

Using ghost lineages to identify diversification events in the fossil record

Lionel Cavin* and Peter L. Forey

Department of Palaeontology, The Natural History Museum, London SW7 5BD, UK

*Author and address for correspondence: Département de géologie et paléontologie, Muséum d'histoire naturelle, CP 6434, 1211 Genève 6, Switzerland (lionel.cavin@ville-ge.ch).

Observed rises in taxic diversity could reflect bias of the fossil record or a genuine diversification. Here we outline a new method that attempts to differentiate between these two possible explanations. The method is based on the calculations of average ghost lineage duration through successive intervals of time. Biases due to variation in preservational conditions affect taxa independently from their position in the tree of life. A genuine radiation event will affect some parts of the tree of life more than others. During periods of rapid diversification, there will be a high proportion of new taxa showing short ghost lineages and therefore the average ghost lineage duration will drop as diversity rises, allowing us to distinguish such events from preservational bias during which ghost lineage duration remains unchanged. We test the method on Aptian–Maastrichtian (Cretaceous) ray-finned fish diversity. The result shows that a peak of diversity in the Cenomanian is associated with a drop in average ghost lineage duration, indicating that a genuine biological radiation occurred at that time.

Keywords: ghost lineages; diversity; fossil record; Actinopterygii

1. INTRODUCTION

Long-term (millions of years) changes in biodiversity patterns are difficult to study because of the inadequacy of the fossil record. The source of the issue centres on our reliance upon the fossil record to mirror real diversity over time (Gale *et al.* 2000; Peters & Foote 2001; Smith *et al.* 2001). There is a belief that, over geologically long terms, biases are randomly distributed and do not significantly alter the signal (Sepkoski 1993), or that corrections for sampling effects may produce reliable diversity estimates (Alroy *et al.* 2001). In the present paper, we outline a method to test for the possible distortion caused by the variable quality of the fossil record. We suggest this as a general method applicable to many groups of organisms. We exemplify the method using the fossil record of Cretaceous marine ray-finned fishes (Actinopterygii).

Electronic supplementary material is available at <http://dx.doi.org/10.1098/rsbl.2006.0602> or via <http://www.journals.royalsoc.ac.uk>.

Received 4 December 2006
Accepted 16 January 2007

2. MATERIAL AND METHODS

Increases in taxic diversity observed in the fossil record may be related to two main causes: (i) a genuine increase in species diversification and (ii) an artefact due to biases related to the nature of the fossil record, such as quantity of the available fossiliferous rocks and sampling effort (including variation in research effort). The latter cause is called a 'Lagerstätten effect', although this name is normally restricted to biases due to the presence of numerous rich assemblages only. We hypothesize that the main causes can be distinguished by assessing the shape of the phylogenetic tree of the taxa present in each chosen time slice. A Lagerstätten effect corresponds to a better sampling of the biodiversity in one time slice relative to another. It increases the observed diversity of taxa that, on the null hypothesis, are assumed to be evenly distributed through the tree of life (we ignore here palaeoecological variations between time slices to simplify the explanation). On the other hand, a genuine radiation increases the diversity of observed taxa, but only in restricted parts of the tree of life reflecting a real biological phenomenon. In short, because the phylogenetic patterns of the cohorts under consideration in each situation are different, we suggest that the causes (Lagerstätten effect or radiation event) can be distinguished by computing the variation of the average ghost lineage duration through time slices.

The average ghost lineage duration for any one time-interval is equal to the sum of the ghost lineage durations of all observed equally ranked taxa in that time-interval divided by the number of those taxa. The ghost lineage as used here includes the range extensions as well as the ghost lineages *sensu* Norell (1992). The range extension is the amount of stratigraphic range of a taxon under study that must be added to comply with a phylogenetic tree (Smith 1994), and the ghost lineage *sensu* Norell (1992) is an ancestral lineage leading to two or more taxa that must be inserted to account for the shape of the phylogenetic tree. Rules used to define ghost lineages, especially those based on uncertain stratigraphic data and/or phylogenetic uncertainty, and calculation of the average ghost lineage duration is exemplified in figures 1 and 2 of electronic supplementary material.

