

INVITED SPECIAL ARTICLE: COMMENTARY

Diversity and skepticism are vital for comparative biology: a response to Donoghue and Edwards (2019)

Jeremy M. Beaulieu^{1,3} and Brian C. O'Meara²

Manuscript received 31 January 2019; revision accepted 4 April 2019.

¹ Department of Biological Sciences, University of Arkansas, Fayetteville, AR 72701, USA

² Department of Ecology and Evolutionary Biology, University of Tennessee, Knoxville, TN 37996-1610, USA

³ Author for correspondence (e-mail: jmbeauli@uark.edu)

Citation: Beaulieu, J. M. and B. C. O'Meara. 2019. Diversity and skepticism are vital for comparative biology: a response to Donoghue and Edwards (2019). *American Journal of Botany* 106(5): 613–617.

doi:10.1002/ajb2.1278

Several objections raised by Donoghue and Edwards (2019) regarding a recent paper of ours (Beaulieu and O'Meara, 2018) provide a good opportunity to discuss the different uses of phylogenies to understand evolutionary history. Are small-scale studies more informative and prone to less error? Or, as we argue, do studies at different scales carry their own benefits and costs? Is *ascertainment bias*, a term we use to describe how individual clades chosen for study may represent a biased sample of life, a potentially critical problem for macroevolutionary studies? Donoghue and Edwards (2019) recently argued that the ascertainment bias we highlighted is not as problematic as we claimed. In fact, they pointed to our own simulations as evidence, which showed that despite the increased variance, analyses of many smaller, variable subclades produced, on average, reasonable estimated rates of the larger and more inclusive clade. We disagree with their interpretation and expand on the ways ascertainment bias may affect our conclusions about the mode and tempo of evolution, and we also highlight areas where more research is needed. Finally, there is a disagreement over whether it is more informative to understand parameters of a process (e.g., rate of trait evolution) or inferred details of a particular history (e.g., ancestral states). We address these issues in reverse order, as each builds toward the next.

MODELS, UNITS, AND NATURAL HISTORY

A key difference between Donoghue and Edwards (2019) and our view centers on the value of parameter estimation. For example,

they are critical that the estimation of a particular rate of, say 42%, is informative (p. 329) and that a “global rate” should not be a goal of comparative biology. These comments reflect two important and undervalued issues that are common in the field: (1) parameters in comparative models typically have units and (2) interpretation of these parameters in the context of the underlying model.

First, we agree with Donoghue and Edwards (2019) in that a “rate of 42%” has very little, if any, meaning. A percentage is, in fact, not a rate at all. As we (e.g., Beaulieu and O'Meara, 2015) and others (e.g., Hansen, 1997) have argued, rates have units (e.g., state transitions per million years, speciation events per million years, accumulated variance per million years) and thus should be discussed in those terms. We acknowledge, however, that units are often glossed over in simulations because the scale of the tree may also lack units (e.g., branch lengths in millions of years). In any event, knowing whether a particular trait (e.g., selfing) has evolved *42 times per million years of evolutionary history*, or one hundredth of that, suggests what has driven the persistence of the different character states. For example, transition rates among a coordinated set of floral traits were low enough that it likely took tens of millions of years for flowers to evolve a diversity-accelerating combination of bilateral symmetry, few stamens, and showy petals (O'Meara, Smith et al., 2016). And, as a result, angiosperms as a whole are not yet at equilibrium with respect to the diversity of their floral trait combinations. Such a discovery comes from understanding transition rates with units, as well as looking at a wide diversity of angiosperms.

Rates can have more intuitive appeal when they are transformed, such as taking the reciprocal to express a rate in units of the expected wait time. A rate of selfing evolution of 42/Myr means a single lineage is expected to change whether it transitions to selfing every 0.024 Myr, or every 24,000 years, which is intuitively a very high rate. The same can be done with diversification rates to give expected times between speciation and extinction events. For example, the highest diversification rate across all life (in table 1a of Lagomarsino et al., 2016) is 3.07 diversification events/Myr, which means every species is expected to speciate into two every 325,000 years—a high rate, but not unreasonable for a recent, presumably ongoing, radiation.

