

Feature Review

Is there room for punctuated equilibrium in macroevolution?

Matthew W. Pennell, Luke J. Harmon, and Josef C. Uyeda

Department of Biological Sciences and Institute for Bioinformatics and Evolutionary Studies, University of Idaho, Moscow, ID 83844, USA

The long-controversial theory of punctuated equilibrium (PE) asserts that speciation causes rapid evolution against a backdrop of stasis. PE is currently undergoing a resurgence driven by new developments in statistical methods. However, we argue that PE is actually a tangle of four unnecessarily conflated questions: (i) is evolution gradualistic or pulsed? (ii) does trait evolution occur mainly at speciation or within a lineage? (iii) are changes at speciation adaptive or neutral? and (iv) how important is species selection in shaping patterns of diversity? We discuss progress towards answering these four questions but argue that combining these conceptually distinct ideas under the single framework of PE is distracting and confusing, and more likely to hinder progress than to spur it.

The resurgence of punctuated equilibrium

The following three quotations were all drawn from abstracts of recent papers purporting to use statistical models to empirically evaluate punctuated equilibrium (PE):

‘A long-standing debate in evolutionary biology concerns whether species diverge gradually through time or by punctuational episodes at the time of speciation. We found that approximately 22% of substitutional changes at the DNA level can be attributed to punctuational evolution, and the remainder accumulates from background gradual divergence’ ([1], p. 119).

‘This controversy, widely known as the ‘punctuated equilibrium’ debate, remained unresolved, largely owing to the difficulty of distinguishing biological species from fossil remains. We analyzed body masses of 2143 existing mammal species on a phylogeny comprising 4510 (i.e., nearly all) extant species to estimate rates of gradual (anagenetic) and speciational (cladogenetic) evolution’ ([2], p. 2195).

Corresponding author: Pennell, M.W. (mwpennell@gmail.com).

Keywords: punctuated equilibrium; macroevolution; comparative methods; phylogenetics; paleobiology.

0169-5347/\$ – see front matter

© 2013 Elsevier Ltd. All rights reserved. <http://dx.doi.org/10.1016/j.tree.2013.07.004>

‘Under such processes, observations at the tips of a phylogenetic tree have a multivariate Gaussian distribution, which may lead to suboptimal model specification under certain evolutionary conditions, as supposed in models of punctuated equilibrium or adaptive radiation’ ([3], p. 193).

These three papers are representative of a substantial number of other recent high-profile studies that have discussed their findings in the context of PE [4–16]. This is somewhat remarkable given that arguably no idea has had such a turbulent history in modern evolutionary thought as PE. In the early 1970s, Eldredge and Gould [17–19] proposed that the predominant pattern of evolution throughout deep time is that of stasis ‘punctuated’ by brief intervals of rapid evolution, which often occurred during speciation events (see [Glossary](#)). This was originally conceived as a way of bridging the gap between

Glossary

Anagenetic change: phenotypic evolution that occurs within a species. Often associated with phyletic gradualism, but can also occur at variable rates.

Brownian motion (BM) model: a statistical model used in macroevolutionary studies in which trait variation accumulates according to a continuous-time random walk at a constant rate through time and across lineages.

Cladogenetic change: phenotypic evolution that occurs during the process of speciation.

Macroevolution: the study of processes and patterns of evolution that occur at or above the level of species.

Ornstein–Uhlenbeck (OU) model: a mean-reverting statistical model used in macroevolutionary studies in which trait change is influenced by both an undirected random walk and a restraining force pulling the trait towards an ‘optimum’ value. Given enough time, this results in a stationary distribution of trait values across lineages.

Phyletic gradualism: a pattern of phenotypic evolution characterized by the accumulation of small changes that occur at a more or less constant rate.

Pulsed evolution: a pattern of phenotypic evolution characterized by brief periods of elevated rates of trait evolution substantially higher than the background rates. These periodic pulses may or may not be associated with speciation events.

Punctuated equilibrium: a macroevolutionary theory that posits that the predominant pattern of phenotypic evolution is that of relative stasis for the majority of species’ ‘lifetimes’, followed by brief intervals of rapid evolution during speciation events.

Species selection: broadly construed as the repeatable effects of a trait on the rate of diversification of species possessing it. This term is sometimes considered to apply only when the trait in question is an emergent property of the species (e.g., geographic range size).

Species sorting: individual-level traits (e.g., body size) influence the diversification rates of species. We view this as being subsumed within the concept of species selection and do not distinguish between these in this paper.

Stasis: a pattern of little net evolutionary change within a lineage.

Wright’s rule: evolutionary changes at speciation are random with respect to the direction of a macroevolutionary trend.



prevailing ideas about speciation (i.e., Mayr's allopatric model [20]) and observations from the fossil record [21]. However, PE has expanded and shifted in definition to become a much more far-reaching hypothesis to many researchers. Consequently, it has been viewed as both a rather innocuous statement about the general patterns found in the fossil record and as an affront to the central tenets of evolutionary theory [22–26]. For some researchers the stakes of the debate over the prevalence of PE could not have been higher:

'If most evolutionary change occurs during speciation events, and if speciation events are largely random, natural selection, long viewed as the process guiding evolutionary change, cannot play a significant role in determining the overall course of evolution. Macroevolution is decoupled from microevolution...' ([22], p. 648).

In the wake of such claims, much of the intellectual history of PE has been characterized by fierce, and often vitriolic, theoretical debates – exhaustively catalogued in [26,27] – and the theory remains divisive (for more details on the history of the idea, see [21]).

The field of macroevolution has recently witnessed a resurgence of interest in PE as paleobiologists and, increasingly, comparative biologists armed with molecular phylogenies, have applied sophisticated statistical models to test quantitatively the major hypotheses of PE. In this review we ask whether these new statistical advances have 'rescued' PE from intellectual extinction. We answer this question in the negative. The challenges inherent in elucidating macroevolutionary processes and patterns from paleontological and comparative data are only exacerbated by the muddled historical legacy of PE. Although a number of studies have indeed discussed their findings in light of PE, they have actually addressed a wide variety of conceptual issues; the studies from which we have quoted above exemplify this – each one asks a fundamentally distinct question.

What then, exactly, defines PE? The central definitions and concepts of PE have shifted substantially over time – including the views of the chief advocates of the theory [21]. We argue that the key to disentangling this Gouldian knot lies not in attempting to parse the literature in search of the true 'essence' of PE, but rather in recognizing that the myriad concepts often associated with the theory can be conceptually dissociated and evaluated independently. We believe that dissociating the different components of PE will lead to a more productive discussion of these ideas and facilitate progress in some of the most fundamental questions in macroevolution. In this review we identify four key questions that have been lumped under the topic of PE, discuss how their association with each other has led to confusion, and comment on recent methodological developments, using a variety of types of data, that may provide novel insights into large-scale patterns of diversity.

Punctuated equilibrium as a conglomerate of concepts

In our view, the theory of PE (and the extensive discussion surrounding it) conflates four separate primary research questions: (i) what is the relative importance of gradualistic versus pulsed evolution? (ii) what is the

role of speciation events (cladogenesis) versus within-lineage evolution (anagenesis) in generating trait divergence? (iii) when change is cladogenetic, are these changes adaptive or driven by neutral processes? and (iv) how important is higher level selection (species selection) in shaping patterns of diversity?