We hypothesize that a peak of taxic diversity with no associated change in the average ghost lineage duration is due to a 'Lagerstätten event', while a peak of taxic diversity associated with a drop in the average ghost lineage duration is caused by a genuine biological radiation reflecting speciation events more closely spaced in time. Figure 1 shows a theoretical application of the method.

Another way to assess the nature of a peak of diversity is to compile total diversities for different clades of actinopterygians. For a given time slice with similar preservational conditions, differences in the shapes of the curves imply different evolutionary patterns between the clades. This way allows us to assess the average ghost lineage duration method.

We chose to use the genus as the unit of our analysis. This is because (i) genera are usually recognized on unique characters and demonstrably monophyletic (cf. species that are often differentiated on body proportions, stratigraphic or geographical criteria), and (ii) the phylogenetic analyses that we used to exemplify this method were usually carried out at the generic level (Forey *et al.* 2004).

3. RESULTS

Aptian–Maastrichtian actinopterygian diversity—information about the database and analysis of fish diversity is available in electronic supplementary material. The database of the Cretaceous actinopterygian fossil record we used here is available in electronic supplementary material from Cavin *et al.* (2006).

Figure 2a shows the variation over time of the different diversity metrics for marine actinopterygian genera. The number of observed occurrences (standing diversity) rises gradually from 15 to 30 genera in the late Early Cretaceous (Late Aptian–Late Albian) to 40 or 50 genera in the Maastrichtian. This trend is interrupted by a dramatic peak in the Cenomanian where over 100 genera are recorded. The total diversity (sum of the standing diversity, Lazarus genera diversity and ghost lineage diversity) curve shows a similar pattern, except that the decrease from the Cenomanian peak is gentler due to the presence of more Lazarus taxa. Separate curves are given for the Lazarus and ghost lineage ranges through time.

