best defined as reproductively isolated units was not without critics. An underlying theme was that the sharp reality built into the BSC conflicts with the idea of speciation as a gradual, emergent Darwinian process. Mayr clearly recognized this: in fact, he believed that gene flow and the lack of reproductive isolation among populations hindered evolutionary progress and prevented speciation. In his view, speciation was "complete" when there was a complete absence of successful gene flow (see also ref. (8)). In a sense, by requiring almost complete reproductive isolation between species, and a complete lack of it within species, speciation became an almost insurmountable barrier for gradual Darwinian evolution to bridge.

Mayr's solution to this difficulty was the proposition that populations would normally diverge into separate species while geographically isolated, or in "allopatry" (3). However, gradual vicariant divergence was acknowledged to be rather slow, maybe too slow to generate the diversity of species we observe on our planet today. To solve the issue of the slow rate of divergence demanded by his view of species as optimized "delicately integrated genetic systems," Mayr popularized another idea, that small populations dispersing to geographically separated locations could speciate more rapidly by means of the "founder effect", leading to a "genetic revolution" via rapid reorganization of coadapted gene complexes (4, 9). Today the founder effect is deemed poorly supported by theory or data, although the more general idea that allopatric speciation is the norm still holds some sway (6, 10-12).

Problems with the Biological Species Concept

A major issue with the BSC concerns its supposed role in evolution. Ehrlich and Raven (13) pointed out that local populations, rather than species, were the actual foci of evolution, and that whether or not such populations belonged to a widely-distributed biological species had little to do with their local evolution, since gene flow was evident mostly on a local, rather than a global scale within widespread biological species. Instead, "selection itself is both the primary cohesive and disruptive force in evolution; the selective regime determines what influence gene flow has on observed patterns of differentiation. Populations will differentiate if they are subjected to different selective forces and will tend to remain similar if they are not" (13). Gene flow undoubtedly constrains divergent natural selection, but only on a very local scale. Populations will diverge on very local spatial scales > $\sigma/\sqrt{s}$, where σ represents gene flow distance, and s represents the strength of divergent selection on a gene. Therefore, populations will diverge, even under relatively weak selection (say s ~ 0.01), at spatial scales of a only handful of gene flow distances (14, 15).

Because the BSC seems to depend on a rather binary view of reproductive isolation (either on: yes -- a species; or off: no -- not a species), it is not operational in taxonomy, especially for far-flung local populations not in contact, where under Mayr's rubric, one would need to assess whether the populations were "potentially interbreeding." Sokal and Crovello (16) "having decided that the BSC is neither operational nor heuristic nor of any practical value," concluded that a more operational "phenetic species as normally described is the desirable species concept to be associated with the taxonomic category 'species,' and that the localized biological population may be the most useful unit for evolutionary study." Their argument for "phenetic" species was a statisticians' version of Darwin's own argument that species were best defined as clusters of individuals separated from other such groups by gaps in the distribution of morphologies (e.g. ref. 5: pp. 58-59). Like Ehrlich and Raven, Sokal & Crovello (16) recognized instead local populations, rather than species, as the main focus for evolution.

Others found different problems. Oak species had been long recognized, but many hybrids were also known to exist naturally. In spite of gene flow, distinct ecologically associated lineages, or species, of oak trees remain identifiable, presumably due to divergent ecological niches. Ecological divergence, rather than reproductive isolation, defined such species (17).

Hugh Paterson pointed out that protection of species from introgression as a function of isolating mechanisms was an unlikely and anti-Darwinian idea: natural selection tends to safeguard the individual's interest, not its population or species. Species, Paterson suggested, should be defined by mate recognition within species, and not protection from invasion by other species (18, 19). While recognizing that the idea that species functioned to protect species integrity under Mayr's BSC was a group selectionist idea, it is arguable that "mate recognition" is not a function of sexual behaviour at the level of species as Paterson supposed; instead, individuals mate to maximize their own individual fitness, and so the idea of mate recognition at the species level is almost as prone to a group selection fallacy as reproductive isolation.