Gradualistic versus pulsed evolution

In principle, it is feasible to distinguish gradualistic versus pulsed evolution using either phylogenetic comparative or paleontological data. Constant-rate gradualism is typically modeled as a random walk or Brownian motion (BM) process (Box 1) in both phylogenetic and paleobiological studies. Several recent studies have examined whether fossil time-series conform better to predictions from constant-rate BM, phenotypic stasis, or directional evolution, which each predict different distributions of trait values through time and can be distinguished using model selection techniques [13]. These studies have found mixed support for each mode of evolution in different lineages and traits [4–6,13,28,29]. An exceptional demonstration of pulsed evolution in the fossil record was examined by Hunt *et al.* [6], who found support for a rapid pulse of evolution in sticklebacks as they colonized a novel adaptive peak. Similar model-fitting approaches have been used to demonstrate that in particular fossil time-series shifts in the mode of evolution (i.e., directional evolution, stasis, or BM) are separated by phenotypic bursts [5]. This pulsed pattern of evolution is supported by large collections of micro- and macroevolutionary data [11,30]. Studies of fossil time-series often include caveats that may complicate inference of evolutionary modes. These include unequal sampling probabilities and uncertain stratigraphic position, as well as issues relating to range shifts, time-averaging, and phenotypic plasticity [31]. Hannisdal [32] developed a sophisticated, albeit data-intensive, approach which can incorporate some of these additional sources of uncertainty in a Bayesian framework; we view this to be an important line of research that deserves increased attention.

Although fossil time-series provide direct observations of phenotypes over time, they are limited by the difficulty in confidently assembling sequences of ancestor–descendant relationships. Phylogenetic comparative methods provide a complementary means to study departures from constant-rate gradualism. These can be applied to both extant and extinct data if the fossil data can be placed in a phylogenetic context [33]. Several methods allow the detection of rate shifts across clades by allowing the BM rate parameter σ^2 to differ across branches of the phylogeny [34,35]. However, these methods model sustained shifts in evolutionary rates, rather than the pulsed patterns suggested by PE. Pure-burst models, in which all change accumulates in pulses, can also be fit to phylogenies and more closely align with PE [11,36]. Landis *et al.* [3] developed a particularly promising approach that models both gradual and punctuational patterns of evolution using jump-diffusion models, in which both jumps and gradual evolution come from a single, long-tailed distribution (see also [37]). In addition to these methods, discrete shifts in adaptive optima separated by long periods of stabilizing

Box 1. Modeling trait evolution

The same basic set of stochastic models are often fit to both fossil time-series and phylogenetic comparative methods. Phyletic gradualism is formulated statistically as constant-rate Brownian motion (BM). This model describes a continuous-time random walk in which the amount of phenotypic change in the population trait mean (\bar{z}) over time-interval t is:

$$\Delta\bar{z} = \sigma dW \quad [I]$$

where dW is a continuous-time stochastic process (BM process) drawn from a normal distribution with mean 0 and variance t . Because change over each time interval is independent of previous time-intervals (i.e., the process is Markovian), the amount of variance among replicate lineages increases linearly through time such that $Var(\bar{z}) = \sigma^2 t$. The covariance between observations is proportional to the shared evolutionary history of samples, which for comparative methods is provided by the phylogeny.

To model discontinuous processes, a shift location is estimated either on a fossil time-series or a phylogeny. For pulsed models, a shift corresponds to a burst in phenotypic evolution, against a background of a single, constant rate BM parameter (σ^2). Multiple bursts can be modeled, for example, as a compound Poisson point process, in which bursts occur stochastically at exponentially distributed time-intervals at rate λt and magnitudes drawn from a normal distribution with parameters $(\mu_{burst}, \sigma^2_{burst})$. However, several comparative methods do not model bursts, but instead fit different parameters or models on either side of the shift, corresponding to either an increased or decreased rate of evolution [5,34,35]. Thus, for k

shifts, there would be $k+1$ BM rate parameters ($\sigma^2_1, \dots, \sigma^2_{k+1}$). These models can also be combined to model jointly both bursts and rate shifts [37].

BM models predict that divergence can increase without bounds, which is unrealistic under adaptive scenarios of trait evolution or under models of stasis, where traits are expected to evolve around adaptive optima. A simple extension of a BM model is an Ornstein–Uhlenbeck (OU) model of trait evolution. The per unit time change in mean phenotype under this model is:

$$\Delta\bar{z} = -\alpha(\bar{z} - \theta) + \sigma dW \quad [II]$$

where σdW is identical to a BM process and contributes stochastically to divergence, θ is the optimum trait value, and α is a pull parameter that governs how strongly the population mean is pulled toward θ . Thus, divergence is a balance between the stochastic diffusion parameter (σ^2) and the deterministic pull parameter (α) toward the optimum value (θ) (Figure 1). As with BM models, discontinuous OU models can allow for shifts in rate parameters (α, σ^2), which has been implemented in a phylogenetic context [80]. Rapid shifts in optima are more naturally included in OU models via shifts in the θ parameter, and have been used extensively in phylogenetic comparative methods [39,79,80]. Population trait means approach a new optimum at a rate proportional to the strength of α . Models with alternative patterns of selective regimes can then be compared via model selection techniques to evaluate adaptive hypotheses [79]. OU models have proven to be very useful for inferring various processes using both phylogenetic comparative [41,127] and paleontological [5,128] data.

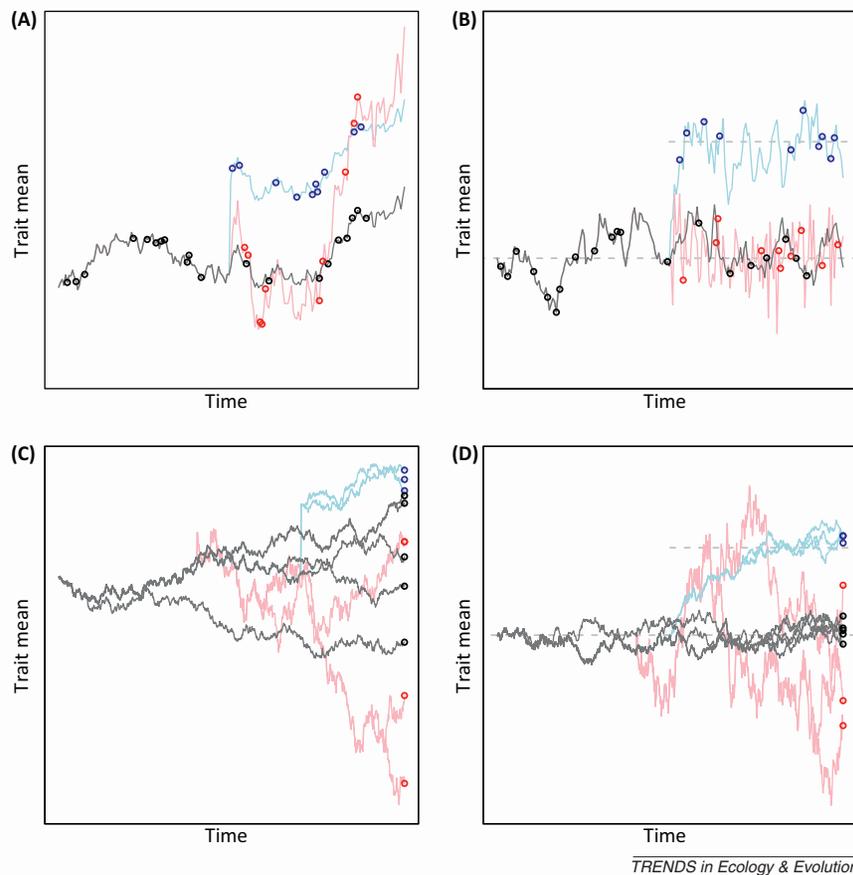


Figure 1. Simulated datasets for different models of trait evolution from fossil time-series (A,B) and phylogenetic comparative data (C,D) under Brownian motion (BM) models (A,C) and Ornstein–Uhlenbeck (OU) models (B,D). Gray lines are simulated as constant-rate BM and OU processes, with circles indicating sampled data. Light-blue lines are discontinuous processes in which a burst of evolution occurs in the form of a single displacement (for BM models) or a walk to a new optimum (θ_2) for OU models. However, all other parameters are kept constant. By contrast, pink lines are models in which rate parameters (σ^2 and α for BM and OU models, respectively) shift to higher values and remain constant thereafter, but are not burst-like (no shift in the expected value of the process).

selection have been extensively implemented in phylogenetic comparative methods by using Ornstein–Uhlenbeck (OU) models. OU models are attractive alternatives to BM that can incorporate stasis, stabilizing selection, and adaptive hypotheses [33] (see [Box 1](#) for further details).