Of course, we emphasize that *transition rates are not the only way to understand biology*. Correlations between traits, detailed studies of ontogeny, understanding phylogenetic relationships, and so forth, are obviously important. Our use of a transition rate as an example in a simulation is different from arguing that the goal of comparative biology is to define processes purely by a single “global rate”. We regret that our intention was apparently unclear. Central to evolutionary biology is to understand the myriad patterns and processes that produced the extraordinary diversity of life on Earth. Sometimes the focus is on particular, unique events (e.g., the evolution of the carpel or what happened after tarweeds

made it to Hawaii), while other questions are on more general scales (e.g., how complex traits evolve, how species diversify in novel habitats). We argue that a continued push is needed to understand these processes on a more general level, while also learning from the exceptions. If one wants to go beyond natural history of a particular group to see how some observable phenomenon works in practice, constructing a model is an important next step—models help explain the world and assess whether predictions from models match reality. Otherwise, we are left with “just-so stories” (Gould, 1978). In our case, we chose to focus on transition rates to illustrate our point. Think of the wide variety of ways transitions in mutation rate, selfing rate, dispersal rate, extinction rate, rate of polyploidization, rate of gain or loss of woodiness or C_4 photosynthesis have shaped biological diversity. Transition rates also tend to be heritable, but with important considerations for heterogeneity across time and taxa, which has been a focus of our research (e.g., O’Meara et al., 2006; Beaulieu et al., 2012, 2013a; O’Meara, 2012; Beaulieu and O’Meara, 2016; O’Meara, Smith et al., 2016; Caetano et al., 2018). Again, the use of transition rates is just one example of a model parameter—not the ultimate goal of biologists.

ANCESTRAL STATE ESTIMATION AS AN ALTERNATIVE

Donoghue and Edwards (2019) reject the utility of models on large trees for anything but identifying focal clades and in essence promote closely examining the details of specific individual transitions observed for a focal trait. Looking at focal transitions has led to many discoveries, and we agree that it is one important way, among others, to understand biology. However, there are three important caveats readers should note for the “examine multiple transitions” approach advocated by Donoghue and Edwards (2019). First, a state transition at some point can only ever be understood through some form of an ancestral state reconstruction—that is, estimating where exactly a character switched from one state to another. A given transition might be so rare and such an obvious change that an implicit parsimony map will suffice. More rigorous approaches to get the same mapping include maximum parsimony, stochastic character mapping, or marginal or joint reconstruction from maximum likelihood or Bayesian approaches. Nevertheless, some sort of reconstruction must have been done to know where on the tree the transitions occurred, so ancestral state estimation still relies on a model. While ancestral state reconstruction methods are fraught with difficulties (e.g., Schluter et al., 1997; Cunningham et al., 1998; Cunningham, 1999; Omland, 1999; Oakley et al., 2000; Salisbury and Kim, 2001; Mossel, 2003; Lucena and Haussler, 2005; Mossel and Steel, 2005; Goldberg and Igic, 2008; Li et al., 2008, 2010; Losos, 2011; Royer-Carenzi et al., 2013; Gascuel and Steel, 2014), they will always remain a tempting enterprise. But, the important point is that, aside from parsimony (though only arguably, given, for example, the model proposed by Tuffley and Steel, 1997), ancestral state reconstruction algorithms use transition rates to produce their estimates, and so state changes cannot be understood without them.

The second caveat, which is related to the first, is that by moving the primary focus to the reconstructions, we ask quite a lot of the data. Consider a 100-taxon tree: To estimate transitions, one is technically using a single value from each of the 100 species to infer not just transition rates for the model, but also the likeliest states at 99 nodes. There is also an interaction between the transition rates and ancestral state reconstructions that has been largely ignored

TABLE 1. Partial list of ascertainment biases affecting comparative methods studies. There are corrections for survivorship bias and clade variability (see text), but impact and corrections for the other biases remain an open research question.