If species evolved from geographic populations, then they passed through earlier phases deemed by Ernst Mayr and others under the BSC to be geographical subspecies. However, many named subspecies have unique trait combinations. Cracraft suggested that these "tips" of the phylogenetic tree should instead be classified as separate species (20). In his view, species are populations with consistently different character combinations. Any diagnosable population could be considered as a separate species under Cracraft's phylogenetic species concept (PSC), regardless of reproductive compatibility, provided that the members of such populations showed "a parental pattern of ancestry and descent."

The fact that many nominal species, especially in plants such as oak trees, engaged in frequent hybridization ("too much sex") and that other species were asexual ("too little sex"), and yet were still classified as species led to the idea that factors other than reproductive isolation combined to cause the "cohesion" of species, in Templeton's "cohesion species concept." Ecological factors and even genealogical relationships that help maintain separate species status were included as mechanisms of cohesion in addition to classical reproductive isolation (21, 22). And there were many other critiques, especially a variety of other conflicting species concepts based on phylogenetic principles, for example a species concept based on gene genealogies (23).

In spite of all these criticisms, many of today's evolutionary biologists still maintain that the BSC is a reasonable definition of species. Of course, "reproductive isolation," today often called "reproductive barriers," does in some sense play a role in keeping taxa that we call species apart. Therefore, to understand speciation in sexual species seems to require investigating something like reproductive isolation so that a pair of taxa can diverge into separate species that can coexist spatially. Thus the BSC has remained one of the more enduring ideas about species, with many modern texts finding it more useful for discussions of evolution and speciation than alternatives (6, 12, 24). The study of speciation therefore appears to be equivalent to a goal of understanding the evolution of reproductive isolation. "Speciation genes" that cause sterility and inviability in hybrids between species are still seen as important in speciation (1, 8, 25, 26).

Species as genotypic clusters, 25 years on

My group's studies on "host races" in the larch budmoth and hybridization among species of *Heliconius* butterflies (27-31) led me, as with Van Valen and Templeton earlier, to a view that species were often distinguishable in spite of continued gene flow. Therefore, while a sort of reproductive isolation was indeed one of the *causes* of speciation, it was not very helpful as a *definition* of species. Attempts have been made to justify a weaker version of the BSC than Mayr's (e.g. ref. 6), but it is unclear how much gene flow would be allowed under this idea. Instead of Templeton's idea of including as "cohesion mechanisms" all possible diverse processes that prevent species from fusing in the definition of species, I argued that we could more simply define what we were talking about operationally, using the pattern of observable traits and genetics of those taxa that did not fuse .

Having realized that species evolve rather than were created, it was obvious to Darwin that species were going to be less "real" than hitherto assumed. What Darwin meant by species in "The Origin of Species" were populations separable because of morphological gaps between them. In population genetics, the equivalent is that we recognize species when there is a deficit of heterozygotes at loci that differ, or more powerfully, when there is a deficit of recombinants among such loci. In modern terms, species are recognized operationally by the gaps between their genetic traits (deficits of intermediates and recombinants) when they overlap in space and time. This Darwinian definition of species is agnostic to the precise processes, such as reproductive isolation or ecology that led to or maintain separateness.

My own paper arguing for a multi-locus "genotypic cluster" definition of species, incorporating these Darwinian ideas, has its 25th anniversary this year (32). Others working on the tractable host-plant races of the apple maggot *Rhagoletis pomonella* had come to similar conclusions at the same time (33). The idea of species as genotypic clusters derived from a viewpoint similar to that of Sokal & Crovello (16), but with a more explicit population genetics focus, for example on linkage disequilibrium (the correlation between differentiated alleles within individual genomes). I called this "a species definition for the Modern Synthesis," because I was concerned to justify existing usage of the taxonomic species, which I knew well from the Lepidoptera I studied, and which seemed usefully to employ a polytypic species delimitation method: geographic subspecies that blended freely in contact zones would be included within species. In contrast, many of the local subspecies that I study would have been divided into multiple separate species under Cracraft's (20) "diagnostic" version of the phylogenetic species concept (PSC). My goal was to ensure that divergent sympatric clusters of genotypes would be classified as separate species, and to suggest operational methods (e.g. studying the genetics of clusters of genotypes and phenotypes in sympatry), rather than to employ the of inferential lip service to reproductive isolation apparently required by the BSC (32).

