This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier’s archiving and manuscript policies are encouraged to visit:

http://www.elsevier.com/copyright
Reinventing species selection with molecular phylogenies

Daniel L. Rabosky1,2 and Amy R. McCune1

1 Department of Ecology and Evolutionary Biology, Cornell University, Ithaca, NY 14853, USA
2 Fuller Evolutionary Biology Program, Cornell Laboratory of Ornithology, 159 Sapsucker Woods Road, Ithaca, NY 14850, USA

Species selection as a potential driver of macroevolutionary trends has been relegated to a largely philosophical position in modern evolutionary biology. Fundamentally, species selection is the outcome of heritable differences in speciation and extinction rates among lineages when the causal basis of those rate differences can be decoupled from genotypic (within-population) fitnesses. Here, we discuss the rapidly growing literature on variation in species diversification rates as inferred from molecular phylogenies. We argue that modern studies of diversification rates demonstrate that species selection is an important process influencing both the evolution of biological diversity and distributions of phenotypic traits within higher taxa. Explicit recognition of multi-level selection refocuses our attention on the mechanisms by which traits influence speciation and extinction rates.

Selection can occur at multiple levels
Natural selection is one of the most important ideas in the history of science. Charles Darwin, writing in the original edition of the Origin, defined natural selection as ‘the preservation of favorable variations and the rejection of injurious variation’ [1]. Although the mechanisms of heredity would remain unknown for decades, Darwin went on to state elsewhere that these ‘favorable variations’ must be passed from parent to offspring for evolution to occur. A modern definition of selection is ‘differential survival and reproduction of replicating entities based on heritable trait differences, which, as a statement about process, is fundamentally the same as that proposed by Darwin.

Identifying the levels at which selection acts has been fraught with controversy, and one of the longer-standing controversies in evolutionary biology has been whether selection can occur at the species level [2]. To the outsider, the literature on species selection is confusing and rich in jargon that frequently seems irrelevant to the pertinent evolutionary issues at hand. Here we clarify the concept of species selection and explain why accepting the central premise of species selection is fundamental to understanding large-scale patterns of biological diversity. We believe that a focus on the distinction between species and organismic selection forces us to contrast the determinants of genotypic, within-population fitness with the causes of differential rates of species origination and extinction. Simply stated, factors promoting diversification are not necessarily the same factors as those maximizing lifetime reproduction. We argue that evidence from molecular phylogenies provides an important demonstration of species selection influencing both patterns of species richness and trait distributions across a range of phylogenetic scales.

What species selection is and is not
Species selection [3] is the outcome of heritable variation in speciation and extinction rates among taxa [4,5]. Such variation in diversification thus reflects differences in the emergent fitness (see glossary) of species [6,7], where emergent fitness is the differential survival and reproduction of species that cannot be extrapolated from analyses of survival and reproduction of lower levels in the biological hierarchy (such as individuals, social groups, or alleles). Traits under species selection cause an alteration of the net diversification rate of a lineage (Figure 1b, c), where the net diversification rate is the difference between the speciation rate and the extinction rate. A particular trait might cause an increase in the rate of speciation, but this need not be related to the fitness of individuals within populations. Even if a trait simultaneously increases fitness at the individual level (by increasing individual survival and

Glossary

Aggregate trait: a trait shared by all individuals within a species. Examples include individual coloration, body size, and timing of sexual reproduction.

Emergent fitness: heritable differences in net diversification rates among species constitute variation in ‘emergent fitness.’ Such differential diversification is an emergent property of species because it generally cannot be extrapolated from analyses of survival and reproduction at the individual level.

Emergent trait: a trait that exists only as a collective property of individuals within a species and that cannot be expressed as a summation of individual trait values. Examples include geographic range size, sex ratio, and intraspecific variability.

Extrinsic influences on diversification: influence rates of speciation and extinction but are not linked to heritable aggregate or emergent species traits. Climate and environmental energy are examples of extrinsic factors that can affect speciation and extinction rates.

Heritability (species level): occurs when traits or diversification rates are inherited across speciation events. Similarity between ancestor and descendant species reflects species-level heritability if they are not simply a result of shared environmental influences.

Intrinsic influence on diversification: occurs when rates of speciation and extinction are influenced by heritable species traits.

Net diversification rate: the net rate of species proliferation of a lineage, equal to the rate of lineage origination (speciation) minus lineage extinction.

