Book Review

Modern Phylogenetic Comparative Methods and Their Application in Evolutionary Biology: Concepts and Practice. — Edited by László Zsolt Garamszegi. Heidelberg: Springer, 2014. xv+554 pp. ISBN 978-3-662-43549-6 $119, £64 95 (hardback). ISBN 978-3-662-43550-2 $90 £81 75 (e-book).

As a result of the process of descent with modification, closely related species share many traits. Phylogenies thus provide information that not only needs to be considered when making inter-specific comparisons but which also can be leveraged to gain insight into macroevolutionary questions. Statistical approaches for utilizing this information, called phylogenetic comparative methods (PCMs), have grown tremendously in the past few decades. The new book, *Modern Phylogenetic Comparative Methods and Their Applications in Evolutionary Biology*, reviews many of these developments.

The book has 22 chapters, arranged in three sections, mostly consisting of reviews of well-trodden topics in comparative biology. These include: phylogenetic regression; fitting models to both continuous and discrete characters; community and coevolutionary phylogenetic methods; as well as some more general chapters on preparing data for inference, tree building, simulating and plotting trees and traits, etc. The primary emphasis of the book, however, is on the mechanics of fitting trait evolutionary models.

While a bit inconsistent in some parts (as is typical of edited collections), overall the various authors do a commendable job of explaining the many statistical nuances and mathematical tricks involved in model fitting and summarizing a growing and often intimidatingly dense body of literature. Very little in here is new (only Nunn and Zhu, in Chapter 21, deviate from this script, and present a novel and intriguing approach for investigating "evolutionary singularities"), but I think the book is a useful resource for both experienced hands and newcomers to the field – though frankly, given the steep price of the book and the fact that it is mostly a review of previously published work, I would recommend checking it out of a library rather than purchasing a copy for one's desk.

However, the authors' focus is almost exclusively on statistical issues. On reading through this collection, I cannot help but recall a sentiment expressed by Houle et al. (2011) in their lucid review of measurement theory and its applications in biology. They criticize statisticians who advocate that data transformations are justifiable whenever they result in distributions that meet the assumptions of a particular analysis: "If that is statistics, we want no part of it, as science is about nature, not numbers" [p. 18].

I thus find the book's perspective limiting, and wish that greater attention was given to the more interesting (and challenging) question of interpretation: what exactly we are inferring when these models are fit to data? Consider for example, regression models of the form:

$$Y = \beta_0 + \beta_1 X + \varepsilon.$$

In phylogenetic regression, it is usually assumed that the tree only enters into the model in the error term ε such that $\varepsilon \sim N(\mathbf{0}, \mathbf{V})$ where $\mathbf{V}$ is the expected variance-covariance matrix for the traits given an evolutionary model. In other words, the evolutionary model is used to model the structure of the residuals and not the actual traits. As discussed in no fewer than nine out of the 22 chapters of the book, formulating the model in such a way allows us to make use of well-established statistical theory from generalized least squares (GLS) and generalized linear mixed-effects (GLM) models. Including the phylogenetic structure in the error variance is no different from including any other type of covariance. By recognizing this equivalence, we can now fit phylogenetic regression models with a variety of distributions for the response variable Y [Ives and Garland, Chapter 9; Villemereuil and Nakagawa, Chapter 11], incorporate measurement error [Garamszegi, Chapter 7], perform model averaging [Garamszegi and Mundry, Chapter 12] and path-analysis [Gonzalez-Voyer and von Hardenberg, Chapter 8], identify outliers [Nunn and Zhu, Chapter 21], and use standard model diagnostics [Mundry, Chapter 6].

A number of the authors suggest that a λ tree transformation (Freckleton et al. 2002) is often more appropriate than simply assuming Brownian motion (BM) for constructing the error variance term $\mathbf{V}$. (The λ transformation involves multiplying the off-diagonals of $\mathbf{V}$ by an estimated parameter between 0 and 1.) This is a purely phenomenological construct — by shrinking every branch except those leading to the tips, it implies that there is something special about extant taxa, which is clearly not the case. Nonetheless, researchers (including the authors of the current volume) often use such models to claim that one trait is adapted to the value of another. In Chapter 14, Hansen clearly articulates (recapitulating arguments he has made elsewhere; see Hansen and Orzack 2005), that these types of models do

not actually capture the process of adaptation at all: "any adaptive process that is sufficiently slow to generate a phylogenetic signal in model residuals will also generate systematic deviations from the optimal state" [p. 360]. Effectively, standard regression models assume that adaptation to a new environment is instantaneous, and that maladaption is phylogenetically structured — closely related species will have similar deviations from the optimal trait value even if the optimum differs between them. From a biological perspective, this seems very odd.