The genic view of speciation

Meanwhile, Chung-I Wu and his group had been investigating "speciation genes," the loci that contribute to hybrid sterility and inviability, in *Drosophila*. As short-fragment DNA sequencing became accessible due to advances in molecular genetics, Wu's group found that a speciation gene, *Odysseus* (*OdsH*), had a genealogy among *D. mauritiana*, *D. sechellia*, and *D. simulans*, that matched a likely speciation tree. However, loci as little as 1.8 kb distant from this gene yielded a different genealogy (i.e. gene tree), suggesting that gene flow beyond species boundaries could have occurred after acquisition of *Odysseus*-related reproductive isolation (34). These findings led Wu (35) to the idea that "speciation is the stage where the gene pools at loci of differential adaptation (i.e. genes that cause speciation) would not mix even when the extrinsic barriers to gene exchange are removed and, furthermore, will be able to continue to diverge. Thus, the very essence does not have to include reproductive isolation."

Wu saw reproductive isolation and speciation as processes that happened to genes, rather than to the whole genome as in Mayr's view. His "genic view of speciation" therefore allowed parts of the genome not involved in divergent selection or hybrid incompatibility to flow between species, even after an irreversible stage of speciation had been reached (35, 36). In that this genic view allowed gene flow among species, this view of species was similar to Mallet (32), except that the "genic view of speciation" stressed that the means of divergence was reproductive isolation due to "speciation genes." In contrast, Mallet's genotypic cluster criterion is agnostic as to means of divergence (32).

The views of Wang et al. in 2020

In their new critique, Wang et al. (1) regard the BSC as "demanding a prolonged period of divergence during which there is no gene flow." The "salient feature" of the BSC, according to Wang et al. is that "a period of strict geographical isolation (or allopatry) ... is needed to complete the process of speciation" (1). Although Mayr and others have regarded speciation as largely occurring via allopatry, and although, as I have pointed out above, allopatry is perhaps more useful for speciation to the BSC than under many alternative species concepts, I think most supporters of the BSC did not mean to preclude the possibility of sympatric or parapatric speciation via their species concept.

Wang et al. then argue that early in the process of speciation, during the evolution of reproductive isolation, there is of course some gene flow, even under the BSC model (1). But subsequently, the BSC allows for no introgression whatsoever (Part A of their Fig. 1). In contrast, under a species concept closer to Wu's genic view of speciation (part B of their Fig. 1), "gene flow continues, possibly diminished, all the way to the completion of speciation when traits of reproductive isolation have evolved."

To test their genic view, the authors carry out simulations of genomes affected by gene flow, with a few "speciation genes" under divergent selection, but separated by many neutral loci that can be recombined away from these. They show that under some combinations of parameters, it is possible to generate genomes in which genomic regions near the speciation genes remain diverged, while other intervening loci recombine relatively freely among the species, and get swamped by introgression. This is reminiscent of the "islands of speciation" idea, as found between *Anopheles gambiae* and *A. coluzzii* (formerly called the M and S strains of *A. gambiae*) (37).

There are now abundant genomic data on the question of continuing gene flow for many of the taxa that we call species. The authors argue that the existing data may allow us to test "between the genomic [i.e. BSC] and genic view of speciation." In the classical BSC view, the authors argue, gene flow is allowed in the very early stages of speciation, when geographic isolation is incomplete, but thereafter a prolonged period of geographic isolation is needed to "complete" speciation. In contrast, under the genic view of speciation, allopatry is not required, and gene flow can continue right up to "completion" of speciation, and indeed maybe long afterwards, until finally, reproductive isolation is complete (i.e. gene flow is negligible).

