The meaning of birth and death (in macroevolutionary birth–death models)

Thomas H. G. Ezard¹,²,*, Paul N. Pearson³, Tracy Aze³ and Andy Purvis¹

¹Division of Biology, Silwood Park Campus, Imperial College London, Ascot, Berkshire SL5 7PY, UK
²Department of Mathematics, University of Surrey, Guildford, Surrey GU2 7XH, UK
³School of Earth and Ocean Sciences, Cardiff University, Cardiff CF10 3YE, UK
*Author for correspondence (t.ezard@surrey.ac.uk).

Birth–death models are central to much macroevolutionary theory. The fundamental parameters of these models concern durations. Different species concepts realize different species durations because they represent different ideas of what birth (speciation) and death (extinction) mean. Here, we use Cenozoic macroperforate planktonic foraminifera as a case study to ask: what are the dynamical consequences of changing the definition of birth and death? We show strong evidence for biotic constraints on diversification using evolutionary species, but less with morphospecies. Discussing reasons for this discrepancy, we emphasize that clarity of species concept leads to clarity of meaning when interpreting macroevolutionary birth–death models.

Keywords: birth; death; extinction; speciation; species concept

1. INTRODUCTION

Birth–death models originated in population biology to address the demise of eminent Victorian families [1]. The models’ substantial use in macroevolution applies the same principles to speciation and extinction [2,3], and they are increasingly being used to investigate the importance of particular traits for diversification [4], the constraints imposed by diversity-dependence [5] and many other questions. Regardless of the particulars of each model, the basic idea remains consistent: the time between events on a phylogenetic tree provides data from which macroevolutionary inferences are made.

We need to know what we mean by births and deaths in this context, since flipping among species concepts alters their definition. For example, what is meant by the birth of a species? The answer may seem obvious at first, yet is anything but. All species trace their ancestry to the origin of life. Hence the birth of a species must be recognized either by the acquisition of some species-defining characteristic or by evolutionary divergence among populations (cladogenesis). Similarly, although extinction would appear to be an obvious endpoint for a species, what if a species evolves into something else or splits into two?

Aze et al. [6] recently proposed dated, fossil-based phylogenies of 339 morphospecies and 210 evolutionary species for Cenozoic macroperforate planktonic foraminifera, whose diversity peaked at 62 and 43 species, respectively. Here, we use these phylogenies and our continuous-time diversification model [7] to show how four palaeontological species concepts can generate very different macroevolutionary dynamics. We begin by outlining the four concepts.

2. THE CONCEPTS

In classical Linnean taxonomy, the members of a ‘species’ are recognized because they share diagnostic characters with a designated type specimen; practical application in the fossil record often results in the delimitation of morphospecies [8]. Consider, for example, a well-known fossil species such as Homo erectus. Fossils are assigned to this species—to the exclusion of all others—if they share characters with the type specimen. As fossil discoveries accumulate, a stratigraphic range for the species is built up. Thus, a temporal duration can be achieved without reliable knowledge of speciation and extinction. A taxonomist may split (by naming new forms) or lump (by synonymy), but the resulting units will still be morphospecies.

In a perfectly complete fossil record (figure 1a), a morphospecies originates with the first specimen to display its diagnostic characters and terminates with the death of the last such specimen. If populations evolve continuously, these diagnostic characters may be polymorphic as morphospecies intergrade for extended periods (shaded rectangles in figure 1b). The appearance or disappearance of morphospecies through such intergradation is termed pseudospeciation and pseudoxinction, to emphasize that there is no change in the number of evolving lineages [8,9] and that extinction can occur through anageneisys. Morphospecies can persist through a ‘budding’ speciation and therefore be paraphyletic. Figure 1b illustrates how this fossil record might be split into five morphospecies: \( \beta \) persists through two speciation events; \( \gamma \) arises through pseudospeciation and terminates through pseudoxinction.

What we term Henigian morphospecies (by analogy with [10]) are similar except that they are not permitted to persist through speciation: the ancestor is presumed to bifurcate into two daughter species and immediately go extinct. However, one of the daughter species will—at least initially—not be diagnosably distinct from its ancestor except by stratigraphic position and assumed phylogeny [10]. Figure 1c shows how the five morphospecies yield nine Henigian morphospecies.

Two further concepts collapse intergrading morphospecies into a single species. Evolutionary species [11] originate at cladogenesis and terminate with extinction of the lineage. The timing of cladogenesis is conservatively inferred to have been when contemporaneous populations no longer intergrade (dotted lines in figure 1); it therefore postdates origination of the descendant morphospecies. The recognition of evolutionary species requires that, when speciation occurs, one of the two species be identified as a continuation of the ancestral species on account of greater similarity...
to the ancestral condition. Figure 1d shows how the five morphospecies collapse to three evolutionary species.

**Hennigian evolutionary species** (termed Hennigian species in [10]) also start at cladogenesis but terminate in either extinction or another cladogenetic event; ancestral species do not persist through speciation. Figure 1e shows how the three evolutionary species of figure 1d translate to five Hennigian evolutionary species.

As a result of these definitions, the average durations differ markedly among species concepts (electronic supplementary material, figure S1): evolutionary species 11.89 Myr (95% parametric CI 10.13–13.95); Hennigian evolutionary species 5.39 Myr (4.84–6.00); morphospecies 8.33 Myr (7.43–9.34); Hennigian evolutionary species 5.39 Myr (4.84–6.00); Hennigian morphospecies 3.90 Myr (3.60–4.21). What impact do these different durations have on macroevolutionary dynamics?