Advances in quantitative model-fitting of evolutionary processes have allowed us to explore a much wider range of evolutionary hypotheses and processes than simple BM (reviewed in [33]), including pulsed evolutionary patterns. However, to what extent does the framework of punctuated equilibrium contribute to interpretation of the results of these models? Note that none of the models described in this section can distinguish between cladogenetic or anagenetic change (see next section). Furthermore, PE is tied to a very specific pattern of evolution and a specific temporal frame: stasis over the lifespan of a species (typically millions of years) followed by geologically brief bursts of phenotypic evolution occurring at speciation [18,19,27]. Therefore, even robust support for a pattern of pulsed evolution – represented by shifts in trait values along branches that are not accounted for by a model of gradual evolution – may be incompatible with PE if the pulses occur too infrequently for conventional PE theory, which predicts pulses at all (or nearly all) speciation events. In addition, exactly as paleontologists have long recognized that repeated burst–stasis episodes can appear gradualistic if viewed at too coarse a scale, gradualistic evolution with variable rates can appear pulse-like at the same coarse scale. A pulsed pattern detected from a phylogeny – which typically has a much longer timescale and coarser sampling than a fossil time-series, may not reflect phenotypic bursts between species. Instead, pulses estimated from model-fits may reflect ‘jumps’ between higher-level niche space or adaptive zones, within which whole clades or groups of species may cluster [37–40]. The observation that groups of species cluster around different phenotypic optima [41] says nothing about whether individual lineages exhibit a pattern of stasis and phenotypic bursts of evolution over the lifespan of individual species. Tying patterns measured at phylogenetic scales to species-level and not clade-level change is fraught with difficulty. However, we can still address other interesting macroevolutionary questions such as whether evolution is characterized by pulses, how often they occur, and what ecological factors may be associated with them [37].

Anagenetic versus cladogenetic change

Although speciation is undoubtedly associated with genetic and trait divergence [42], its relative importance compared to evolutionary change within a lineage is currently poorly understood. Several studies have attempted to investigate the contribution of cladogenetic change to trait evolution using paleontological data [43–46]. This is evaluated by determining whether the stratigraphic ranges of descendant species overlap with their progenitor species, indicative of coexistence and cladogenesis. However, robustly distinguishing between cladogenetic and anagenetic changes using fossil data crucially depends on several assumptions, such as the accurate reconstruction of ancestor–descendant relationships, the equivalency of species

concepts applied to fossil and extant taxa, robust estimation of the temporal ranges of species, and enough sampling to eliminate the possibility of gradual evolution [44]. Disputes over the validity of these assumptions have been well played out in the PE literature, such as in the case-study of the Turkana Basin molluscs [47–49]. Approaches have been developed to account for potential biases (such as estimating ancestor–descendant relationships [50] and stratigraphic ranges [51,52], while accounting for sampling), but difficulties remain. As a recent example, Strotz and Allen [46] found a predominance of cladogenetic change among fossil Foraminifera, using assumed ancestor–descendant relationships assembled from stratigraphic and phenotypic data [45]. We view such claims with considerable skepticism because it is impossible to detect cryptic speciation – which is increasingly being inferred in many extant groups [53] – in the fossil record, and therefore distinguish decisively between anagenetic and cladogenetic change.

Another tactic to assess the question of comparing anagenetic and cladogenetic change has been to look for correlations between speciation rates (or species richness, as a proxy for speciation rates) and rates of evolution using phylogenetic comparative data. This has been done by fitting regression models between inferred lineage-specific rates of evolution and diversification. Several studies have demonstrated such a correlation using a variety of characters, including genetic substitutions [1,7,54], morphological traits [14,55,56], and linguistic characters [16]. However, demonstrating a correlation between speciation rates and trait evolution does not demonstrate that the actual speciation events themselves are coincident with changes in the traits. For example, higher rates of speciation and trait evolution might both be driven by a common cause [12] (see below and [Box 2](#)).

A more statistically robust avenue for partitioning out the influence of cladogenetic versus anagenetic change is to use statistical models, which explicitly parameterize both of these components and simultaneously estimate them using maximum likelihood or Bayesian inference. Early forms of such models assumed that all speciation events were captured by the reconstructed phylogeny. These methods partition the variance in trait values between the speciation events and the background evolution occurring within a lineage [57–59]. More sophisticated approaches attempt to model simultaneously the diversification process together with trait evolution [2,8,9,15,60–62] to account for the fact that extinction has erased many of the speciation events in the inferred phylogeny (see [Box 3](#) for details). Such model-based approaches are not without caveats. Importantly, violations of simplifying assumptions may strongly affect inferences, and methods to evaluate model adequacy are sorely needed. Furthermore, unaccounted measurement error may be erroneously folded into estimates of cladogenetic change. In fact, we should expect samples of recently diverged species to differ substantially regardless of whether evolution is punctational or gradual – even after accounting for simple forms of sampling error – due to within-lineage processes such as local adaptation [11]. These processes may or may not be important for macroevolutionary patterns [63,64] and are difficult to model using current comparative methods [65].

Box 2. An example of how PE can mislead inferences

From a historical perspective it is undoubtedly accurate that numerous comparative methods owe their genesis to the framework of PE. However, the temptation to frame these methods as tests of PE is, in our opinion, unwarranted. For example, Webster *et al.* [7] developed a method to correlate the total genetic distance between the root and the tip of the tree (hereafter, the path length) with the number of nodes along that path (Figure 1). A significant correlation between path length and the number of nodes rejects constant-rate gradualism in molecular evolution, purportedly in favor of a PE model. This correlation has been repeatedly demonstrated in a variety of datasets in traits ranging from molecular sequences to human languages [1,7,16,54] (but see [129]).

However, to what extent is there evidence in these cases for PE? We argue: very little. Eldredge and Gould [18] hypothesized that allopatric speciation causes pulsed phenotypic divergence. However, the direction of causality can just as easily be reversed. Genetic divergence is expected to promote speciation under many models of speciation [42]. Alternatively, divergence and speciation may result indirectly from causal links with a third factor, such as shorter generation times, higher fecundity, or increased genetic variation, to

name a few [129]. Furthermore, trait evolution need not be pulsed for a positive correlation to exist. This effect was demonstrated by Rabosky [12], who showed that correlations between path length and speciation are expected whenever trait evolutionary rates are correlated with rates of speciation, even under purely gradual models. Finally, trait change may not be correlated with speciation at all, but instead with extinction rates. This may occur, for example, if higher evolvability decreases extinction risk [54]. It is certainly a worthwhile avenue of research to establish a correlation between diversification and trait evolutionary rates, but the available tests demonstrate nothing about whether or not trait evolution is pulsed, whether trait change accumulates anagenetically or cladogenetically, or the direction of causality. Taking the ‘off-the-shelf’ interpretation of these macroevolutionary patterns in the form of PE only obfuscates understanding, and worse, could lead to recapitulating four decades worth of often unproductive and contentious debates. Instead, we argue that we should focus on inferences that may be tested effectively using our available statistical tools. These tools should be integrated with more narrowly defined theories that are free of the unwanted assumptions of PE.