Organismic selection: natural selection that results from differential survival and reproduction of individuals within a population.

Sister clades: a pair of clades resulting from a single speciation event in a phylogenetic tree. Sister clades are by definition the same age.

Species selection: results from heritable differences in net diversification rates among species, when the causal basis of those rate differences is not a simple function of differential survival and reproduction at the individual level.

Corresponding author: Rabosky, D.L. (DLR28@cornell.edu).
reproduction) and the species level (by increasing the speciation rate), the causal mechanisms underlying fitness at these two levels will be different [8,9], because the means by which a trait maximizes lifetime reproduction will differ from the means by which the same trait influences diversification.

One example of a trait that seems to be favored by both organismic and species selection is floral symmetry. Within populations, plants with more symmetric flowers appear to have greater reproductive success than less symmetric individuals, because floral symmetry facilitates pollination efficiency and reduces pollen waste [10,11]. However, such symmetry is also associated with greater pollinator specificity, which in turn facilitates rapid species divergence owing to the potential for pollinator isolation among geographically isolated populations; plant clades with symmetric flowers have diversified at greater rates than those with more asymmetric flowers [12]. Selection at the individual level presumably led to the evolution of more symmetric flowers within populations, which then served as the raw variation for species selection. The net outcome is that floral symmetry has become more frequent both within populations (organismic selection on pollination efficiency) and within clades (species selection on pollinator specificity).

Some have previously argued that species selection can only operate on emergent characters [13], such as geographic range and population genetic structure [5,14], where emergent characters are species traits that cannot be reduced to individual traits. Although there is consensus that differential speciation and extinction owing to emergent traits is species selection [14], the situation has traditionally been more contentious for aggregate traits [6,9,15], which are individual-level traits shared by all individuals within a species (see glossary). However, we have already argued that an aggregate trait can influence both lifetime reproduction (natural selection) as well as rates of speciation and extinction (species selection):
floral symmetry is an aggregate trait and it increases lifetime reproduction by reducing pollen fitness. As a simple example, imagine that all individual organisms above a certain threshold body size have low absolute fitness, and that all genotypes within a particular species result in body sizes above the threshold. Individuals thus die because they possess a particular trait value (large body size), and the fact that all individuals die results in the extinction of the species. In this case extinction is arguably due only to organismic selection. However, such a clear case of organismic selection leading to species extinction is only possible for a very limited set of scenarios; specifically, those where extinction is a direct effect of selection on individual organisms. For such a process to drive a trend, traits under selection at the individual level would have to arise repeatedly, become fixed within populations, and then drive the species to extinction by reducing the absolute fitness of all genotypes to below replacement (R < 1). The evidence that such pure organismal selection directly mediates extinction dynamics is far from convincing. Aggregate traits like body size are usually correlated with emergent traits such as species abundance and geographic range size [61,65,79], for which selection can only occur at the species level. Moreover, aggregate traits can even be favored at the individual level while at the same time increasing extinction probability [80]. For example, organismic selection for resource specialization might incidentally reduce the total population size of a species and thus increase the probability of extinction via chance demographic fluctuations. This would be a clear case of species selection because of the potential for conflict between individual and species level processes [2].

floral symmetry is an aggregate trait and it increases lifetime reproduction by reducing pollen waste and increases diversification by a different mechanism, increasing pollinator specificity.

We would not argue that organismic selection is unrelated to species selection. Selection at the individual level contributes to trait variation between species by transforming intraspecific variation into species differences that might result in species selection (Figure 1: a to b, or a to c). However, the mechanism by which a trait becomes fixed within a species, whether through selection or drift, need not be the same as the mechanism by which the trait influences diversification. Moreover, traits can be favored by selection at the individual level while exerting a negative influence on diversification rates, leading to a decline in trait frequency within the clade (Figure 1a, c).

To summarize: understanding large-scale evolutionary patterns requires that we understand both the mechanism by which traits cause differential species diversification as well as differential reproduction of individuals. Even a perfect understanding of the dynamics of selection within populations can be uninformative with respect to the trait’s effect on speciation rates [4], and we would argue that this is often true for extinction rates as well (Box 1). If a trait causes a higher rate of speciation in lineages possessing the trait, a greater percentage of species through time will possess that trait (Figure 1d). Species selection is an important concept because it forces us to recognize that traits can influence patterns of speciation and extinction independent of their effects on survival and reproduction at the individual level (Figure 1: a vs. b and c). We cannot even begin to address questions of why some groups of organisms are so much more diverse than other groups without accepting that selection can be decoupled at the individual and species level, because there is no necessary reason why traits that promote survival and reproduction of individuals need correspond to those that influence speciation and extinction. Furthermore, even if the same trait influences both individual fitness within populations and diversification rates (e.g. floral symmetry), the mechanisms by which traits increase in frequency within populations and clades can be entirely distinct.