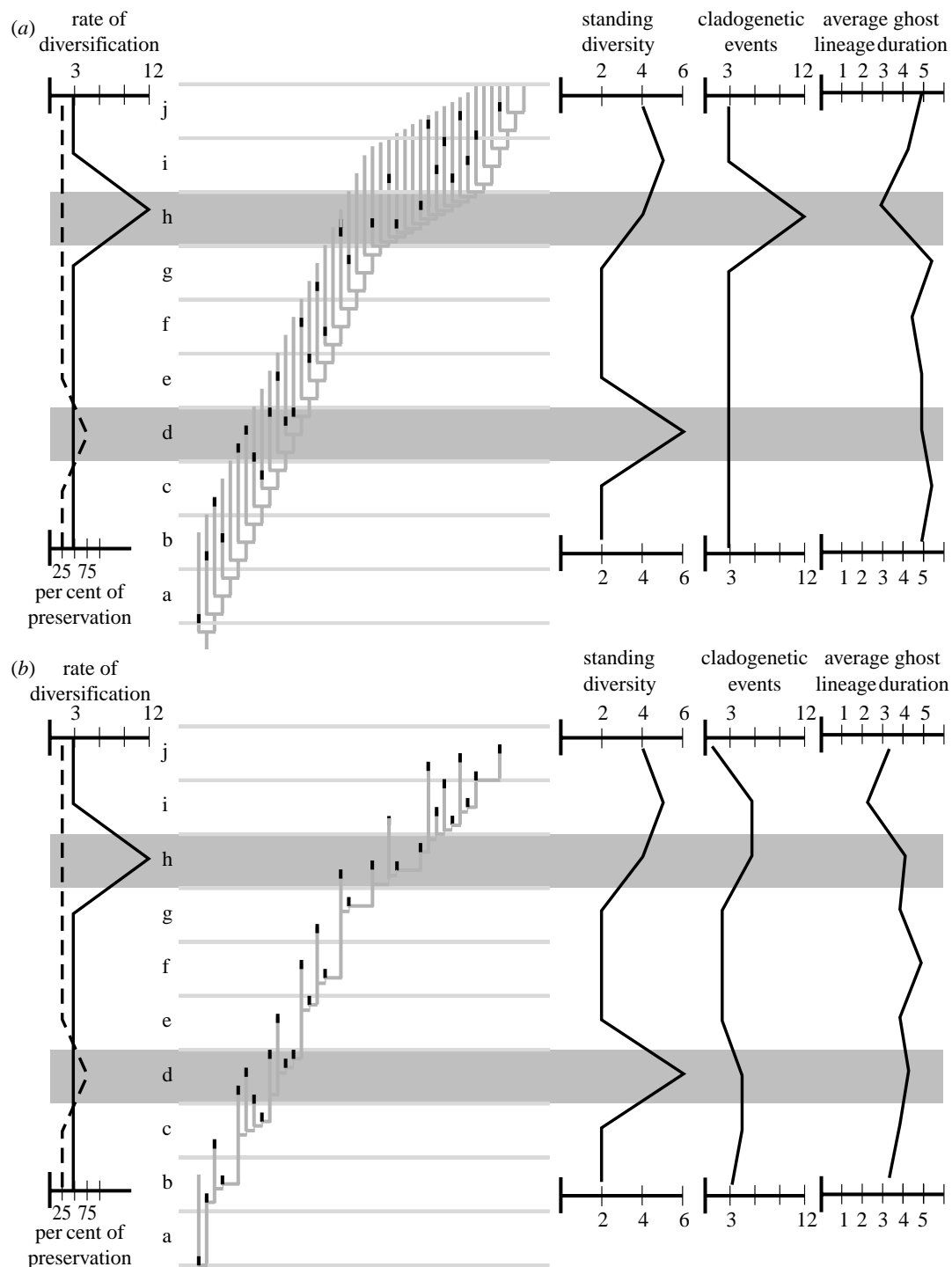


Figure 1. The average ghost lineage duration method. (a) The ‘true’ complete phylogeny is shown as a pectinated cladogram in light grey, plotted against time (10 slices, a–j). The observed occurrences represent the standing diversity and are shown as solid black vertical bars. The per cent preservation remains at 25% (observed occurrences divided by true occurrences) in each time slice except for time slice ‘d’ (75%) that represents a Lagerstätten event (dotted line). The rate of diversification (number of cladogenetic events per time slice) remains constant except in time slice ‘h’ (solid line). The average ghost lineage duration per time slice drops during the period of rapid diversification. (b) Here the phylogeny is constrained to only include known occurrences and average ghost lineage durations per time slice are calculated according to the rules specified in the electronic supplementary material. This would be the normal study situation. Here, the cladogenetic event curve does not discriminate between events in time slices ‘d’ and ‘h’. The average ghost lineage duration is a better indicator of the origin of the peaks, because no significant variation of the average ghost lineage duration is observed in association with the Lagerstätten event, but a drop in the average ghost lineage duration is associated, although slightly postponed, with the radiation event. This model shows that the origin of both kinds of diversity peaks can be separated on the basis of the variations in the average ghost lineage duration.

Assessing how accurately the diversity curve mirrors the genuine fish diversity curve is done here by looking at the variation in the number of cladogenetic events and the average ghost lineage duration through time

(figure 2b). The number of cladogenetic events shows a peak in the Late Albian that, we suggest, led to the peak of standing diversity observed in the Cenomanian. Meanwhile, the average ghost lineage duration drops