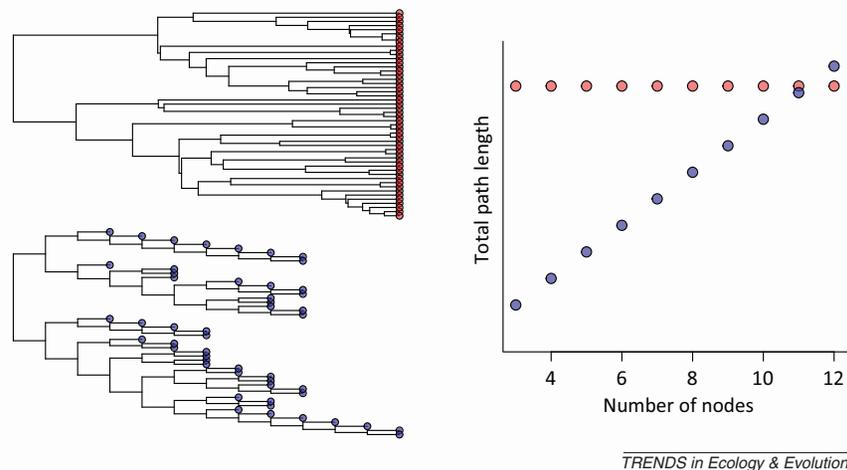


Figure 1. Illustration of a method for correlating evolutionary divergence with speciation. Branch lengths for both phylogenies are in units of evolutionary change. The total path length from the root of the tree to the tips is plotted against the total number of nodes along that path. A positive correlation (blue) is indicative of a relationship between the number of speciation events and evolutionary change, whereas under constant-rate gradualism no such relationship exists (red). Adapted from [1].

Even when speciation is inferred to be associated with trait divergence, a broader conceptual issue remains: what are the causal mechanisms that could generate such an association against a general backdrop of apparent stasis [66,67]? Speciation has long been thought of as a major driver of phenotypic change, both in the context of PE and in evolutionary biology more broadly (but see [68]). In their original conception of PE, Eldredge and Gould [17,18] viewed the pattern posited by PE as a consequence of Mayr’s model of allopatric speciation [20]; as such, speciation is considered a mechanism that interrupts stasis [63,64]. However, the causes of stasis in macroevolutionary data are still unclear [30,64,69,70]. In particular, the direction of causality cannot be elucidated from the statistical methods we have described. Alternative explanations remain such that trait evolution often generates reproductively isolated lineages or that both species diversification and trait evolution rates increase with greater evolutionary ‘versatility’ or ‘evolvability’ [14,71]. There is some evidence that there is variation in the ability of lineages to evolve novel phenotypes [72] although the causes of this variation are still poorly understood and deserve further

research. Regardless, it is important to recognize that a central tenet of PE theory – that speciation causally leads to phenotypic evolution – remains impossible to evaluate from either phylogenetic comparative or paleontological data.

Adaptive versus neutral evolution at speciation

One of the most contentious ideas surrounding PE is that changes associated with speciation are random or neutral; this is what led Gould, Stanley, and others to claim that macro- and microevolution were effectively ‘decoupled’. There are actually two specific versions of this question and these, similarly to many of the ideas we discuss throughout the paper, have often been conflated. The first version is that the changes that occur are random with respect to the direction of a macroevolutionary trend. This is referred to as ‘Wright’s rule’ in the paleobiology literature [19] and has been evaluated by testing whether trait differences between ancestors and descendants are directionally biased. More precisely, this has been looked at by evaluating whether the mean of the distribution of changes are significantly different than zero (the null expectation

Box 3. Modeling species selection and cladogenetic change on phylogenies

In a ground-breaking paper, Maddison and colleagues [103] developed a statistical framework that has opened up investigation into two major components of PE – the influence of traits on diversification ('species selection', *sensu* [86,90]) and cladogenetic character change. The premise of the approach is that instead of specifying a full likelihood of the model, one needs only to describe the probabilities of all possible events that could occur in a very short time interval, Δt (see Figure 1 in Box 3), solve a differential equation, and then use numerical integration to evaluate the likelihood of the model given the phylogeny and trait data at the tips (see [103] for full details). The initial model considered by Maddison *et al.* [103] was the binary state speciation extinction ('BiSSE') model in which different states for a single character resulted in different diversification rates.

Consider that lineage diversification can be modeled by a birth–death process [107] in which there is a constant rate of speciation (λ) and extinction (μ) across the clade. Lineages with a trait in state 0 diversify at rates λ_0 and μ_0 , and lineages in state 1 diversify at rates λ_1 and μ_1 . Transitions (anagenetic evolution) between states $0 \rightarrow 1$ occur at rate θ_{01} and transitions from $1 \rightarrow 0$ occur at rate θ_{10} . The probabilities of all possible events that can occur during Δt can be described as a set of differential equations. One can then use the integration machinery (as described in [103]) to estimate simultaneously all parameters using either maximum likelihood or Bayesian inference to test for a statistical difference between λ_0 and λ_1 (or between μ_0 and μ_1) to infer the strength of species selection.

The BiSSE model was extended by Magnuson-Ford and Otto ('BiSSE-ness' [62]) to allow for the possibility of character transitions at speciation (cladogenetic change); (an identical model was independently derived by Goldberg and Igić [61], and related approaches, which we will not describe in detail here, were also developed by Bokma [8,9,60].) In addition to the six parameters of the BiSSE model (λ_0 , λ_1 , μ_0 , μ_1 , θ_{01} , θ_{10}), their model includes the probabilities of a change occurring at a speciation event (p_{0c} and p_{1c} , for the two states, respectively) as well as the probabilities that the character changes

are asymmetric, where the change only occurs in one of the two daughter lineages, p_{0a} and p_{1a} (often referred to as 'budding cladogenesis' in the paleobiological literature). This allows one to evaluate simultaneously the importance of species selection as well as the relative importance of cladogenetic versus anagenetic change. This model also highlights the general message of our paper; the questions can be evaluated independently of each other if parameter sets are constrained:

$\lambda_0 = \lambda_1$, $\mu_0 = \mu_1$ Estimate cladogenetic and anagenetic rates only
 θ_{01} , $\theta_{10} = 0$ Estimate species selection with only cladogenetic change
 p_{0c} , p_{1c} , p_{0a} , $p_{1a} = 0$ Estimate species selection with only anagenetic change

thus making it an excellent statistical framework (though certainly not the only one) for evaluating the questions associated with PE.