**Box 1. Extinction, individual death, and species selection**

Some have argued that selective extinction does not necessarily reflect species selection [24]. It is true that if a species becomes extinct because of low absolute fitness of all genotypes within the species, then extinction is reducibly specified by genotypic (within-population) fitness. As a simple example, imagine that all individual organisms above a certain threshold body size have low absolute fitness, and that all genotypes within a particular species result in body sizes above the threshold. Individuals thus die because they possess a particular trait value (large body size), and the fact that all individuals die results in the extinction of the species. In this case extinction is arguably due only to organismic selection. However, such a clear case of organismic selection leading to species extinction is only possible for a very limited set of scenarios; specifically, those where extinction is a direct effect of selection on individual organisms. For such a process to drive a trend, traits under selection at the individual level would have to arise repeatedly, become fixed within populations, and then drive the species to extinction by reducing the absolute fitness of all genotypes to below replacement (R < 1). The evidence that such pure organismal selection directly mediates extinction dynamics is far from convincing. Aggregate traits like body size are usually correlated with emergent traits such as species abundance and geographic range size [61,65,79], for which selection can only occur at the species level. Moreover, aggregate traits can even be favored at the individual level while at the same time increasing extinction probability [80]. For example, organismic selection for resource specialization might incidentally reduce the total population size of a species and thus increase the probability of extinction via chance demographic fluctuations. This would be a clear case of species selection because of the potential for conflict between individual and species level processes [2].

Molecular phylogenies provide a growing body of evidence for species selection

Although it is not widely acknowledged, questions about variation in diversification rates among clades are fundamentally questions about species selection and the broader role of the process in generating biological diversity [7,14,16]. Species selection has in fact been the focus of an ever-expanding research program in comparative biology for nearly two decades, fueled by the availability of both molecular phylogenies and rigorous statistical methods for diversification analysis. Our perspective is that species selection can serve as a bridge between the paleobiological and molecular phylogenetic literature [7,14], where researchers have been generating complementary insights into the diversification process from very different types of data. A renewed focus on this frequently mischaracterized concept should bring the paleobiological literature to the attention of researchers who study diversification rates using molecular phylogenies of extant taxa.

At least two patterns should immediately indicate the possibility of species selection. The first is when speciation and extinction rates show heritable differences among lineages [17], where heritability is the tendency of progeny species to resemble their ‘parents’. The second involves any repeatable effect of a trait on the rate of diversification of the species possessing the traits (Figure 2). Molecular phylogenies provide a rich source of material with which to assess these patterns. Rates of species diversification can be reconstructed from time-calibrated molecular phylogenies [18–20], and even trees with topological information only can be used to identify shifts in diversification as well as correlations between traits and diversification rates [21].

A substantial literature has developed around methods for testing the relationship between traits and diversification rates [22–25], but formal tests for the heritability of diversification rates have been limited [17,26]. This might suggest that few tests for heritability of diversification rates have been performed. However, all studies of which we are aware that have documented a relationship between a trait and diversification rate have implicitly demonstrated heritability of rates, whether this relationship was stated or not. Most traits show a tendency to be inherited across speciation events; were this not the case, there would be no need for phylogenetic independent contrasts and other comparative methods. If a trait is associated with increased diversification rates, and the trait is heritable in this sense, diversification rates must themselves be heritable.
Many phylogenetic trees show evidence for dramatic differences in species richness among even closely related clades, a pattern that can be caused by species selection. This is especially true for large phylogenetic trees, which frequently contain a heterogeneous mix of clades with both high and low diversification rates [26–29]. Apparent also at larger phylogenetic scales are the profound effects of geographic area, environmental energy, and other ecological covariates on patterns of diversification [30–33]. Extrinsic environmental factors might cause increased rates of diversification [34] or influence ecological limits on clade growth [19,31,35], but differential diversification mediated solely by such ecological factors would not constitute species selection, because variation in speciation and extinction is not attributable to heritable trait differences among species.