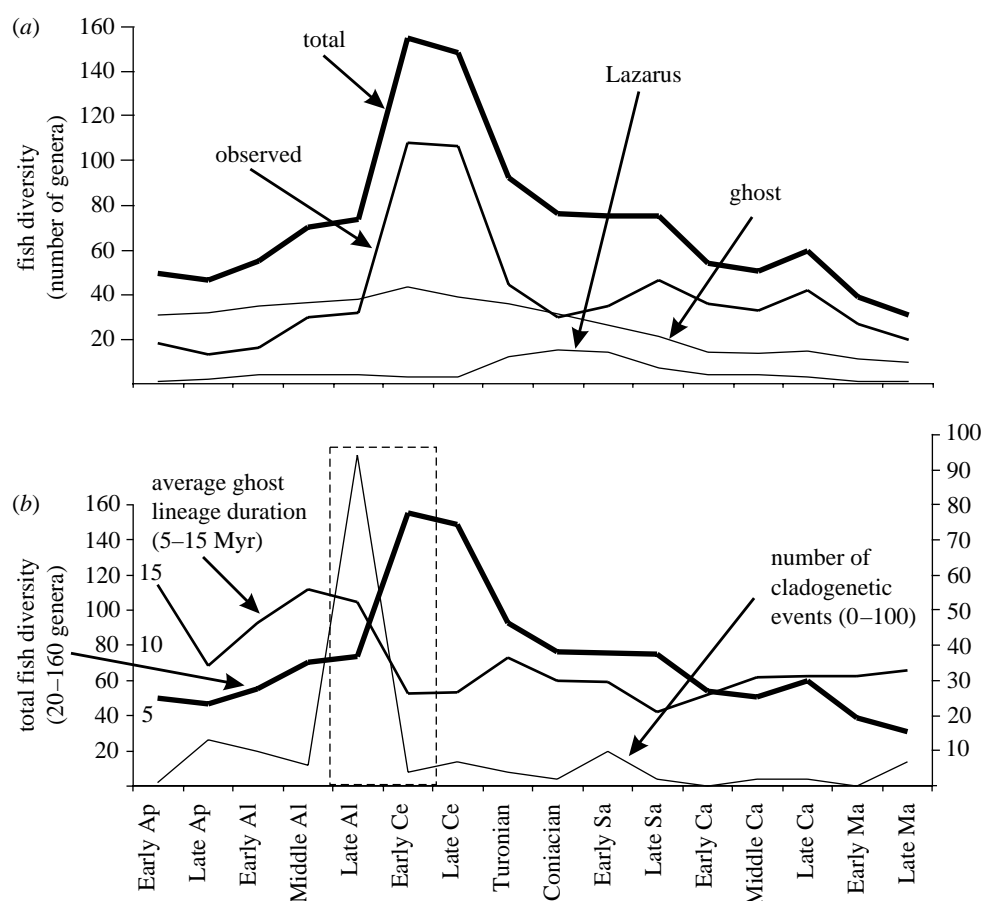


Figure 2. (a) Amount of observed, ghost, Lazarus and total actinopterygian marine genera from the Aptian to the Maastrichtian. (b) Number of cladogenetic events, actinopterygian total diversity and average ghost lineage duration of marine actinopterygians computed with the phylogeny of all actinopterygians from the Early Aptian to the Maastrichtian available in electronic supplementary material from Cavin *et al.* (2006). The dotted frame bounds the Late Albian–Early Cenomanian time interval. Ap, Aptian; Al, Albian; Ce, Cenomanian; Sa, Santonian; Ca, Campanian and Ma, Maastrichtian.

from the Albian to the Cenomanian, confirming that a radiation occurred at that time. The actinopterygian radiation event may well have occurred in the Early Cenomanian rather than in the Late Albian: the peak may be artificially smeared backwards because of the coarse-grained stratigraphic resolution of the actinopterygian database.

Another approach to assess the nature of the variation of diversity curves is to compare the diversity between different clades. We computed the variation of diversity for five clades from the Early Aptian to the Late Maastrichtian (figure 4 in electronic supplementary material). The total diversity of pycnodontiforms and ichthyodectiforms, two clades that appear in the fossil record prior to the Cretaceous, is remarkably constant from the Aptian to the Maastrichtian with no increase in the Cenomanian. Tselfatiiforms, aulopiforms and beryciforms, on the other hand, are clades that diversified in the mid-Cretaceous. Tselfatiiforms and aulopiforms are known by rare fossil occurrences in the Albian and by ghost lineages in the Aptian, while the first fossil occurrence of a beryciform is Cenomanian in age, but its ghost lineage extends to the Early Albian. These three orders show an important radiation starting in the Late Albian through the Cenomanian, after which their respective diversity decreases progressively during the rest of the Late Cretaceous. The marked difference in the diversity

patterns during the mid-Cretaceous event between the ‘old’ clades, which remained constant in number, and the ‘new’ clades, which rapidly diversified at that time, strongly support the hypothesis that the Cenomanian peak of diversity is not due to a Lagerstätten event.