The authors review existing literature for gene flow among named species, and while acknowledging its existence, they say "we find no case of large scale introgression in late stages of speciation, when postmating reproductive isolation is evident." Instead, all or at least most existing examples of introgression are considered "early stage events." To take an example, they review the data on the *Anopheles gambiae* group. As well as there being abundant evidence for genomic islands of speciation and gene flow between *A. gambiae* and *A. coluzzii* (37, 38), more recent genomic data have shown that gene flow between the *A. arabiensis* and *A. gambiae* lineages has affected ~98% of the genome, causing the average gene tree today to differ from the inferred original branching order of species (39-41). The view that the average gene tree in the *Anopheles gambiae* group of mosquitos is distorted away from the species tree is agreed on by the three sets of authors, in spite of the fact that the precise inferred species tree differs among all of these three papers based on the same data! Nonetheless, Wang et al. argue that "it is evident that the introgressions in these species are early-stage events" (1), and therefore that these observations do not constitute proof against the BSC.

An attempt to resolve the issues

As we have seen, there is a lot to agree with in the "genic view of speciation" and its critique of the BSC, as shown by the obvious overlap with my own view of species as "genotypic clusters" between which gene flow is possible. However, the "genic view" as propounded here suffers, in my view, from a reiteration of some major problems associated with the BSC. For example, the current article repeatedly mentions the "completion of speciation." While in the authors' stated view speciation need not necessarily preclude a trickle of continued gene flow (35), "complete speciation" in the new paper seems to mean total cessation of gene flow that would affect the whole genome (1). Arguably, in nature, speciation is not like that. There is in fact a continuum of gene flow that declines but continues, at least in some parts of the genome, long after taxonomists and evolutionary biologists generally recognize species. In the simplest "light-bulb failure" model, there is a linear, albeit stochastic accumulation of incompatible substitutions with time, and species compatibility declines exponentially (42). If epistasis is involved, then compatibility decline may speed up over time, compared to exponential decline, as in the "snowball" model (8, 43). However, the compatibility predicted under all such idealized probabilistic models never actually reaches zero (42).

Empirical results appear to bear this out. The hybrid between horse and donkey, the mule, has long held an almost iconic status as a completely sterile hybrid between species, and at least since Buffon (44) has supported the view that species are real entities that lack the capacity to cross. However, mules do occasionally have offspring, and this has been known at least since the 18th Century (45, 46). This potentially leads to occasional gene flow between horses and mules and vice-versa. Today, we can examine their genomes, which show a strong signal of introgression between the horse lineage and the donkey + zebra lineages of *Equus* (47).

Heliconius butterflies provide further examples: occasional gene flow continues at every step of the way between *H. cydno/timareta* and *H. melpomene*, in spite of a lack of "islands of divergence" in the genome between these two lineages. While male hybrids are fertile, female hybrids are usually sterile, an example of Haldane's Rule, so that there is strong evidence of "postmating isolation" between these two lineages (48). However, we have now found a pair of species in widespread sympatry, *H. pardalinus* and *H. elevatus*, that show evidence for substantial introgression, leading to large patches of near panmixia across the genome, with occasional genomic islands of divergence presumably containing divergently selected speciation genes (49), very similar to the pattern shown by Turner et al. for *Anopheles* "islands of speciation" (37).

As with *Anopheles*, we also find evidence for massive gene flow deep in the estimated species tree of the *Heliconius erato* group of species, and this again causes distortion of the average gene tree away from the presumed species tree (50). Wang et al. (1) suggest that "speciation genes" that resist introgression may not yet have developed in this clade, which they argue means that this gene flow is taking place early in speciation (their Table 1). But contrary to Wang et al., the existence of frequent Haldane's Rule postzygotic hybrid sterility in *Heliconius* (49, 51) demonstrates considerable "postmating isolation." Furthermore, positive correlation of introgression probability with recombination rate along the *Heliconius* genome in this study suggests that there are actually a very large number of "postmating isolation" loci that select against introgression, and that these loci are widely scattered across all 21 chromosomes (50).