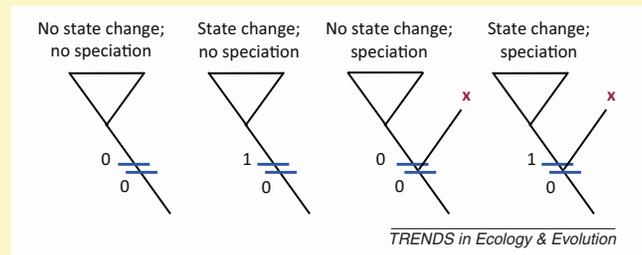


Figure 1. A conceptual figure demonstrating how the BiSSE/BiSSE-ness model simultaneously models species selection and trait change. The four trees depict four possible events that could occur during a short time interval (Δt ; the area between the two blue bars) along a given branch: no events happen; the trait changes; speciation occurs; both speciation and trait change occur (these do not depict all possible scenarios.) If each event can be assigned a probability of occurring over Δt , one can derive differential equations describing the entire model, and use numerical integration to calculate the likelihood of the model given a phylogenetic tree and trait data at the tips. Adapted from [103].

under most models of trait evolution) [73,74]. For example, if there is a trend of increasing body size throughout the history of a clade (Cope's rule [75]), then Wright's rule requires that daughter species, at speciation, are on average no bigger (or smaller) than their ancestor. To Gould and Eldredge [19] (as well as to other researchers [22,23]), Wright's rule was a key justification for including species selection in a PE framework; if change only occurs at speciation, and change is random with respect to macroevolutionary trends, those trends can only be explained by species selection. However, if we recognize that the nature of change at speciation is independent of species selection (see below), then establishing Wright's rule has no bearing on the strength of higher level selection in a group [15]. At the same time, biased transmission may still be involved in macroevolutionary trends [73,76].

The second, broader version of the claim is that change at speciation is driven by neutral processes rather than adaptive evolution [23,24,27]. This is much more complex to address. There have been several attempts to investigate the hypothesis that past trait changes were adaptive using phylogenetic comparative or paleobiological data (see, e.g., [77]). For example, phylogenetic methods can attempt to associate trait changes with changes in the selective regimes experienced by those lineages [78–80], or studies of functional morphology can provide specific hypotheses about relationships between trait states and the environment, which can then be tested statistically [81]. However, these necessarily rely on either detailed information about form, function, and the environment

(e.g., [82] and references therein) or *a priori* hypotheses regarding what was adaptive at some period in the past. We know from studies of wild populations that the direction of selection is often temporally and spatially variable [83,84], and it is therefore extremely tenuous to draw conclusions regarding the adaptive value of changes during speciation from comparative or paleontological data alone. There has been a great deal of study investigating the patterns of evolution throughout the course of speciation using natural populations, experimental systems, and mathematical models [42,85–87]. In particular, many recent studies have explored the distinction between ecological speciation, where speciation is driven by divergent natural selection between lineages, and other forms of speciation such as mutation–order speciation, speciation driven by sexual selection, and others (reviewed in [42]). As a result of these studies, we have learned a great deal about the mechanisms involved in speciation and are beginning to understand the relative importance of adaptive and neutral processes during speciation across a broad suite of taxa – although it is much too early to draw any sweeping conclusions. We strongly suggest that this avenue of research is far more appropriate for addressing this aspect of PE than analyzing either phylogenetic comparative or paleontological data alone.

Species selection as a macroevolutionary process

Though long controversial in its own right [88], the idea that natural selection can act on species-level characteristics is becoming more widely appreciated [86,88–90].

Here we follow the lead of other authors [86,90,91] and define species selection as ‘repeatable effects of that trait on the rate of diversification of species possessing it’ ([86], p. 444), regardless of whether or not the trait is an emergent property of the lineage or the aggregate of individual-level traits. We therefore ignore the (in our view, unnecessary) distinction between ‘species selection and ‘species sorting’ (*sensu* [92]). For an alternative perspective on this issue, see Jablonski’s excellent review of the topic [89].

The idea that the tempo and mode of evolutionary change is inexorably linked to selection at the lineage level is an old and persistent one and is, in the minds of at least some researchers, part and parcel of a broader macroevolutionary theory [19,22–27]. The reasoning behind this is that, in some researchers’ conception of the process, selection can only act on ‘evolutionary individuals’ [93] and species can only operate as such if they have a definite beginning and end [19,27] – a pattern that is produced if evolutionary change only occurs at cladogenesis. Although this may seem intuitive, such an association is logically (and mathematically) unnecessary. Species selection does not require any particular mode of evolutionary change and it certainly does not require the majority of change to be concentrated at speciation events [15,94–98] (see discussion in [89], pp. 513–514). The conflation of species selection with punctuated change has been cited by some authors to be a cause of antagonism towards species selection [88].

Species selection has been recently reviewed in depth [89,90] and we will not attempt to be comprehensive here. Instead, we focus on recent methodological developments that have improved our ability to detect species selection. Conventionally, inference regarding the influence of a trait on diversification rate from molecular phylogenies has been carried out by comparing the diversities or diversification rate between independent pairs of sister taxa [99,100]. However, this is problematic for several reasons including statistical power [101] and that asymmetries in character transition rates can confound asymmetries in diversification rate (and vice versa) [102]. A major innovation to deal simultaneously with this issue and investigate the correlation between traits and speciation and extinction was made by Maddison *et al.* (the binary state speciation and extinction, or ‘BiSSE’, model [103]; see Box 3). This has been extended beyond binary traits to investigate the effect of multi-state discrete traits [104], quantitative traits [105], and geographic range [106] on lineage diversification. We note that although these are certainly very promising statistical approaches, they rely on large sample sizes and potentially dubious assumptions, such as that diversification can be modeled as constant-rate branching process (i.e., a ‘birth–death’ model [107]), that rates of evolution are constant across the phylogeny, and that the directionality and strength of species selection is consistent. There is substantial evidence that suggests that diversification rates are not constant through time or across clades [12,14,108], perhaps owing to diversity-dependent diversification [109,110], and this has been a focus of modeling work in both paleobiology [111,112] and phylogenetic comparative methods [113,114]. Similarly, rates of trait evolution are likely often to be heterogeneous

[35,115] and the vector of species selection has been inferred to be variable in some groups [116–118]. Some of these assumptions can potentially be relaxed [119] but, in general, the robustness of these methods to severe violations awaits further investigation. In paleobiology there has been increasing development of multivariate methods to partition out the effect of various correlated traits on speciation and or extinction, which is key to elucidating causal mechanisms. This has been accomplished using statistical techniques such as general linear models (i.e., predict lineages’ diversification rates or durations in the fossil record from lineage-specific traits) [118,120,121]. Alternatively researchers have used the Price equation [98,122] to examine the covariance between traits and diversification rates. The Price equation was first proposed for the purposes of studying macroevolution by Arnold and Frisrup [96], and this has recently been expanded upon by Simpson and colleagues [117,123]. Adopting the Price equation also allows for the possibility of a unified approach to the study of species selection across data types that could potentially be applied to both phylogenetic comparative data and fossil time-series [15,89].

Concluding remarks

In this paper we have described quantitative approaches to addressing four fundamental macroevolutionary questions that have long been conflated with each other in the literature on PE. Confusion among these disparate and independent questions has led many researchers to consider PE as being robustly verified, whereas others believe the theory bankrupt of empirical support. Either view may be justified depending on which component an individual researcher considers the essence of PE theory. If macroevolutionary researchers dissociate these concepts, the fact that some may be more difficult to evaluate or are less theoretically sound should not impede progress on other questions. Although we argue throughout that the questions that make up PE can be addressed independently, this does not preclude synthesis. Instead, multiple processes could be important (and often, probably are) to understanding the accumulation of diversity and disparity through deep time. For example, Goldberg *et al.* [124] used the BiSSE model to demonstrate that species selection was important for the maintenance of self-incompatibility in the plant family Solanaceae (the ‘nightshades’). In a subsequent paper, Goldberg and Igić [61] reanalyzed the same data but used a model that allowed for trait evolution within a lineage, species selection, and (additionally) trait evolution occurring at cladogenesis. They found that all three processes appear to be important in this group. Nonetheless, these are independent processes that may or may not be linked mechanistically, in this group or others. Instead of bringing new insight into PE – and thereby rescuing the term from its historical problems – novel developments have demonstrated that the terminology associated with PE can be problematic. We believe that emerging statistical models and datasets are best suited for testing independent components of PE theory. Evaluation of these methods in the context of PE will only lead to confusion. Although PE undoubtedly served as a catalyst in the development of concepts and methods discussed

above, we think it is time to move on, and encourage researchers in macroevolution to look forward rather than look back.