Bias	Description
Survivorship	When extinction is present, the probability that the focal clade should have survived to the present is less than 1. Focal clades will also often appear more diverse than expected based on the set of rates that produced them, in the same way that a study of gamblers who have yet to go bankrupt would suggest their odds of winning are better than expected, unless accounted for.
Variable clades	Does the focal clade have enough variation in the trait(s) of interest? For discrete traits, variability may be the ratio of states across taxa or the number of state changes on the tree; for continuous traits, variability is more heuristic. For studying adaptation to freezing, examining variable groups is important, but groups that never evolve to tolerate a hard freeze may also provide insights.
Named-based	A focal group is chosen, intentionally or not, because it has a formal name—i.e., a genus, a family, or even a group without a rank but with a particular name. There are millions of possible clades to name, but only some are morphologically distinct enough, and with sufficient diversity, to warrant a formal name.
Practicality	Is a study feasible—i.e., for a student dissertation, can the student make a tree for this clade that might be sufficiently powerful to answer the question at hand?
Charismatic clades	Certain groups are evidently more attractive for study than others, which affects the amount of information available for them and potential impact of papers.
“Organism first”	The focal clade is chosen first, and the research question is developed later, based on perceived unique or important features of the group.

by the field. As character change becomes increasingly labile, the underlying Markov process actually makes it increasingly difficult to infer ancestral states accurately (Sober and Steel, 2002). In the absence of fossil data from right around the transition, we may also make the implicit assumption that descendants differing in the focal trait have not undergone state change in any other trait since then. Finally, if we want to see what additional factors affect individual transitions of our focal traits, we can still use models. For example, the method of Pagel (1994) might be a good approach for testing whether high herbivore pressure leads to higher rates of becoming deciduous (but see Maddison and FitzJohn, 2015—there are significant problems with this approach).

ASCERTAINMENT BIAS IS A GENUINE PROBLEM

In Beaulieu and O'Meara (2018), we expressed our concerns about ascertainment bias in general. That is, perhaps by focusing on certain clades, those with variation in a trait of interest, we are often misled about general processes because we only look at peculiar subsets of life. The example we used was to focus on variable clades only (but this is not the only source of ascertainment bias, as we expand upon below and in Table 1). Donoghue and Edwards (2019) responded that ascertainment bias is not a problem in practice. First, to suggest that ascertainment bias is not a serious problem is to ignore all the ways in which it is already accounted for in many statistical and analytical tools. For example, the issue of ascertainment bias is a well-known problem in phylogenetic inference. Felsenstein (1992) first noted that when certain restriction sites are absent from all species, the entire site is omitted from the data matrix. The data set, now biased and unrepresentative because it contains only variable sites, can potentially inflate inferred branch lengths and generate erroneous trees. Felsenstein (1992) proposed a simple but clever modification to the likelihood calculation at a site. This modification was later adopted by Lewis (2001) in phylogenetic inference based on morphological matrices, which also often omit invariant characters (e.g., “cell wall present” is never a trait used in plant phylogenetics). Similar biases and associated modifications to the underlying likelihood calculations have been proposed for single nucleotide polymorphisms (Leaché et al., 2015) and restriction-site-associated DNA-restriction data sets (Peterson et al., 2012).

Correcting for ascertainment bias is also central to properly estimating diversification rates (i.e., speciation and extinction). With nonzero extinction rates, there is a probability that a clade observed today could have, alternatively, gone completely extinct at some point in the past. The effect is that the clades that have survived to the present often get off to a running start initially and are, therefore, part of a distribution of clades whose diversity is not a representative sample of the typical clade sizes for a given set of speciation and extinction parameters (Nee et al., 1994; Magallón and Sanderson, 2001). A simple way of thinking about this type of ascertainment bias is to imagine a busload of patrons arriving at a casino, each clutching a shiny quarter. The patrons still playing several hours later must have had an unusually good run of luck at the beginning of the night, even though the chances of winning were the same for everyone. In the same way, not accounting for survival probability can produce spurious estimates of the diversification process.

