Introduction

Modelling the past: new generation approaches to understanding biological patterns in the fossil record

The history of life on this planet is gleaned from analysing how fossils are distributed through time and space. While these patterns are now rather securely known, at least for well-studied parts of the world, their interpretation remains far from simple. Fossils preserve only partial data from which to reconstruct their biology and the geological record is incomplete and biased, so that taxonomic ranges and palaeocommunity structure are imperfectly known. To better understand the often highly complex deep-time processes that gave rise to the empirical fossil record, palaeontologists have turned to modelling the past. Here, we summarize a series of 11 papers that showcase where modelling the past is being applied to advance our understanding across a wide spectrum of current palaeontological endeavours.

Keywords: palaeontology; modelling; evolutionary history; sampling biases

1. INTRODUCTION

The geological record provides a wealth of data about when and where organisms once lived and, through functional analysis of their fossilized remains, about their modes of life [1]. The associations of fossil organisms inform us about the structure of ancient communities, while the pattern of appearances and disappearances of taxa in the fossil record provides the key data from which evolutionary dynamics are deduced [2]. When integrated into a phylogenetic framework, these data can generate a remarkable picture of how life has evolved through time. Yet, interpreting the fossil record correctly is not simple, as it has been shaped by complex processes involving many variables that are often incompletely known. Nor is the fossil record entirely reliable, as there are distortions and biases in the geological record that need accounting for [3]. The great work of the nineteenth and twentieth centuries was to document the patterns that exist in the fossil record, and these now seem robustly established [4]. Furthermore, over the last 40 years, huge strides have been made in developing detailed phylogenies and rigorously analysing and interpreting the emergent patterns. But palaeontologists are now realizing just how complex a task faces them in interpreting process from pattern. In systems where multiple variables interact to create the observed pattern of change over geological time, modelling provides a powerful way to explore the potential role played by individual contributing factors. Whether exploring the biological advantages behind shape evolution, the history of biodiversity, fossil calibration of molecular clock analysis or the dynamics of species–species interactions, modelling provides testable and repeatable numerical results that sharpen and clarify ideas about evolutionary processes. The 11 papers assembled here provide clear examples of how modelling is being applied to provide greater insight into the processes that may have gone into generating the patterns seen in the fossil record.

2. MODELLING FORM AND FUNCTION

Fossils provide hard evidence about the static morphology of their skeletal elements but interpreting the biological function of these structures, and thereby determining what might be driving evolutionary change over time, is fraught with problems. In their papers, Hutchinson [5] and Anderson et al. [6] highlight the problems that palaeontologists face when interpreting the evolution of form and function in fossils. An individual structure often serves a number of different roles, and a suite of interacting biomechanical functions and selective factors need to be considered when assessing what might be driving change. Furthermore, many of the important parameters that a biologist can measure directly are missing in fossils. Nevertheless, employing a modelling approach to assess the performance of existing and hypothetical morphologies using quantitative methods is helping make functional analysis of extinct life forms more rigorous and precise, especially when tied into ever better phylogenetic trees. Major challenges remain, as both Hutchinson and Anderson et al. acknowledge. All models make assumptions, and palaeontological models, in particular, require additional assumptions to study unobservable events in deep time, so the treatment of uncertainty and the assessment of a model’s sensitivity in the face of uncertainty are critical. Nevertheless, with due care and realistic goals, modelling can provide better constraints on functional interpretations even if it cannot always provide definitive answers. The message of both papers is clear: focus on tractable questions and define realistic, justifiable limits with clear realization of where errors might occur in biomechanical models.

3. MODELLING DIVERSITY

For over 150 years, palaeontologists have been endeavouring to document how diversity has changed over geological time. The pattern we see in the fossil record is a mixed signal, the result of biological processes that cause new species to arise and old species to go extinct, overprinted by a sampling bias shaped by geological drivers that have created a non-uniform fossil record [3]. Modelling provides an obvious way to begin to disentangle the effect of these two factors on the empirical fossil record. Two papers focus on this theme, both looking at the dinosaur record. Lloyd [7] uses a proxy for sampling and a modelling approach to isolate the signal in the fossil record that cannot be simply assigned to variation in sampling or...
the nature of the geological record through time. He shows that a long-term decline in dinosaur diversity during the Cretaceous is probably valid. Benson & Mannion [8] tackle the same problem but in a different way, using multivariate regression. They compare specific models of how diversity might have changed over time to test which fits the observed pattern best. In both cases, it is clear that accurate parametrization of sampled diversity and sampling opportunity are the first essential steps in developing better models of diversity over geological time.