### 3. THE MODEL

We used parametric survival analysis with censoring [12] to obtain extinction and speciation functions for all macroperforate planktonic foraminifer species that originated during the Cenozoic [6]. Survival analysis generates ‘hazards’, which here are instantaneous rates of speciation or extinction. We modelled time to extinction and speciation as Weibull random variables, which tests for changes in hazard as species age [12]. Constant extinction (i.e. Van Valen’s law [13]) is a special case where the time between events is a Poisson process. We flipped among species concepts by applying different censoring rules. Censoring is typically used when an individual survives to the end of an experiment: an individual’s ‘undeath’ is valuable information on the risk of death and therefore species’ duration (electronic supplementary material, figure S2).

Species’ duration was regressed against overall diversity (ln-transformed), palaeoclimate (the deep-sea oxygen isotope ratio [14] as a single proxy), depth habitat (open tropical/subtropical ocean mixed layer; open ocean thermocline; and open ocean sub-thermocline) and morphological types (presence/absence of keels, spines and symbionts). The minimum adequate model for each concept was obtained using the corrected Akaike Information Criterion [15] from combinations of main effects and two-way interactions (for more details, see [7]).

Once the speciation and extinction functions had been obtained independently, we united them in the multivariate Euler–Lotka equation [16] to provide a continuous-time estimate of per capita growth $r$ for an ecologically structured clade [7]. The ecological structure results from significant differences in speciation or extinction hazards among morphological types (see the electronic supplementary material). Since a species of one type can produce descendants of another, we used the matrix of observed, type-specific speciation probabilities to relate all potential ancestors [7]. In this ‘multi-type’ scenario, $r$ is the dominant eigenvalue of the system [16].

Analysis was performed in R [17] using the ‘deSolve’ package [18].

### 4. RESULTS

Species with persistent ancestry show increasing extinction hazard with age, unlike Hennigian concepts (figure 2). While the species most likely to speciate are the youngest under all four concepts, the concept-specific hazards show marked variation (figure 2a–d: compare the heights of the speciation curves at, say, 10 Myr). These differences in speciation and extinction hazards generate very different estimates of $r$ (figure 2e): evolutionary species strongly indicate constraints on diversification, but estimated limits to morphospecies diversity were much higher than those observed (the inferred ‘carrying capacities’ were slightly over 100 morphospecies and nearly 4000 Hennigian morphospecies). All parameters and model comparisons are given in the electronic supplementary material; the importance of species’ ecology for macroevolution was strongly evident for all concepts (electronic supplementary material, tables S1–S3).

### 5. DISCUSSION

Changing the meaning of birth and death by changing the species concept has substantial dynamical consequences for macroevolutionary inference (figure 2) because each species concept generates a different distribution of species’ durations (electronic supplementary material, figure S1), provoking alternative analytical interpretations.

The species most likely to speciate are the youngest for all four concepts (figure 2a–d), which is consistent with the ‘early-burst’ pattern of early, rapid diversification followed by diversity-dependent slowdowns [5]. The estimates of nonlinearity in extinction hazards have overlapping confidence intervals (electronic supplementary material, tables S1–S3).
supplementary material, tables S1 and S2) for evolutionary species and morphospecies, suggesting that increased extinction risk with age is not strongly related to anagenesis as might be expected under a Red Queen scenario [13]. Death comes more quickly when species are not permitted to persist through speciation; Hennigian morphospecies are particularly prone to generating macroevolutionary dynamics with high turnover rates or punctuated bursts of diversification. Hennigian concepts complicate interpretation because it is hard to disentangle speciation and extinction from the additional births and deaths: there are clear differences in extinction hazards with and without persistent ancestry (figure 2a versus 2b and figure 2c versus 2d). By treating speciation and extinction independently, we obtain different distributions (e.g. times to extinction), different hazards (e.g. each speciation generates two Hennigian species but one evolutionary species) and therefore different estimates of clade growth with and without persistent ancestry, even though the number of species is the same.

The constraint of diversity-dependence is lower for morphospecies than for evolutionary species because morphospecies’ numbers can fluctuate owing to anagenesis, i.e. pseudoextinction and pseudospeciation. If morphological change is continual, analyses of morphospecies are biased towards finding that extinction hazard increases with species’ age [19], which is a pattern we detect (figure 2b). If we suppose a changing climate provokes—at least sometimes—anagenesis, then the significant interaction between climate and diversity for all morphospecies’ hazards suggests a less-tightly constrained biotic limit to morphospecies diversity than to the numbers of evolutionary species (electronic supplementary material, table S3). We are therefore even less likely to detect a fixed limit on biodiversity using morphospecies than with evolutionary species (evidence for a variable limit in evolutionary species’ diversity is presented in [7]). With evolutionary species, fluctuations owing to anagenesis are not visible, as they occur within the same lineage.

Although higher taxa (e.g. genera) can represent evolutionarily significant units [20], fossil records that are complete enough to support species-level analyses are particularly valuable for studying macroevolutionary dynamics. Our delimitations of the species concepts assume known phylogenetic relationships [6] and that contemporaneous species can be distinguished morphologically (Darling & Wade [21] review cryptic and near-cryptic diversity in this clade). Because our models are deterministic, they are expected to slightly overestimate growth rates of assemblages exposed to a changing environment [22], but can nevertheless recover key features of evolutionary history [7]. Species-level completeness highlights the choice among species concepts because different meanings of birth and death lead to different interpretations of the same evolutionary history.

We thank Albert Phillimore, Lynsey McInnes and three anonymous reviewers for valuable comments that improved our work and NERC for funding (NE/E015956/1 to A.P. and P.N.P.).