If it were just diversification rates being uniformly biased up or down, that would not be ideal, but it might be acceptable (see Rabosky et al., 2017 on this point). However, this is not the case, and

in fact, we can be misled as to which clades are diversifying faster if we choose not to correct for the ascertainment bias. We can demonstrate the problem through a simple simulation (scripts available from the Dryad Digital Repository: <https://doi.org/10.5061/dryad.n9c0c8m>). We generated 50 pairs of clades under a birth–death process, each with a different set of speciation and extinction rates (though with extinction never exceeding speciation), that were terminated after 20 Myr, so that we have 50 comparisons between sister groups of the same age but different *true* diversification rates. A common question is which of the two clades has a higher net diversification rate. Choosing at random, we would be right 50% of the time; ignoring ascertainment bias, we would be right 68% of the time, but incorporating it properly, we would be right 81% of the time. Clearly, it is better to incorporate ascertainment bias. Also note that this is just one simulation, and understanding ascertainment bias in the context of diversification is a tricky problem that is still unresolved (see Stadler, 2013 for more details). Nevertheless, our demonstration argues against dismissing it as unimportant.

In Beaulieu and O'Meara (2018), one intention was to highlight the many additional types of ascertainment bias at work in most comparative analyses, each capable of producing data sets that are not representative of the overall evolutionary process. We focused on ascertainment bias of clades that exhibit variable characters, but it is worth expanding on the various kinds of ascertainment bias affecting evolutionary studies (Table 1). For instance, it is much easier to communicate evolutionary patterns if the focal group is named, like “angiosperms” (Angiospermae; Cantino et al., 2007), instead of, say, the clade that excludes *Amborella* and Nymphaeales. The name “angiosperms” is immediately recognizable because it denotes a morphologically distinct and seemingly more natural group of plants. The existence of a name has the unintended consequence of driving research questions that are motivated simply by the synapomorphies that define a named group (see Smith et al., 2011 for a more thorough discussion of this topic). There are additional biases that reflect other practicalities of choosing a clade for study. Clades are often chosen because they are tractable, neither too big nor too small, already have a fair bit known about them, and are often times geographically convenient. Sometimes clades are chosen because of their apparent charisma (e.g., silverswords) or because they seem atypical with respect to their observed trait variation. Extrapolating from these particular clades to processes operating more generally in less compelling clades may be problematic. We also worry that this issue is compounded by the fact that only a portion of possible clades are examined closely, while others are ignored completely. For example, we ran an analysis of the most recent 1000 papers published in the *American Journal of Botany* to extract any plant genera discussed anywhere in any paper (genus is the finest scale reliably retrieved, as tools involved cannot yet always identify *Q. rubra* as *Quercus rubra*; see Appendix S1 for analysis details). There is strong phylogenetic signal in which clades were even mentioned (Pagel's λ [Pagel, 1999] was 0.93), showing that botanical interests are clumped and many key groups are relatively unexamined.

DO WE EXTRAPOLATE?

Donoghue and Edwards (2019) contend that extrapolations from specific clades to broader ones are not general practice. We disagree. In fact, we argue that it is standard practice for researchers to contextualize their findings by showing whether they reinforce

or depart from general trends—that is, ones broadly shared across diverse species or ecosystems. A desire to uncover general patterns, and whether those patterns were generated by shared or divergent mechanisms, is the primary motivator of most scientific research. For example, Edwards et al. (2017) contextualized their detailed study of 20 Asian forest species sampled across *Viburnum*, combined with a broader *Viburnum* comparative study of 120 species, against the backdrop of biome assembly, suggesting that the relatively balanced proportion of deciduous and evergreen species they examined “may explain the massive convergence of adaptive strategies that characterizes the world’s biomes.” Even though *Viburnum* does not occur in the tundra, deserts, or deep ocean, a detailed study of their evolution may indeed help us understand evolution in these biomes. Such extrapolation is useful, as it generates predictions to test elsewhere and helps formulate broader principles of how traits and biogeography interact on evolutionary time scales. On the other hand, Beaulieu et al. (2013b), contrasted their results of a southern hemisphere origin of campanulid angiosperms and potential for Gondwana vicariance, with the findings of the many small-scale studies that suggested the break-up of Gondwana was not an important event for angiosperms as whole (e.g., Sanmartín and Ronquist, 2004). Interestingly, this prior generalization came from an aggregation of smaller studies of groups that were far too young, and not at the right phylogenetic scale, for this type of question—the question actually required a different approach and a much broader scale.