Paleontologists may view our discussion of PE as being too harsh on theoretical constructs from their discipline – and, admittedly, all of us were raised in the traditions of population biology and evolutionary genetics. However, we will also note that, in our reading of much of the literature using phylogenetic comparative methods, we have found a recurrent theme of comparative biologists adopting concepts from the paleobiological literature (including, but not limited to, PE), but doing so rather blithely. Although it is widely recognized that incorporating fossil data into comparative studies will dramatically improve the inferences we can draw from them [33,125,126], concordant attention has not been paid, in our opinion, to the conceptual foundations which underlie the studies. Comparative biologists have much to gain by engaging more seriously with the arguments and ideas from the rich literature in paleontology on rates of evolution, macroevolutionary trends, species selection, adaptive radiations, and so forth. A truly synthetic macroevolutionary research program will involve the melding of data and theory from different disciplines, and a thoughtful examination of precisely what the fundamental questions are and how we can go about answering them.

Acknowledgments

We would like to thank Paul Craze for inviting us to contribute this paper and for his encouragement and editorial guidance throughout the whole process. We thank Peter Wagner, Dan Rabosky, Rob Lanfear, and two anonymous reviewers for thought-provoking comments on previous versions of the manuscript. We would also like to thank Jon Eastman, David Bapst, Carl Simpson, and Paul Harnik for stimulating discussions of these topics. M.W.P. was supported by Bioinformatics and Computational Biology Graduate Fellowship from the University of Idaho.

References

- 1 Pagel, M. *et al.* (2006) Large punctuational contribution of speciation to evolutionary divergence at the molecular level. *Science* 314, 119–121
- 2 Mattila, T.M. and Bokma, F. (2008) Extant mammal body masses suggest punctuated equilibrium. *Proc. R. Soc. B: Biol. Sci.* 275, 2195–2199
- 3 Landis, M.J. *et al.* (2013) Phylogenetic analysis using Lévy processes: finding jumps in the evolution of continuous traits. *Syst. Biol.* 62, 193–204
- 4 Hunt, G. (2007) The relative importance of directional change, random walks, and stasis in the evolution of fossil lineages. *Proc. Natl. Acad. Sci. U.S.A.* 104, 18404–18408
- 5 Hunt, G. (2008) Gradual or pulsed evolution: when should punctuational explanations be preferred? *Paleobiology* 34, 360–377
- 6 Hunt, G. *et al.* (2008) Evolution toward a new adaptive optimum: phenotypic evolution in a fossil stickleback lineage. *Evolution* 62, 700–710
- 7 Webster, A.J. *et al.* (2003) Molecular phylogenies link rates of evolution and speciation. *Science* 301, 478
- 8 Bokma, F. (2002) Detection of punctuated equilibrium from molecular phylogenies. *J. Evol. Biol.* 15, 1048–1056
- 9 Bokma, F. (2008) Detection of ‘punctuated equilibrium’ by bayesian estimation of speciation and extinction rates, ancestral character states, and rates of anagenetic and cladogenetic evolution on a molecular phylogeny. *Evolution* 62, 2718–2726
- 10 Ingram, T. (2011) Speciation along a depth gradient in a marine adaptive radiation. *Proc. R. Soc. B: Biol. Sci.* 278, 613–618
- 11 Uyeda, J.C. *et al.* (2011) The million-year wait for macroevolutionary bursts. *Proc. Natl. Acad. Sci. U.S.A.* 108, 15908–15913
- 12 Rabosky, D.L. (2012) Positive correlation between diversification rates and phenotypic evolvability can mimic punctuated equilibrium on molecular phylogenies. *Evolution* 66, 2622–2627
- 13 Hunt, G. (2012) Measuring rates of phenotypic evolution and the inseparability of tempo and mode. *Paleobiology* 38, 351–373
- 14 Rabosky, D.L. *et al.* (2013) Rates of speciation and morphological evolution are correlated across the largest vertebrate radiation. *Nat. Commun.* 4, 1958
- 15 Simpson, C. (2013) Species selection and the macroevolution of coral coloniality and photosymbiosis. *Evolution* 67, 1607–1621
- 16 Atkinson, Q.D. *et al.* (2008) Languages evolve in punctuational bursts. *Science* 319, 588
- 17 Eldredge, N. (1971) The allopatric model and phylogeny in paleozoic invertebrates. *Evolution* 25, 156–167
- 18 Eldredge, N. and Gould, S.J. (1972) Punctuated equilibria: an alternative to phyletic gradualism. In *Models in Paleobiology* (Schopf, T., ed.), pp. 82–115, Freeman Cooper
- 19 Gould, S.J. and Eldredge, N. (1977) Punctuated equilibria: the tempo and mode of evolution reconsidered. *Paleobiology* 3, 115–151
- 20 Mayr, E. (1942) *Systematics and the Origin of Species from the Viewpoint of a Zoologist*. Harvard University Press
- 21 Sepkoski, D. (2012) *Rereading the Fossil Record: The Growth of Paleobiology and an Evolutionary Discipline*. University of Chicago Press
- 22 Stanley, S.M. (1975) A theory of evolution above the species level. *Proc. Natl. Acad. Sci. U.S.A.* 72, 646–650
- 23 Stanley, S. (1979) *Macroevolution*. Johns Hopkins University Press
- 24 Gould, S.J. (1980) Is a new and general theory of evolution emerging? *Paleobiology* 6, 119–130
- 25 Charlesworth, B. *et al.* (1982) A neo-Darwinian commentary on macroevolution. *Evolution* 36, 474–498
- 26 Levinton, J.S. (2001) *Genetics, Paleontology and Macroevolution*. Cambridge University Press
- 27 Gould, S.J. (2002) *The Structure of Evolutionary Theory*. Harvard University Press
- 28 Grey, M. *et al.* (2008) Variation in evolutionary patterns across the geographic range of a fossil bivalve. *Science* 322, 1238–1241
- 29 Hopkins, M.J. and Lidgard, S. (2012) Evolutionary mode routinely varies among morphological traits within fossil species lineages. *Proc. Natl. Acad. Sci. U.S.A.* 109, 20520–20525
- 30 Estes, S. and Arnold, S.J. (2007) Resolving the paradox of stasis: models with stabilizing selection explain evolutionary divergence on all timescales. *Am. Nat.* 169, 227–244
- 31 Patzkowsky, M.E. and Holland, S.M. (2012) *Stratigraphic Paleobiology: Understanding the Distribution of Fossil Taxa in Time and Space*. University of Chicago Press
- 32 Hannisdal, B. (2007) Inferring phenotypic evolution in the fossil record by bayesian inversion. *Paleobiology* 33, 98–115
- 33 Pennell, M.W. and Harmon, L.J. (2013) An integrative view of phylogenetic comparative methods: connections to population genetics, community ecology, and paleobiology. *Ann. N. Y. Acad. Sci.* 1289, 90–105
- 34 O’Meara, B.C. *et al.* (2006) Testing for different rates of continuous trait evolution using likelihood. *Evolution* 60, 922–933
- 35 Eastman, J.M. *et al.* (2011) A novel comparative method for identifying shifts in the rate of character evolution on trees. *Evolution* 65, 3578–3589
- 36 Hansen, T.F. and Martins, E.P. (1996) Translating between microevolutionary process and macroevolutionary patterns: The correlation structure of interspecific data. *Evolution* 50, 1404–1417
- 37 Eastman, J.M. *et al.* (2013). Simpsonian ‘evolution by jumps’ in an adaptive radiation of *Anolis* lizards. *ArXiv* 1305.4216
- 38 Simpson, G.G. (1944) *Tempo and Mode of Evolution*. Columbia University Press
- 39 Hansen, T.F. (1997) Stabilizing selection and the comparative analysis of adaptation. *Evolution* 51, 1341–1351
- 40 Hansen, T.F. (2012) Adaptive landscapes and macroevolutionary dynamics. In *The Adaptive Landscape in Evolutionary Biology* (Svensson, E. and Calsbeek, R., eds), pp. 205–221, Oxford University Press
- 41 Mahler, D.L. *et al.* (2013) Exceptional convergence on the macroevolutionary landscape in island lizard radiations. *Science* 341, 292–295