Perhaps even more confusing is the use of Ornstein-Uhlenbeck (OU) models to construct the error variance term. OU is attractive for modeling the residual variance because, unlike the λ transformation, it is a coherent stochastic process and is directly analogous to a population level model from quantitative genetics — quadratic stabilizing selection on a fixed adaptive landscape (Lande 1976; Hansen and Martins 1996). While the λ transformation is obviously just a statistical construct, OU *seems* to be biologically motivated. Indeed, a number of authors suggest that including an OU error variance captures "constraints" [Paradis, p. 9], "stabilizing selection" [Ives and Garland, p. 234], or "selective regimes" [Symonds and Blomberg, p. 122]; but this does not get around Hansen's criticisms. These models still assume phylogenetically structured maladapation, and they do not allow researchers to make specific inferences about stabilizing selection or evolutionary constraints. OU error structures may often fit data better than BM error structures, but it is likely that this is simply because OU can accommodate more variance towards the tips of the phylogeny than a BM model can (including λ has a similar effect). The evolutionary argument here seems to be merely window dressing for a purely statistical argument.

OU models are further treated in depth in three different chapters. Each of these chapters [Hansen, Chapter 14; O'Meara and Beaulieu, Chapter 15; Mahler and Ingram, Chapter 18] offers an interpretation as to what the parameters of an OU model actually represent. The differences between them are nuanced (and I will not dissect them here), but importantly they all share the perspective that a simple quantitative genetics explanation — i.e., clade-wide stabilizing selection where some species are further from the optima than are others — is almost certainly naïve and unreasonable. Rather, OU models likely reflect in some way the structure and dynamics of the macroevolutionary adaptive landscape (Simpson 1944; Arnold et al. 2001; Hansen 2012), upon which lie population-level adaptive landscapes.

How are we to reconcile these different uses and interpretations of the same core models, and make sense of comparative analyses? In my view, there are three possible frameworks with which to think about comparative biology. First, we can take the view that what we are measuring are strictly patterns, and that we are not necessarily making inferences about specific evolutionary processes. This is certainly a defensible position: the patterns may be interesting in and of themselves, and documenting commonalities and differences among clades and through time may provide a broader picture of the history of life on earth. In practice, this is what researchers are often actually doing, even if they are hesitant to admit it. A benefit of openly adopting this perspective is that we can consider a much broader suite of models that may provide a much better fit to our data and more predictive power than current models — if we are not interested in making specific evolutionary inferences, then we need not be beholden to specific evolutionary models. Such alternatives may include macroevolutionary diffusion processes (e.g., Clauset and Erwin 2008), models derived from macroecological theories, or making use of statistical learning approaches divorced from any process whatsoever.

The second framework is the quantitative genetics view: the models we fit in comparative biology should be taken as literally representing microevolutionary hypotheses. Many of the commonly used models can be directly interpreted in terms of population-level parameters (Hansen and Martins 1996; Pennell and Harmon 2013). We can compare the estimated model parameters to within-population measures, in order to test whether macroevolutionary divergences are consistent with evolution by drift, stabilizing selection, etc. This project is certainly interesting and worth pursuing. But given the results of studies that have explicitly examined this connection using rather simple models (Lynch 1990; Estes and Arnold 2007; Hohenlohe and Arnold 2008), it appears that translating the parameters estimated from comparative data to the terms of quantitative genetics (i.e., if we assume that BM is strictly a model of drift, the estimated rate parameter σ^2 is equal to the additive genetic variance G divided by the effective population size N_e) will often result in nonsensical numbers.

The third perspective is to take seriously the idea that macroevolutionary models reflect the dynamics of adaptive landscapes through deep time (Arnold et al. 2001; Hansen 2012). This is in line with the views of chapter authors Hansen, O'Meara & Beaulieu, and Ingram & Mahler. Comparative biologists have a tendency to discuss many of these ideas in quotation marks. The optimum of OU models is referred to as "clade level optimum". A model with decelerating rates of change depicts an "early burst". I argue that a much richer and more meaningful connection can potentially be made. Theoretical work over the last century has produced a beautiful and fairly comprehensive understanding of how populations move across adaptive landscapes, and empiricists have tested the theoretical predictions in a wide variety of systems and contexts. In contrast, we have only a preliminary understanding of how the landscapes themselves evolve

at longer time scales. This is a fundamentally important question in evolutionary biology, and one which I believe phylogenetic comparative biology and paleobiology can help address.