Science thrives on bold ideas flung out into the world, so using a charismatic clade to make predictions about larger groups is important. However, Donoghue and Edwards (2019) seem to advocate the immoderate view that studies at larger scales are not very relevant for understanding evolution, explicitly rejecting a consensus view that studies at different scales can be complementary and valuable. In their view, “large” studies can merely identify patterns, while “small” studies can help identify mechanisms. (Note that there is no guidance on the location of the cutoff, though studies of 20 species [Edwards et al., 2017] are considered small.) We strongly object to this view. The motivation for the simulations presented in Beaulieu and O’Meara (2018) was to simply investigate the value, as well as the potential costs, of large-scale studies in dealing with many of the biases in comparative biology. We expand on these biases here (Table 1). For instance, examining larger, comprehensive clades will naturally include relatively obscure and unstudied groups often overlooked in favor of compelling, charismatic clades. Large-scale studies also provide the context by which to judge what may be exceptional at smaller scales, and they allow for exploration of patterns and for identifying locations of important changes that may or may not correspond to any formally named group. Stating that studies of different scales can all be important and that appropriate scale can relate to what questions are being addressed would normally be thought of as banalities, except for the persistent view, expressed by Donoghue and Edwards (2019) among others, about the primacy of small scale studies.

CONCLUSIONS

We need a variety of approaches, at a diversity of scales, to truly understand evolution. There are problems with large-scale studies, but there are problems with small-scale studies, too. The type of analyses at one scale may not be useful at another. For example, at larger scales, one may have to reduce biological diversity to

discrete traits with a single value for a species, whereas at a smaller scale one can embrace the variation across individuals that powers so much of evolution, at the cost of power for detecting variation across species. There has been salutary attention paid to the potential pitfalls of working with large phylogenies, but very little attention given to limits of studies of small clades. Of course, it is worth bearing in mind that the “large” vs. “small” distinction is artificial. Once one is applying comparative methods on a set of taxa to understand biology, there is a smooth continuum from a study of 20 species to one of 20,000. As scientists, we all spend our time at the margin of what is known and unknown, trying to expand the illuminated area of knowledge. Some use candles, some spotlights, and some use lasers, which we think makes for a remarkably effective glow overall.

ACKNOWLEDGEMENTS

We thank Michael Donoghue for the great conversations over the years, starting when we were students, which helped form many of the ideas presented here; discussions with Erika Edwards have also proved very insightful. We are grateful to Pam Diggle for giving us the opportunity to respond and to Stacey Smith and an anonymous reviewer for their constructive feedback. We also thank Andrew Alverson, Teo Nakov, and members of the Beaulieu and O’Meara labs for the thoughtful discussions related to these topics.

DATA ACCESSIBILITY

The data and all scripts associated with this article are available from the Dryad Digital Repository: <https://doi.org/10.5061/dryad.n9c0c8m> (Beaulieu and O’Meara, 2019).

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

LITERATURE CITED

- Beaulieu, J. M., and B. C. O’Meara. 2015. Extinction can be estimated from moderately sized molecular phylogenies. *Evolution* 69: 1036–1043.
- Beaulieu, J. M., and B. C. O’Meara. 2016. Detecting hidden diversification shifts in models of trait-dependent speciation and extinction. *Systematic Biology* 65: 583–601.
- Beaulieu, J. M., and B. C. O’Meara. 2018. Can we build it? Yes we can, but should we use it? Assessing the quality and value of a very large phylogeny of campanulid angiosperms. *American Journal of Botany* 105: 417–432.
- Beaulieu, J. M., and B. C. O’Meara. 2019. Data from: Diversity and skepticism are vital for comparative biology: a response to Donoghue and Edwards (2019). Dryad Digital Repository. <https://doi.org/10.5061/dryad.n9c0c8m>.
- Beaulieu, J. M., D.-C. Jhwhueng, C. Boettiger, and B. C. O’Meara. 2012. Modeling stabilizing selection: Expanding the Ornstein–Uhlenbeck model of adaptive evolution. *Evolution* 66: 2369–2383.
- Beaulieu, J. M., B. C. O’Meara, and M. J. Donoghue. 2013a. Identifying hidden rate changes in the evolution of a binary morphological character: the