- 42 Nosal, P. (2012) *Ecological Speciation*. Oxford University Press
- 43 Wagner, P.J. and Erwin, D.H. (1995) Phylogenetic patterns as tests of speciation models. In *New Approaches to Speciation in the Fossil Record* (Erwin, D.H. and Anstey, R.L., eds), pp. 87–122, Columbia University Press
- 44 Jackson, J.B. and Cheetham, A.H. (1999) Tempo and mode of speciation in the sea. *Trends Ecol. Evol.* 14, 72–77
- 45 Aze, T. *et al.* (2011) A phylogeny of Cenozoic macroperforate planktonic Foraminifera from fossil data. *Biol. Rev.* 86, 900–927
- 46 Strotz, L.C. and Allen, A.P. (2013) Assessing the role of cladogenesis in macroevolution by integrating fossil and molecular evidence. *Proc. Natl. Acad. Sci. U.S.A.* 110, 2904–2909
- 47 Williamson, P. (1981) Palaeontological documentation of speciation in Cenozoic molluscs from Turkana Basin. *Nature* 293, 437–443
- 48 Fryer, G. *et al.* (1983) Punctuated equilibria, morphological stasis and the palaeontological documentation of speciation: a biological appraisal of a case history in an African lake. *Biol. J. Linn. Soc.* 20, 195–205
- 49 Van Bocxlaer, B. *et al.* (2008) Gradual versus punctuated evolution in the Turkana Basin molluscs: evolutionary events or biological invasions? *Evolution* 62, 511–520
- 50 Foote, M. (1996) On the probability of ancestors in the fossil record. *Paleobiology* 22, 141–151
- 51 Marshall, C.R. (1994) Confidence intervals on stratigraphic ranges: partial relaxation of the assumption of randomly distributed fossil horizons. *Paleobiology* 20, 459–469
- 52 Marshall, C.R. (1997) Confidence intervals on stratigraphic ranges with nonrandom distributions of fossil horizons. *Paleobiology* 23, 165–173
- 53 Fujita, M.K. *et al.* (2012) Coalescent-based species delimitation in an integrative taxonomy. *Trends Ecol. Evol.* 27, 480–488
- 54 Lanfear, R. *et al.* (2010) Mutation rate is linked to diversification in birds. *Proc. Natl. Acad. Sci. U.S.A.* 107, 20423–20428
- 55 Adams, D.C. *et al.* (2009) Are rates of species diversification correlated with rates of morphological evolution? *Proc. R. Soc. B: Biol. Sci.* 276, 2729–2738
- 56 Rabosky, D.L. and Adams, D.C. (2012) Rates of morphological evolution are correlated with species richness in salamanders. *Evolution* 66, 1807–1818
- 57 Pagel, M. (1997) Inferring evolutionary processes from phylogenies. *Zoolog. Scr.* 26, 331–348
- 58 Mooers, A.Ø. *et al.* (1999) Using phylogenies to test macroevolutionary hypotheses of trait evolution in cranes (Gruinae). *Am. Nat.* 154, 249–259
- 59 Wagner, P.J. and Marcot, J.D. (2010) Probabilistic phylogenetic inference in the fossil record: current and future applications. In *Quantitative Methods in Paleobiology* (Alroy, K. and Hunt, G., eds), pp. 195–217, The Paleontological Society
- 60 Bokma, F. (2010) Time, species, and separating their effects on trait variance in clades. *Syst. Biol.* 59, 602–607
- 61 Goldberg, E.E. and Igić, B. (2012) Tempo and mode in plant breeding system evolution. *Evolution* 66, 3701–3709
- 62 Magnuson-Ford, K. and Otto, S.P. (2012) Linking the investigations of character evolution and species diversification. *Am. Nat.* 180, 225–245
- 63 Futuyma, D.J. (1987) On the role of species in anagenesis. *Am. Nat.* 130, 465–473
- 64 Futuyma, D.J. (2010) Evolutionary constraint and ecological consequences. *Evolution* 64, 1865–1884
- 65 Stone, G.N. *et al.* (2011) Controlling for non-independence in comparative analysis of patterns across populations within species. *Philos. Trans. R. Soc. B: Biol. Sci.* 366, 1410–1424
- 66 Benton, M.J. and Pearson, P.N. (2001) Speciation in the fossil record. *Trends Ecol. Evol.* 16, 405–411
- 67 Eldredge, N. *et al.* (2005) The dynamics of evolutionary stasis. *Paleobiology* 31, 133–145
- 68 Sætre, G.P. (2013) Hybridization is important in evolution, but is speciation? *J. Evol. Biol.* 26, 256–258
- 69 Hansen, T.F. and Houle, D. (2004) Evolvability, stabilizing selection and the problem of stasis. In *The Evolutionary Biology of Complex Phenotypes* (Pigliucci, M. and Preston, K., eds), pp. 130–150, Oxford University Press
- 70 Walsh, B. and Blows, M.W. (2009) Abundant genetic variation + strong selection = multivariate genetic constraints: a geometric view of adaptation. *Annu. Rev. Ecol. Evol. Syst.* 40, 41–59
- 71 Vermeij, G.J. (1973) Biological versatility and earth history. *Proc. Natl. Acad. Sci. U.S.A.* 70, 1936–1938
- 72 Martin, C.H. and Wainwright, P.C. (2011) Trophic novelty is linked to exceptional rates of morphological diversification in two adaptive radiations of Cyprinodon pupfish. *Evolution* 65, 2197–2212
- 73 Wagner, P.J. (1996) Contrasting the underlying patterns of active trends in morphologic evolution. *Evolution* 50, 990–1007
- 74 Wagner, P.J. (2001) Rate heterogeneity in shell character evolution among Lophospiroid gastropods. *Paleobiology* 27, 290–310
- 75 Rensch, B. (1948) Historical changes correlated with evolutionary changes of body size. *Evolution* 2, 218–230
- 76 McShea, D.W. (1994) Mechanisms of large-scale evolutionary trends. *Evolution* 48, 1747–1763
- 77 Rose, M.R. and Lauder, G.V. (1996) *Adaptation*. Academic Press
- 78 Baum, D.A. and Larson, A. (1991) Adaptation reviewed: a phylogenetic methodology for studying character macroevolution. *Syst. Biol.* 40, 1–18
- 79 Butler, M.A. and King, A.A. (2004) Phylogenetic comparative analysis: a modeling approach for adaptive evolution. *Am. Nat.* 164, 683–695
- 80 Beaulieu, J.M. *et al.* (2012) Modeling stabilizing selection: expanding the Ornstein–Uhlenbeck model of adaptive evolution. *Evolution* 66, 2369–2383
- 81 Wainwright, P.C. (2007) Functional versus morphological diversity in macroevolution. *Annu. Rev. Ecol. Evol. Syst.* 38, 381–401
- 82 Vermeij, G.J. (1987) *Evolution and Escalation*. Princeton University Press
- 83 Grant, P.R. and Grant, B.R. (2002) Unpredictable evolution in a 30-year study of Darwin's finches. *Science* 296, 707–711
- 84 Siepielski, A.M. *et al.* (2009) It's about time: the temporal dynamics of phenotypic selection in the wild. *Ecol. Lett.* 12, 1261–1276
- 85 Schluter, D. (2000) *The Ecology of Adaptive Radiation*. Oxford University Press
- 86 Coyne, J.A. and Orr, H.A. (2004) *Speciation*. Sinauer Associates
- 87 Gavrillets, S. (2004) *Fitness Landscapes and the Origin of Species*. Princeton University Press
- 88 Fitzjohn, R.G. (2012) *What Drives Biological Diversification? Detecting Traits Under Species Selection*. PhD thesis University of British Columbia
- 89 Jablonski, D. (2008) Species selection: theory and data. *Annu. Rev. Ecol. Evol. Syst.* 39, 501–524
- 90 Rabosky, D.L. and McCune, A.R. (2010) Reinventing species selection with molecular phylogenies. *Trends Ecol. Evol.* 25, 68–74
- 91 Williams, G.C. (1992) *Natural Selection: Domains, Levels and Challenges*. Oxford University Press
- 92 Vrba, E.S. and Gould, S.J. (1986) The hierarchical expansion of sorting and selection: Sorting and selection cannot be equated. *Paleobiology* 12, 217–228
- 93 Hull, D.L. (1980) Individuality and selection. *Annu. Rev. Ecol. Syst.* 11, 311–332
- 94 Van Valen, L. (1975) Group selection, sex, and fossils. *Evolution* 29, 87–94
- 95 Slatkin, M. (1981) A diffusion model of species selection. *Paleobiology* 7, 421–425
- 96 Arnold, A.J. and Fristrup, K. (1982) The theory of evolution by natural selection: a hierarchical expansion. *Paleobiology* 8, 113–129
- 97 Rice, S.H. (1995) A genetical theory of species selection. *J. Theor. Biol.* 177, 237–245
- 98 Okasha, S. (2006) *Evolution and the Levels of Selection*. Oxford University Press
- 99 Mitter, C. *et al.* (1988) The phylogenetic study of adaptive zones: has phytophagy promoted insect diversification? *Am. Nat.* 132, 107–128
- 100 Sargent, R.D. (2004) Floral symmetry affects speciation rates in angiosperms. *Proc. R. Soc. Lond. Ser. B: Biol. Sci.* 271, 603–608
- 101 Slowinski, J.B. and Guyer, C. (1993) Testing whether certain traits have caused amplified diversification – an improved method based on a model of random speciation and extinction. *Am. Nat.* 142, 1019–1024
- 102 Maddison, W.P. (2006) Confounding asymmetries in evolution diversification and character change. *Evolution* 60, 1743–1746
- 103 Maddison, W.P. *et al.* (2007) Estimating a binary character's effect on speciation and extinction. *Syst. Biol.* 56, 701–710