The "genic view" of Wang et al. (1) versus the BSC therefore seems very hard to test, at least using the reasoning suggested in their paper. In contrast, under a more operational genotypic cluster approach, species are usefully defined as able to maintain their distinctness at multiple unlinked loci (not, of course, at all loci, so allowing for introgression of some parts of the genome) while in sympatry with sister species. Today, a test of sympatric species status could be carried out using multiple genetic markers and a Bayesian assignment test such as STRUCTURE (52, 53). It is even arguable that Dobzhansky and Mayr may have meant something like this when they proposed the BSC and coined "reproductive isolating mechanisms" as the key. In the 1930s-1950s the unavailability of genetic markers would have led them to propose a heuristic, "reproductive isolation" instead of, say, strong linkage disequilibria among unlinked markers in sympatry, when defining species.

If we instead adopt a sympatric genotypic cluster criterion, speciation can occur much earlier than in the "completeness" view of speciation espoused by Wang et al., and we no longer have the problem about whether the gene flow is early or late during the "process of speciation." It may occur at any time after speciation, albeit with diminishing probability over time since the split between a pair of species.

Conclusions

The species concept engendered by Wu's "genic view of speciation" (35) originated in order to accommodate the possibility of gene flow after speciation, while also recognizing that reproductive isolation was involved in speciation. However, Wang et al. (1) show that it is difficult to test genomic data under this view against the alternative view of the biological species concept. In my view, the reason for the difficulty of their test is that it is hard to know when speciation is "complete" under either of these views. Under my own genotypic cluster view of species, the completeness of speciation doesn't matter. I argue that what we call species in sexual taxa are merely populations that can overlap spatially without fusing (and this was the original intent of the biological species concept).

Under all three views outlined above, reproductive isolation and divergent selection are probable causes of speciation in sexual species worth investigating.

However, it would in my view be helpful to avoid including these processes in the definition of the taxa we call species. It is becoming especially difficult to know what supposedly "reproductively isolated species" are when we know, as in both *Anopheles* and *Heliconius*, that gene flow is so extensive as to re-write the history of their genes towards an entirely different topology than the original species tree. It is more helpful to separate cause and effect, and therefore, to focus on maintenance of divergence of at least some loci in sympatry as the most useful criterion of species.

References

1. X. Wang, Z. He, S. Shi, C.-I. Wu, Genes and speciation - is it time to abandon the Biological Species Concept? *National Science Review* **xxx**, nwz220 (2020).
2. T. Dobzhansky, *Genetics and the Origin of Species* (Columbia Univ. Press, New York, 1937), pp. -.
3. E. Mayr, *Systematics and the Origin of Species* (Columbia Univ. Press, New York, 1942), pp. -.
4. E. Mayr, *Animal Species and Evolution* (Harvard University Press, Cambridge, Mass., 1963), pp. -.
5. C. R. Darwin, *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life* (John Murray, London, 1859), pp. 1-x.
6. J. A. Coyne, H. A. Orr, *Speciation* (Sinauer Associates, Sunderland, Mass., 2004), pp. 1-xiii+545.
7. A. Berry, H. E. Hoekstra, Teaching the Origin. *Current Biology* **19**, R9- (2009).
8. H. A. Orr, M. Turelli, The evolution of postzygotic isolation: accumulating Dobzhansky-Muller incompatibilities. *Evolution* **55**, 1085-1094 (2001).
9. E. Mayr, "Change of genetic environment and evolution" in Evolution as a Process*,* J. Huxley, A. C. Hardy, E. B. Ford, Eds. (Allen and Unwin, London, 1954), pp. 157-180.
10. N. H. Barton, B. Charlesworth, Genetic revolutions, founder effects, and speciation. *Annu. Rev. Ecol. Syst.* **15**, 133-164 (1984).
11. M. Turelli, N. H. Barton, J. A. Coyne, Theory and speciation. *Trends in Ecology and Evolution* **16**, 330-343 (2001).
12. D. I. Bolnick, B. M. Fitzpatrick, Sympatric speciation: models and empirical evidence. *Annual Review of Ecology, Evolution, and Systematics* **38**, 459-487 (2007).
13. P. R. Ehrlich, P. H. Raven, Differentiation of populations. *Science* **165**, 1228-1232 (1969).
14. M. Slatkin, Gene flow and selection in a cline. *Genetics* **75**, 733-756 (1973).
15. N. H. Barton, The dynamics of hybrid zones. *Heredity* **43**, 341-359 (1979).
16. R. R. Sokal, T. J. Crovello, The biological species concept: a critical evaluation. *American Naturalist* **104**, 107-123 (1970).
17. L. Van Valen, Ecological species, multispecies, and oaks. *Taxon* **25**, 233-239 (1976).
18. H. E. H. Paterson, The continuing search for the unknown and unknowable: a critique of contemporary ideas on speciation. *S. Afr. J. Sci.* **77**, 113-119 (1981).
19. H. E. H. Paterson, "The recognition concept of species" in Species and Speciation*,* E. S. Vrba, Ed. (Transvaal Museum, Pretoria, 1985), pp. 21-29.
20. J. Cracraft, "Speciation and its ontology: the empirical consequences of alternative species concepts for understanding patterns and processes of differentiation"*,* D. Otte, J. A. Endler, Eds. (Sinauer Associates, Sunderland, Mass., 1989), pp. 28-59.