Yet these extrinsic influences on diversification are frequently entangled with lineage-specific factors that can promote speciation. This is one explanation for the observation that different groups inhabiting the same geographic region can show dramatically different patterns of diversification. For example, Hawaii’s honeycreeper finches have undergone a spectacular evolutionary radiation, whereas the remaining four passerine bird lineages in the islands have not [36]. These finches did not arrive earlier on the islands than other groups, and their high rate of diversification has been attributed to evolutionary lability of feeding morphology [36]. Similar arguments have been made to explain the radiation of Darwin’s finches and the failure of other clades on the Galapagos to radiate [37] as well as the radiations of cichlids, but not other fishes, in the African rift lakes [38]. In Australia, a single clade of arid-adapted scincid lizards has diversified into nearly 200 species and is far more diverse than all other Australian skink clades [39], even though several clades that did not radiate occupy similar ecological settings. This indicates that intrinsic traits, perhaps associated with thermal physiology, have been a key factor in the explosive diversification of the group. Clearly, we have much to learn about the intrinsic factors that promote diversification in these and other groups.

Studies of species selection are most informative when the causal basis of differential diversification can be identified (Figure 2). A variety of traits are known to result in differential diversification and have been reviewed at length in several recent papers [7,14,40]. Among the first examples of traits to show putative correlations with diversification rates were those associated with sexual selection [41]. Under this model, species that are prone to rapid divergence in traits that influence mating success will, all else being equal, be more likely to undergo speciation than those showing low potential for divergence in such traits. This is probably because of the fact that traits under sexual selection serve an incidental role in species recognition: geographically isolated populations that diverge in sexually selected traits might no longer recognize each other as potential mates following secondary contact [42]. Many studies have reported positive effects of such traits on diversification rates, in groups that include agamid lizards [43], birds [44,45], flies [46], and fishes [47]. The evidence from birds is mixed, with other studies reporting no relationship between sexual selection and diversification [48,49]. However, Price argues that the lack of effect is perhaps attributable to the fact that sexual selection is so widespread within birds [42], indicating that other factors might be more important influences on species richness.

In addition to floral symmetry, a number of other traits associated with pollinator specificity have been shown to correlate with diversification rates. For example, lineages
with nectar spurs have higher diversification rates than unspurred lineages, across multiple clades of plants [50,51]. Nectar spurs restrict the pool of possible pollinator species and are hypothesized to facilitate more rapid evolution of reproductive isolation among geographically isolated populations. However, at the individual level, nectar spurs seem to evolve because of selection for improved pollen transfer [52]; their effects on speciation rates are incidental with respect to their consequences for within-population fitness. Other floral characters associated with pollination biology seem to show similar effects, including shifts from abiotic to biotic (animal) agents [51,53].

We do not intend to be exhaustive, but numerous other traits show positive (or negative) effects on diversification rates, ranging from behavioral flexibility in birds [54] to the evolution of a climbing habit in plants [55]. Whereas much of our discussion above has emphasized discrete traits, many continuous traits are associated with differential diversification, including rates of molecular evolution [56,57], population density in accipitriform hawks [44], and limb complexity in the Crustacea [58]. Of course, another class of traits might mediate differences in species richness without directly influencing rates of speciation or extinction. For example, some traits might constitute key innovations that enable lineages to occupy and radiate within a novel adaptive zone [59], leading to the rapid increase in species diversity that is characteristic of adaptive radiation [60]. Such traits might increase the effective ‘carrying capacity’ of a clade without directly influencing diversification rates [35].

**Challenges in the study of species selection**

The central challenges faced by these and related studies are fourfold. First, almost all ‘traits’ are in fact highly correlated with a suite of traits that confound simple interpretations of patterns. This is amply demonstrated by traits associated with extinction risk in extant species [61], which can be so inseparably bound with other traits that it is difficult to assign causation. However, much progress has been made recently in teasing apart factors that impact emergent fitness at the species level (by influencing speciation and extinction rates) from those correlated traits that are not explicit foci of selection. For example, Jablonski [62] reported that geographic range size in Cretaceous mollusks influenced species survivorship and was thus under species selection, but some researchers proposed that this pattern could result from selection on correlated traits, such as mode of larval development [63]. Jablonski and Hunt [64] used generalized linear models to tease apart these potentially confounding factors and demonstrated that geographic range is the dominant factor influencing species survivorship and shows significant species-level heritability even after controlling for larval mode. In another recent study, Cooper et al. [65] tested whether extinction risk in frogs is predicted by geographic range size and/or life history traits. Their findings indicate that range size is the primary determinant of extinction risk in extant species and raise the intriguing possibility that evolutionary trends in organismic traits such as body size and fecundity could be driven by species selection on range size.