- evolution of plant habit in campanulid angiosperms. *Systematic Biology* 62: 725–737.
- Beaulieu, J. M., D. C. Tank, and M. J. Donoghue. 2013b. A Southern Hemisphere origin for the campanulid angiosperms, with traces of the break-up of Gondwana. *BMC Evolutionary Biology* 13: 80.
- Caetano, D. S., B. C. O'Meara, and J. M. Beaulieu. 2018. Hidden state models improve the adequacy of state-dependent diversification, including biogeographical models. *Evolution* 72: 2308–2324.
- Cantino, P. D., J. A. Doyle, S. W. Graham, W. S. Judd, R. G. Olmstead, D. E. Soltis, P. S. Soltis, et al. 2007. Towards a phylogenetic nomenclature of Tracheophyta. *Taxon* 56: E1–E44.
- Cunningham, C. W. 1999. Some limitations of ancestral character-state reconstruction when testing evolutionary hypotheses. *Systematic Biology* 48: 665–674.
- Cunningham, C. W., K. E. Omland, and T. H. Oakley. 1998. Reconstructing ancestral character state: a critical reappraisal. *Trends in Ecology and Evolution* 13: 361–366.
- Donoghue, M. J., and E. J. Edwards. 2019. Model clades are vital for comparative biology, and ascertainment bias is not a problem in practice: a response to Beaulieu and O'Meara. *American Journal of Botany* 106: 327–330.
- Edwards, E. J., D. S. Chatelet, B. Chen, J. Y. Ong, S. Tagane, J. Kanemitsu, K. Tagawa, et al. 2017. Convergence, consilience, and the evolution of temperate deciduous forests. *American Naturalist* 190: S87–S104.
- Felsenstein, J. 1992. Phylogenies from restriction sites: a maximum-likelihood approach. *Evolution* 46: 159–173.
- Gascuel, O., and M. Steel. 2014. Predicting the ancestral character changes in a tree is typically easier than predicting the root state. *Systematic Biology* 63: 421–435.
- Goldberg, E. E., and B. Igic. 2008. On phylogenetic tests of irreversible evolution. *Evolution* 62: 2727–2741.
- Gould, S. J. 1978. Sociobiology: the art of storytelling. *New Scientist* 80: 530–533.
- Hansen, T. F. 1997. Stabilizing selection and the comparative analysis of adaptation. *Evolution* 51: 1341–1351.
- Lagamarsino, L. P., F. L. Condamine, A. Antonelli, A. Mulch, and C. C. Davis. 2016. The abiotic and biotic drivers of rapid diversification in Andean bellflowers (Campanulaceae). *New Phytologist* 210: 1430–1442.
- Leaché, A. D., B. L. Banbury, J. Felsenstein, A. Nieto-Montes de Oca, and A. Stamatakis. 2015. Short tree, long tree, right tree, wrong tree: new acquisition bias corrections for inferring SNP phylogenies. *Systematic Biology* 64: 1032–1047.
- Lewis, P. O. 2001. A likelihood approach to estimating phylogeny from discrete morphological character data. *Systematic Biology* 50: 913–925.
- Li, G., M. Steel, and L. Zhang. 2008. More taxa are not necessarily better for the reconstruction of ancestral character states. *Systematic Biology* 57: 647–653.
- Li, G., J. Ma, and L. Zhang. 2010. Greedy selection of species for ancestral state reconstruction on phylogenies: elimination is better than insertion. *PLoS ONE* 5: e8985.
- Losos, J. B. 2011. Seeing the forest for the trees: the limitations of phylogenies in comparative biology. *American Naturalist* 177: 709–727.
- Lucena, B., and D. Haussler. 2005. Counterexample to a claim about the reconstruction of ancestral character states. *Systematic Biology* 54: 693–695.
- Maddison, W. P., and R. G. FitzJohn. 2015. The unsolved challenge to phylogenetic correlation tests for categorical characters. *Systematic Biology* 64: 127–136.
