Editor’s note on ‘Putting fossils in trees’ special issue

April M. Wright1,2

1Ecology, Evolution and Organismal Biology, Iowa State University, Ames, IA 50010, USA
2The Field Museum, Chicago, IL 60605, USA

This special feature has its origins with a 2014 symposium organized by myself and my co-editors David W. Bapst, Graeme T. Lloyd and Nicholas J. Matzke at the Society for Vertebrate Paleontology meetings in Berlin, Germany. In the years prior to 2014, there had been several interesting and important papers published [1–4] about estimating phylogenetic trees incorporating fossil taxa, and particularly involving Bayesian analyses. Some of these papers were written from a neontological perspective, and appeared in journals not often read by palaeontologists. We convened our session with a simple idea: to get different types of researchers from both neontology and palaeontology together to share work and ideas concerning phylogenetic analyses in palaeontology. The session’s success inspired my co-editors and I to propose a special feature for Biology Letters, to which we invited the session participants and other contributors. The Biology Letters short format is perfect for focused vignettes about the actual practice of phylogenetics with fossils.

For years, those working primarily with fossil data, and those working primarily with molecular data, took different approaches to both building phylogenetic trees, and scaling them to absolute time. Palaeontologists have often applied a maximum parsimony paradigm under which potential trees are scored according to how many character changes are required in the dataset to produce that tree. The tree implying the fewest changes being ultimately preferred, returning either the single most parsimonious tree or a sample of equally parsimonious trees. Neontologists, working primarily with molecular data, more often apply likelihood-based methods, including Bayesian methods, which evaluate phylogenetic trees by applying a model of evolution to the data, and calculating the likelihood of the data given the tree and the model. Maximum likelihood methods return a point estimate of the phylogeny with branch lengths equal to the expected number of substitutions per site along that branch. Bayesian methods allow for researchers to quantify and incorporate their existing knowledge about phylogenetic relationships and model parameters using probability distributions called priors. Under Bayesian methods, a sample of trees is returned, allowing a more natural incorporation of uncertainty in phylogenetic estimation.

Communication between palaeontological and neontological workers has been limited. No matter the methodology for estimating them, phylogenetic trees have become crucial to comparative biology. Phylogenetic trees have been used to formulate, give evidence in favour of or against hypotheses about the evolutionary history of groups and traits. Because the field of phylogenetics is both crucial and rapidly changing, we sought to incorporate a diverse cross-section of the approaches biologists use to incorporate fossils in trees, and how those trees are eventually used to inform the study of evolution.

The bulk of these papers have focused on the estimation of phylogenetic trees. Building a quality phylogenetic dataset remains a challenge for researchers; Guillerme & Cooper [5] address this issue in their study on availability of both fossil and extant data in mammals. This paper takes a new look at the very
old question of if and how missing data are injurious to phylogenetic inference, concluding that the distribution of missing data is more important than the ultimate amount. O’Reilly & Donoghue [6] take a broad look at different paradigms for dating phylogenetic trees, using fossils either as tips in the analysis or as node calibrations, ultimately concluding that these approaches can be used concurrently. Matzke & Wright [7] evaluate the importance of the assumptions made in Bayesian divergence dating analyses, and demonstrate how violations of these assumptions can result in poor inferences. Together, these papers address some of the aspects of building and dating phylogenetic trees incorporating fossils that may be opaque to researchers.

Some papers built trees as a way to understand the biology of a specific clade, or to address a question about how traits have evolved, and how methodological advances can allow us to explore these questions more thoroughly. Researchers in this issue tackled two phylogenetic questions that hold not just the interest of scientists, but the general public. Lloyd et al. [8] use the fossil record of dinosaurs in conjunction with model-based methods for a posteriori time-scaling to evaluate when the bird clade, and flight, originated. Lee [9], seeking to understand the timing of mammal evolution, implemented automated approaches for discovering which parts of the dataset evolve under potentially different evolutionary rates. He concludes that automated approaches can be useful for understanding the biology that underlies phylogeny.

Lastly, some of the papers in this issue explored the downstream use of phylogenetic trees in trait-based or biogeographic analysis. Halliday & Goswami [10] and Bapst et al. [11] both evaluate multiple phylogenetic trees in phylogenetic comparative analyses. Halliday & Goswami built both parsimony and Bayesian phylogenies of placental mammals to look at body-size evolution. Bapst et al. perform a similar analysis in birds. Both papers ultimately conclude that careful attention to the assumptions underlying the production of dated phylogenies is crucial, and that the inferential method chosen can greatly affect the results of comparative analyses. Gorscak & O’Connor [12] build a Bayesian time-scaled phylogeny to understand the palaeobiogeography of titanosaur sauropod dinosaurs, in the process posing several large taxonomic rearrangements. They then use a model selection approach to conclude that titanosauras were globally distributed, but experienced regionalization as the continents drifted apart in the Cretaceous. Together, these three papers suggest that to learn more about evolutionary patterns via phylogenies with fossil tips, we must closely examine how we reconstruct our dated trees.

It has been a pleasure to assemble this collection. I am proud that seven of the papers in this special feature come from early-career researchers, including graduate students. I am particularly thankful to *Biology Letters*, and editors Surayya Johar and Kamberd Shergill for assisting and allowing us to gather a diverse and talented group of researchers to make this special feature. This field has moved faster than anyone could have expected. In a few short years, the models and methods available to analyse combined datasets have expanded greatly. My co-editors and I look forward to seeing how methods for putting fossils in trees will continue to develop into the future.

**Competing interests.** I declare I have no competing interests.

**Funding.** This article was funded by National Science Foundation (grant no. 1612858).

### References