There is a lot of work to be done before we will really be able to get at these types of questions. Once we recognize that some of the classic concepts in evolutionary biology — such as adaptive zones, adaptive radiations and key innovations — are actually hypotheses about the structure and dynamics of adaptive landscapes, we can start developing statistical models that actually capture their essential properties. Current models are, at best, loosely tied to these ideas (and hence the scare quotes). Additionally, there are a number of existing mathematical frameworks that make predictions about these higher-order processes and trait evolution over longer time periods (see e.g., Gavrilets 2004; Doebeli 2011); but there is currently no way to estimate the relevant parameters of these models from comparative data.

Both the development of new PCMs and the interest in using them has grown tremendously over the past decade. Nevertheless, I feel that we, as a field, are somewhat stuck. First, the same handful of statistical models are employed over and over again, with most of the progress representing relatively minor variations on similar themes (that is not to say that such improvements are not challenging or worthwhile). Second, we are often much too vague about what exactly we want to explain with PCMs — this is apparent in both this current book collection and in the literature more broadly. I argue that these two problems are deeply intertwined. The standard collection of models available today, namely those based on BM and OU, have had such staying power in part because they can be useful for detecting patterns, can be interpreted in light of evolutionary genetics, and can loosely be tied to questions about adaptive landscapes. Requiring this sort of conceptual flexibility is also a limitation. More focused, question-specific approaches to modeling that are directly tied to the inferences we actually want to make will likely get us much further than sticking to models that are more general but address no questions particularly well.

REFERENCES

Arnold S.J., Pfrender M.E., Jones A.G. 2001. The adaptive landscape as a conceptual bridge between micro- and macroevolution. Genetica 112:9–32.

Clauset, A., Erwin D.H. 2008. The evolution and distribution of species body size. Science 321:399–401.

Doebeli M. 2011. Adaptive diversification. Princeton, NJ: Princeton University Press.

Estes S., Arnold S.J. 2007. Resolving the paradox of stasis: models with stabilizing selection explain evolutionary divergence on all timescales. Amer. Nat. 169:227–244.

Freckleton R.P., Harvey P.H., Pagel M. 2002. Phylogenetic analysis and comparative data: a test and review of evidence. Amer. Nat. 160:712–726.

Gavrilets S. 2004. Fitness landscapes and the origin of species. Princeton, NJ: Princeton University Press.

Hansen T.F. 2012. Adaptive landscapes and macroevolutionary dynamics. In: Svensson E., Calsbeek R., editors. The adaptive landscape in evolutionary biology. Oxford: Oxford University Press, pp. 205–221.

Hansen T.F., Martins E.P. 1996. Translating between microevolutionary process and macroevolutionary patterns: the correlation structure of interspecific data. Evolution 50:1404–1417.

Hansen T.F., Orzack S.H. 2005. Assessing current adaptation and phylogenetic inertia as explanations of trait evolution: the need for controlled comparisons. Evolution 59:2063–2072.

Hohenlohe P.A., Arnold S.J. 2008. MIPoD: A hypothesis-testing framework for microevolutionary inference from patterns of divergence. Amer. Nat. 171:366–385.

Houle D., Pelabon C., Wagner G.P., Hansen T.F. 2011. Measurement and meaning in biology. Quart. Rev. Biol. 86:3–34.

Lande R. 1976. Natural selection and random genetic drift in phenotypic evolution. Evolution 30:314–334.

Lynch M. 1990. The rate of morphological evolution in mammals from the standpoint of the neutral expectation. Amer. Nat. 136:727–741.

Pennell M.W. Harmon L.J. 2013. An integrative view of phylogenetic comparative methods: connections to population genetics, community ecology, and paleobiology. Ann. New York Acad. Sci. 1289:90–105.

Simpson G.G. 1944. Tempo and mode in evolution. New York: Columbia University Press.

Matthew W. Pennell, Institute for Bioinformatics & Evolutionary Studies, University of Idaho, Moscow, ID 83844, U.S.A; E-mail: mwpennell@gmail.com