- Magallón, S., and M. J. Sanderson. 2001. Absolute diversification rates in angiosperm clades. *Evolution* 55: 1762–1780.
- Mossel, E. 2003. On the impossibility of reconstructing ancestral data and phylogenies. *Journal of Computational Biology* 10: 669–678.
- Mossel, E., and M. Steel. 2005. How much can evolved characters tell us about the tree that generated them? In O. Gascuel [ed.], *Mathematics of evolution and phylogeny*, 384–412. Oxford University Press, NY, NY, USA.
- Nee, S., R. M. May, and P. H. Harvey. 1994. The reconstructed evolutionary process. *Philosophical Transactions of the Royal Society of London, B, Biological Sciences* 344: 305–311.
- Oakley, T. H., and C. W. Cunningham. 2000. Independent contrasts succeed where ancestor reconstruction fails in a known bacteriophage. *Evolution* 54: 397–405.
- O'Meara, B. C. 2012. Evolutionary inferences from phylogenies: a review of methods. *Annual Review of Ecology, Evolution, and Systematics* 43: 267–285.
- O'Meara, B. C., C. Ané, M. J. Sanderson, and P. C. Wainwright. 2006. Testing for different rates of continuous trait evolution using likelihood. *Evolution* 60: 922–933.
- O'Meara, B. C., S. D. Smith, W. S. Armbruster, L. D. Harder, C. R. Hardy, L. C. Hileman, L. Hufford, et al. 2016. Non-equilibrium dynamics and floral trait interactions shape extant angiosperm diversity. *Proceedings of the Royal Society, B, Biological Sciences* 283: 20152304.
- Omland, K. E. 1999. The assumptions and challenges of ancestral state reconstructions. *Systematic Biology* 48: 604–611.
- Pagel, M. 1994. Detecting correlated evolution on phylogenies: a general method for the comparative analysis of discrete characters. *Proceedings of the Royal Society, B, Biological Sciences* 255: 37–45.
- Pagel, M. 1999. Inferring the historical patterns of biological evolution. *Nature* 401: 877–884.
- Peterson, B. K., J. N. Weber, E. H. Kay, H. S. Fisher, and H. E. Hoekstra. 2012. Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS ONE* 7: e37135.
- Rabosky, D. L., J. S. Mitchell, and J. Chang. 2017. Is BAMM flawed? Theoretical and practical concerns in the analysis of multi-rate diversification models. *Systematic Biology* 66: 477–498.
- Royer-Carenzi, M., P. Pontarotti, and G. Didier. 2013. Choosing the best ancestral state reconstruction method. *Mathematical Biosciences* 242: 95–109.
- Salisbury, B. A., and J. Kim. 2001. Ancestral state estimation and taxon sampling density. *Systematic Biology* 50: 557–564.
- Sanmartín, I., and F. Ronquist. 2004. Southern hemisphere biogeography inferred by event-based models: plant versus animal patterns. *Systematic Biology* 53: 216–243.
- Schluter, D., T. Price, A. Ø. Mooers, and D. Ludwig. 1997. Likelihood of ancestor states in adaptive radiation. *Evolution* 51: 1699–1711.
- Smith, S. A., J. M. Beaulieu, A. Stamatakis, and M. J. Donoghue. 2011. Understanding angiosperm diversification using small and large phylogenetic trees. *American Journal of Botany* 98: 404–414.
- Sober, E., and M. Steel. 2002. Testing the hypothesis of common ancestry. *Journal of Theoretical Biology* 218: 395–408.
- Stadler, T. 2013. How can we improve accuracy of macroevolutionary rate estimates? *Systematic Biology* 62: 321–329.
- Tuffley, C., and M. Steel. 1997. Links between maximum likelihood and maximum parsimony under a simple model of site substitution. *Bulletin of Mathematical Biology* 59: 581–607.