- 104 FitzJohn, R.G. (2012) Diversitree: comparative phylogenetic analyses of diversification in R. *Methods Ecol. Evol.* 3, 1084–1092
- 105 FitzJohn, R.G. (2010) Quantitative traits and diversification. *Syst. Biol.* 59, 619–633
- 106 Goldberg, E.E. *et al.* (2011) Phylogenetic inference of reciprocal effects between geographic range evolution and diversification. *Syst. Biol.* 60, 451–465
- 107 Kendall, D. (1948) On the generalized 'birth-and-death' process. *Ann. Math. Stat.* 19, 1–15
- 108 Phillimore, A.B. and Price, T.D. (2008) Density-dependent cladogenesis in birds. *PLoS Biol.* 6, e71 41
- 109 Sepkoski, J.J.J. (1984) A kinetic model of Phanerozoic taxonomic diversity. iii. Post-Paleozoic families and mass extinctions. *Paleobiology* 10, 246–267
- 110 Rabosky, D.L. (2009) Ecological limits and diversification rate: alternative paradigms to explain the variation in species richness among clades and regions. *Ecol. Lett.* 12, 735–743
- 111 Roy, K. (1996) The roles of mass extinction and biotic interaction in large-scale replacements: a reexamination using the fossil record of stromboidean gastropods. *Paleobiology* 22, 436–452
- 112 Sepkoski, J.J.J. *et al.* (2000) Competitive displacement among post-Paleozoic cyclostome and cheilostome bryozoans. *Paleobiology* 26, 7–18 865
- 113 Rabosky, D.L. and Lovette, I.J. (2008) Explosive evolutionary radiations: decreasing speciation or increasing extinction through time? *Evolution* 62, 1866–1875
- 114 Etienne, R.S. *et al.* (2012) Diversity-dependence brings molecular phylogenies closer to agreement with the fossil record. *Proc. R. Soc. B: Biol. Sci.* 279, 1300–1309
- 115 Beaulieu, J.M. *et al.* (2013) Identifying hidden rate changes in the evolution of a binary morphological character: the evolution of plant habit in campanulid angiosperms. *Syst. Biol.* <http://dx.doi.org/10.1093/sysbio/syt034>
- 116 Jablonski, D. (1986) Background and mass extinctions: the alternation of macroevolutionary regimes. *Science* 231, 129–133
- 117 Simpson, C. (2010) Species selection and driven mechanisms jointly generate a large-scale morphological trend in monobathrid crinoids. *Paleobiology* 36, 481–496
- 118 Harnik, P.G. *et al.* (2012) Long-term differences in extinction risk among the seven forms of rarity. *Proc. R. Soc. B: Biol. Sci.* 279, 4969–4976
- 119 Rabosky, D.L. and Glor, R.E. (2010) Equilibrium speciation dynamics in a model adaptive radiation of island lizards. *Proc. Natl. Acad. Sci. U.S.A.* 107, 22178–22183
- 120 Jablonski, D. and Hunt, G. (2006) Larval ecology, geographic range, and species survivorship in Cretaceous mollusks: organismic versus species-level explanations. *Am. Nat.* 168, 556–564
- 121 Harnik, P.G. (2011) Direct and indirect effects of biological factors on extinction risk in fossil bivalves. *Proc. Natl. Acad. Sci. U.S.A.* 108, 13594–13599
- 122 Price, G.R. (1972) Extension of covariance selection mathematics. *Ann. Hum. Genet.* 35, 485–490
- 123 Simpson, C. and Harnik, P.G. (2009) Assessing the role of abundance in marine bivalve extinction over the post-Paleozoic. *Paleobiology* 35, 631–647
- 124 Goldberg, E.E. *et al.* (2010) Species selection maintains self incompatibility. *Science* 330, 493–495
- 125 Quental, T.B. and Marshall, C.R. (2010) Diversity dynamics: molecular phylogenies need the fossil record. *Trends Ecol. Evol.* 25, 434–441
- 126 Slater, G.J. *et al.* (2012) Integrating fossils with molecular phylogenies improves inference of trait evolution. *Evolution* 66, 3931–3944
- 127 Hansen, T.F. *et al.* (2008) A comparative method for studying adaptation to a randomly evolving environment. *Evolution* 62, 1965–1977
- 128 Reitan, T. *et al.* (2012) Phenotypic evolution studied by layered stochastic differential equations. *Ann. Appl. Stat.* 6, 1531–1551
- 129 Goldie, X. *et al.* (2011) Diversification and the rate of molecular evolution: no evidence of a link in mammals. *BMC Evol. Biol.* 11, 286