4. DISCUSSION

The main feature of Aptian–Maastrichtian ray-finned fish diversity is the very high peak of diversity observed in the Cenomanian. This stage witnessed one of the highest sea levels of the Phanerozoic leading to the formation of numerous fossil fish deposits (the Cenomanian localities worldwide yield many more fossil fish specimens than any other stage of the Cretaceous). A high sea level certainly accounts in part for the peak of diversity. Marine ray-finned fish diversity as a whole as well as diversity of various more inclusive marine taxa are correlated with upper ocean temperature. The very high temperature recorded for Cenomanian stage (Pucéat *et al.* 2003) probably triggered fish diversification (Cavin *et al.* 2006).

The two lines of evidence mentioned above indicate that the Cenomanian explosion of fish diversity is due to a real biological diversification. Previous studies have shown that some aspects of biodiversity patterns are artefacts of the stratigraphic record (Peters & Foote 2001; Smith 2001). The

phylogeny-based test of the genuine versus artefactual origin of the Cenomanian peak leads us to reject this for this particular group at this particular time.

Funding for this study was provided by a Marie Curie Individual Fellowship funded by the Swiss Federal Office for Education and Science (L.C., grant no. 02.0335).

- Alroy, J. *et al.* 2001 Effects of sampling standardization on estimates of Phanerozoic marine diversification. *Proc. Natl Acad. Sci.* **98**, 6261–6266. (doi:10.1073/pnas.111144698)
- Cavin, L., Forey, P. L. & Lecuyer, C. 2006 Correlation between environment and Late Mesozoic ray-finned fish evolution. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* (doi:10.1016/j.palaeo.2006.08.010)
- Forey, P. L., Fortey, R. A., Kenrick, P. & Smith, A. B. 2004 Taxonomy and fossils: a critical appraisal. *Phil. Trans. R. Soc. B* **359**, 639–653. (doi:10.1098/rstb.2003.1453)
- Gale, A., Smith, A. B., Monks, N. E. A., Young, J. A., Howard, A., Wray, D. S. & Huggett, J. M. 2000 Marine biodiversity through the Late Cenomanian—Early Turonian: palaeoceanographic controls and sequence stratigraphic biases. *J. Geol. Soc. Lond.* **157**, 745–757.
- Norell, M. A. 1992 Taxic origin and temporal diversity: the effect of phylogeny. In *Extinction and phylogeny* (eds M. J. Novacek & Q. D. Wheeler), pp. 88–118. New York, NY: Columbia University Press.
- Peters, S. E. & Foote, M. 2001 Biodiversity in the Phanerozoic: a reinterpretation. *Paleobiology* **27**, 583–601. (doi:10.1666/0094-8373(2001)027)
- Pucéat, E., Lecuyer, C., Sheppard, S. M. F., Dromart, G., Reboulet, S. & Grandjean, P. 2003 Thermal evolution of Cretaceous Tethyan marine waters inferred from oxygen isotope composition of fish tooth enamels. *Paleoceanography* **18**, 1–12. (doi:10.1029/2002PA000823)
- Sepkoski, J. J. 1993 Ten years in the library: new data confirm paleontological patterns. *Paleobiology* **19**, 43–51.
- Smith, A. B. 1994 *Systematics and the fossil record. Documenting evolutionary patterns*. Oxford, UK: Blackwell Scientific.
- Smith, A. B. 2001 Large-scale heterogeneity of fossil record: implications for Phanerozoic biodiversity studies. *Phil. Trans. R. Soc. B* **356**, 351–367. (doi:10.1098/rstb.2000.0768)
- Smith, A. B., Gale, A. S. & Monks, M. E. A. 2001 Sea-level change and rock-record bias in the Cretaceous: a problem for extinction and biodiversity studies. *Paleobiology* **27**, 241–253. (doi:10.1666/0094-8373(2001)027)