Olsweski [9] is faced with a similar problem of having to disentangle the roles of complex interacting factors, but at a smaller scale. He is interested in the composition of ancient communities but is faced with the problem that the assemblages of organisms preserved in one bed are not necessarily a true representation of the living community from which they are derived. This is because death assemblages are time averaged, and taxa differ in their preservation potential. If palaeontologists are to interpret communities in the geological past correctly, they need to be able to understand the processes that go on during the fossilization process. Modelling offers a powerful approach to isolate and investigate the potential effects of single processes, and Olsweski develops a model in which influx and loss through decay are specified for individual species. By comparing his model predictions against empirical data, a better understanding emerges of the factors that drive variation in species abundance ratios in death assemblages.

4. MODELLING EVOLUTIONARY PROCESSES
Recovering information about evolutionary processes from the temporal distribution of fossils has been the bread and butter of palaeobiologists for decades, and the modelling of rates of origination and extinction has long been applied. Five papers fall into this category. Foote [10] recognizes that the analysis of taxonomic structure (i.e. number of species nested within a genus or higher taxon) is an untapped source of data that are potentially informative about underlying evolutionary processes. He uses modelling to show that the ratio of monospecific genera present in a taxon is determined by the interplay of origination and extinction rates. He also shows that the species to genus ratios in extant faunas is also as predicted from the evolutionary dynamics deduced from their fossil record. Although there are other confounding factors that need to be taken into account, such as the effect of taxonomic interest [11], this corroboration from two approaches is welcome and reassuring.

One confounding factor that might affect our understanding of evolutionary dynamics is the role systematists play in defining species and higher taxa. Ezard et al. [12] make an important contribution to this debate by showing how different species concepts affect the outcome of birth–death modelling of evolutionary dynamics. Birth–death modelling is widely used in palaeontology to try to understand evolutionary dynamics through time, but little consideration has been given previously to the ways in which species durations are defined and evolutionary trees constructed. Ezard et al. use one the best fossil records we have, that of planktonic foraminifera, to show just how complex interpreting species dynamics can be when the way we define species can vary so much.

Wagner [13] tackles the thorny question of how morphological change within a phylogenetic framework can best be modelled. This is important for both correct phylogenetic reconstruction and for correctly reconstructing character state changes over time. By using inverse modelling, Wagner is able to show that, of the three models compared, a log-normal distribution of character rate change provides the best fit to empirical data. Intriguingly, he concludes that the preponderance of log-normal rate distributions suggests that morphological evolution is more than a collection of Poisson processes. Palaeontological data are here being used to infer that simple stochastic processes affect when change happens, but some other set of rules (such as integration) affect what change can occur.

The evolution of community structure is the subject of Roopnarine & Angielczyk’s contribution [14]. For them, the evolution of communities over geological time is the result of highly complex interactions. Species–species interactions are paramount, and the success of one species often negatively affects other species in the same community. Thus, patterns of taxonomic rise and fall at higher ecological levels are both driven by, and drive, similar patterns at lower evolutionary levels. Their modelling of trophic networks among Permo–Triassic vertebrate communities suggests that the success of individual species is constrained by the overall success of the communities in which they are living.

Finally, Bush & Novack-Goshall [15] advocate adopting a modelling approach to investigate how ecological diversification of marine Metazoa proceeded through geological time. Instead of treating characters individually, as has been done before, their approach is to model within a multi-dimensional framework. Ecologically important features that can be deciphered from fossils (e.g. mode of feeding and mobility) provide the starting point, and the distribution of these characters in the fossil record then defines ecospace occupation through time. From this, various models of ecospace diversification can be tested against the empirical data to discover which best explains the observed pattern.

5. MODELLING FOSSIL CALIBRATION EFFECTS ON MOLECULAR CLOCKS
A final area, where modelling has an important and growing role to play, is the integration of fossil data into molecular clock studies. Molecular data allow divergence times of living organisms to be estimated, but require the rates of molecular change to be calibrated. This calibration is best done by reference to the fossil record. However, the fossil record is imperfect and this imperfection needs to be taken into account. Warnock et al. [16] present a careful analysis of the different models that are used to define uncertainty in the dating of fossil calibration points. Their modelling shows that the way in which fossil uncertainty is treated can make a huge difference to the
results. This shows that the accuracy and the precision of molecular clock divergence time estimates will only improve if divergence time priors are developed that better summarize the information in the fossil record. This highlights the need for a greater appreciation of the density distribution of samples in the fossil record and a better appreciation of sampling biases through time.

6. CONCLUSION

All of these papers take palaeontological data as their starting point and use models to disentangle the potential effects of individual parameters in situations where there are many confounding factors at play. This is undoubtedly transforming our view of how we approach palaeontology, and ultimately will lead to a better and clearer view of the history of life.

We thank all of the contributors to this volume for their timely and exciting papers. We also extend our deep appreciation to Fiona Pring and Charlotte Wray for their hard work and attention to detail in putting together this Special Feature, and Brian Charlesworth for his invitation to compile it.

Andrew B. Smith* and Paul M. Barrett
Department of Palaeontology, The Natural History Museum, London SW7 5BD, UK
*Author for correspondence (abs@nhm.ac.uk).