21. A. R. Templeton, "The meaning of species and speciation: a genetic perspective" in Speciation and its Consequences, D. Otte, J. A. Endler, Eds. (Sinauer Associates, Sunderland, Mass., USA, 1989), pp. 3-27.
22. A. R. Templeton, "Species and speciation. Geography, population structure, ecology, and gene trees" in Endless Forms. Species and Speciation, D. J. Howard, S. H. Berlocher, Eds. (Oxford Univ. Press, New York, 1998), pp. 32-43.
23. D. A. Baum, K. L. Shaw, "Genealogical perspectives on the species problem", P. C. Hoch, Ed. (Missouri Botanical Garden, St. Louis, Missouri, 1995), pp. 289-303.
24. D. Futuyma, M. Kirkpatrick, *Evolution. Fourth Edition* (Sinauer, Sunderland, Mass., 2017), pp. xx + 700.
25. H. A. Orr, The genetic basis of reproductive isolation: insights from *Drosophila*. *Proc. Natl. Acad. Sci. USA* **102**, 6522-6526 (2005).
26. H. A. Orr, Dobzhansky, Bateson, and the genetics of speciation. *Genetics* **144**, 1331-1335 (1996).
27. I. Emelianov, J. Mallet, W. Baltensweiler, Genetic differentiation in the larch budmoth *Zeiraphera diniana* (Lepidoptera: Tortricidae): polymorphism, host races or sibling species? *Heredity* **75**, 416-424 (1995).
28. I. Emelianov, M. Drès, W. Baltensweiler, J. Mallet, Host-induced assortative mating in host races of the larch budmoth. *Evolution* **55**, 2002-2010 (2001).
29. I. Emelianov, F. Marec, J. Mallet, Genomic evidence for divergence with gene flow in host races of the larch budmoth. *Proc. Roy. Soc. Lond. B Biol. Sci.* **271**, 97-105 (2004).
30. W. O. McMillan, C. D. Jiggins, J. Mallet, What initiates speciation in passion-vine butterflies? *Proc. Natl. Acad. Sci. USA* **94**, 8628-8633 (1997).
31. J. Mallet, M. Beltrán, W. Neukirchen, M. Linares, Natural hybridization in heliconiine butterflies: the species boundary as a continuum. *BMC Evolutionary Biology* **7**, 28 (2007).
32. J. Mallet, A species definition for the Modern Synthesis. *Trends in Ecology and Evolution* **10**, 294-299 (1995).
33. J. L. Feder, "The apple maggot fly, *Rhagoletis pomonella*: flies in the face of conventional wisdom" in Endless Forms. Species and Speciation, D. J. Howard, Ed. (Oxford University Press, New York, 1998), pp. 130-144.
34. C. T. Ting, S. C. Tsaur, C. I. Wu, The phylogeny of closely related species as revealed by the genealogy of a speciation gene, *Odysseus*. *Proc. Natl. Acad. Sci. USA* **97**, 5313-5316 (2000).
35. C. I. Wu, The genic view of the process of speciation. *J. Evol. Biol.* **14**, 851-865 (2001).
36. C. I. Wu, C. T. Ting, Genes and speciation. *Nature Reviews Genetics* **5**, 1-9 (2004).
37. T. L. Turner, M. W. Hahn, S. V. Nuzhdin, Genomic islands of speciation in *Anopheles gambiae*. *PLoS Biology* **3**, e285 (2005).
38. Y. Lee *et al.*, 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**, 19854-19859 (2013).
39. M. C. Fontaine *et al.*, Extensive introgression in a malaria vector species complex revealed by phylogenomics. *Science* **347**, 1258524 (2015).
40. D. Wen, Y. Yu, M. W. Hahn, L. Nakhleh, Reticulate evolutionary history and extensive introgression in mosquito species revealed by phylogenetic network analysis. *Mol. Ecol.* **25**, 2361-2372 (2016).
41. Y. Thawornwattana, D. Dalquen, Z. Yang, Coalescent analysis of phylogenomic data confidently resolves the species relationships in the *Anopheles gambiae* species complex. *Mol. Biol. Evol.* **35**, 2512–2527 (2018).
42. S. Gourbière, J. Mallet, Are species real? The shape of the species boundary with exponential failure, reinforcement, and the "missing snowball". *Evolution* **64**, 1-24 (2010).