Second, because numerous traits can influence diversification rates, traits can oppose one another at a given level of the selection hierarchy. The complexity of these interactions might explain, in part, why traits show contrasting effects on diversification among even closely related groups of higher taxa [66,67].

A third concern is strictly methodological: with some simple methods of analysis, it can be difficult to distinguish between the repeated effects of a trait on diversification and asymmetric rates of character change [68]. In particular, simple sister clade contrasts (Figure 2) can lead one to infer that diversification rates differ with respect to a particular trait, when the pattern is solely attributable to the characters themselves evolving at different rates. If one character state evolves frequently, but reversals to the previous state occur rarely, then simple contrasts will reveal an apparent association between that character state and the diversification rate. This is perhaps less problematic for newer methods that use all of the information in phylogenetic trees [22,23,69,70], including branch length data, but researchers need to bear this potential pitfall in mind when interpreting diversification analyses.

Finally, that we can ascribe a given pattern of diversification to a particular trait tells us nothing about the actual mechanisms by which the trait influences speciation and/or extinction dynamics. We argue in favor of a species selectionist viewpoint precisely because we feel that this level of analysis focuses attention on identifying these mechanisms.

Whereas we have strong reasons to believe that the above examples represent species selection, we still do not know in most cases whether the effects are driven by variation in speciation or extinction rates. Disentangling these processes is an important step towards understanding causation. Because speciation and extinction leave different signatures in phylogenies of extant taxa only [71], it is sometimes possible to determine whether increased diversification rates are mediated by increased speciation, decreased extinction, or both [20]. The binary state speciation–extinction model (BiSSE) of Maddison et al. [69] has considerable promise in this regard and enables simultaneous estimates of speciation, extinction, and transition rates for a binary character on a molecular phylogenetic tree.

Also promising is the use of theory to derive a priori predictions for the effects of traits on individual and species fitness. This approach has been used to explore the evolution of specialist versus generalist pollination [72] as well as the evolution of sexually dimorphic floral displays [73]. Experimental studies of the evolution of reproductive isolation can also help refine interpretations of higher-level patterns. For example, Arnqvist et al. [74] reported increased diversification rates in clades of insects where females mate with multiple males relative to those where females mate only once, indicating that postmating sexual selection might have caused differential diversification. However, several recent experimental studies on *Drosophila* found no evidence that manipulation of postmating sexual selection led to more rapid evolution of reproductive isolation [75,76]. This indicates that postmating sexual
selection influences extinction but not speciation rates, or alternatively that other correlated traits have caused differential diversification in insect clades with contrasting mating systems.

Prospects and conclusions

Whether species selection can and does occur is, in our minds, an issue settled. However, many issues remain, in addition to those outlined above. We consider the following to be of paramount importance. First, how much of the variance in species richness among clades can be explained by species selection relative to extrinsic influences on diversification? These explanations are not mutually exclusive and there is clear evidence for interactions between intrinsic and extrinsic factors [14,36,39,40]. The answer to this question is almost certainly complex, but at present, our understanding of these factors and their interactions is deeply unsatisfying. Second, could species selection drive large-scale trends in organismal design? It is possible that some of the most striking trends in evolution are attributable in part to species selection, such as the evolution of phenotypic complexity [58] and even the capacity of species to evolve [77]. Finally, how are human activities interacting with species traits to influence extinction rates? Species traits with high heritability currently explain much of the variation in extinction risk among threatened species [61,78], leading Jablonski to propose that we are unwittingly conducting a massive experiment in species selection [14]. An explicit recognition of selection at the species level is central to our ability to answer these and other questions, because it focuses our attention on the actual mechanisms by which traits influence the dynamics of speciation and extinction.

Acknowledgements

We thank D. Jablonski, P. Buston, R. Harrison, A. Agrawal, K. Wagner, I. Lovette, B. Cramer, and the BK discussion group for discussion of these topics and/or comments on the manuscript. This research was supported by NSF-OSIE-0612855 and NSF-DEB-0614277.

References

1 Darwin, C. (1859) On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life, John Murray