43. H. A. Orr, The population genetics of speciation: the evolution of hybrid incompatibilities. *Genetics* **139**, 1805-1813 (1995).
44. C. d. Buffon, Georges-Louis Leclerc, "Description de la partie du Cabinet qui a rapport à l'histoire naturelle du cheval. L'asne" in Histoire Naturelle, Générale et Particuliére, avec la Description du Cabinet du Roy. (Paris, 1753), vol. 4, pp. 377-403.
45. J. Hunter, Observations tending to shew that the wolf, jackal, and dog, are all of the same species. *Philosophical Transactions of the Royal Society of London* **77**, 253-266 (1787).
46. C. Lyell, *Principles of Geology, Being an Inquiry how far the Former Changes of the Earth's Surface Are Referable to Causes now in Operation. Volume 2. 5th Edition* (John Murray, London, 1837), pp. ii + 442.
47. H. Jónsson *et al.*, Speciation with gene flow in equids despite extensive chromosomal plasticity. *Proc. Natl. Acad. Sci. USA* **111**, 18655-18660 (2014).
48. R. E. Naisbit, C. D. Jiggins, M. Linares, J. Mallet, Hybrid sterility, Haldane's rule, and speciation in *Heliconius cydno* and *H. melpomene*. *Genetics* **161**, 1517-1526 (2002).
49. N. Rosser *et al.*, Geographic contrasts between pre- and postzygotic barriers are consistent with reinforcement in Heliconius butterflies. *Evolution* **73**, 1821-1838 (2019).
50. N. B. Edelman *et al.*, Genomic architecture and introgression shape a butterfly radiation. *Science* **366**, 594-599 (2019).
51. R. E. Naisbit, C. D. Jiggins, J. Mallet, Disruptive sexual selection against hybrids contributes to speciation between *Heliconius cydno* and *H. melpomene*. *Proc. Roy. Soc. Lond. B Biol. Sci.* **268**, 1849-1854 (2001).
52. D. Falush, M. Stephens, J. K. Pritchard, Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Mol. Ecol. Notes* **7**, 574-578 (2007).
53. J. Mallet, "Species, concepts of" in Encyclopedia of Biodiversity, Second Edition. Volume 6*,* S. A. Levin, Ed. (Academic Press, Waltham, Mass., 2013), pp. 679-691.