Mesozoic marine tetrapod diversity: mass extinctions and temporal heterogeneity in geological megabiases affecting vertebrates

Roger B. J. Benson[1],*, **Richard J. Butler**[2],*, **Johan Lindgren**[3]
and Adam S. Smith[4]

[1]*Department of Earth Sciences, University of Cambridge, Downing Street, Cambridge CB2 3EQ, UK*
[2]*Bayerische Staatssammlung für Paläontologie und Geologie, Richard-Wagner-Straße 10, 80333 Munich, Germany*
[3]*GeoBiosphere Science Centre, Department of Geology, Lund University, 223 62 Lund, Sweden*
[4]*Natural History Division, National Museum of Ireland, Merrion Street, Dublin 2, Ireland*

The fossil record is our only direct means for evaluating shifts in biodiversity through Earth's history. However, analyses of fossil marine invertebrates have demonstrated that geological megabiases profoundly influence fossil preservation and discovery, obscuring true diversity signals. Comparable studies of vertebrate palaeodiversity patterns remain in their infancy. A new species-level dataset of Mesozoic marine tetrapod occurrences was compared with a proxy for temporal variation in the volume and facies diversity of fossiliferous rock (number of marine fossiliferous formations: FMF). A strong correlation between taxic diversity and FMF is present during the Cretaceous. Weak or no correlation of Jurassic data suggests a qualitatively different sampling regime resulting from five apparent peaks in Triassic–Jurassic diversity. These correspond to a small number of European formations that have been the subject of intensive collecting, and represent 'Lagerstätten effects'. Consideration of sampling biases allows re-evaluation of proposed mass extinction events. Marine tetrapod diversity declined during the Carnian or Norian. However, the proposed end-Triassic extinction event cannot be recognized with confidence. Some evidence supports an extinction event near the Jurassic/Cretaceous boundary, but the proposed end-Cenomanian extinction is probably an artefact of poor sampling. Marine tetrapod diversity underwent a long-term decline prior to the Cretaceous–Palaeogene extinction.

Keywords: palaeodiversity; Mesozoic biodiversity; rock record bias; mass extinction; marine reptiles

1. INTRODUCTION

Fossils are our only direct record of the history of animal diversification, providing a window onto processes that have shaped the history of life on Earth, such as mass extinctions, clade-replacement events and the tempo of diversification through time. Large-scale studies of marine invertebrate diversity through the Phanerozoic designed to address these points quantitatively were first carried out during the 1970s and 1980s (e.g. Raup 1972, 1976; Sepkoski *et al.* 1981). Initially, little effort was made to correct for biases introduced to palaeodiversity signals by uneven sampling (but see Raup 1972, 1976) owing to differences in the area of rock preserved at outcrop, in preservation potential of ancient organisms through geological time, and in the collection history of different regions of the globe or geological formations. Recent years have seen increasing investigation into the effects of these factors, and uneven sampling is now recognized as a major confounding influence on our view of

palaeodiversity (e.g. Peters & Foote 2001; Smith 2001, 2007; Crampton *et al.* 2003; Peters 2005; Smith & McGowan 2007; McGowan & Smith 2008; Alroy *et al.* 2008).

Studies of geological biases affecting vertebrate palaeodiversity are less common. Recent investigations of anomodont synapsids (Fröbisch 2008), dinosaurs (Lloyd *et al.* 2008; Barrett *et al.* 2009) and pterosaurs (Butler *et al.* 2009) found correlations between palaeodiversity and geological sampling in terrestrial clades. By contrast, there is apparently no statistically significant correlation between Cenozoic marine mammal diversity and rock area at outcrop (Marx 2008), suggesting that marine vertebrate diversity patterns may not be subject to the sampling biases affecting marine invertebrates and terrestrial vertebrates.

Here, we present the first thorough investigation of the palaeodiversity patterns and geological sampling biases in Mesozoic marine tetrapods, often considered as an adaptive assemblage (e.g. Massare 1987). These taxa represent six taxonomically and morphologically diverse clades that formed an important component of Mesozoic marine ecosystems for *ca* 185 Ma (figure 1). Several major extinction events affecting these clades have been proposed, and examination of these provides insight

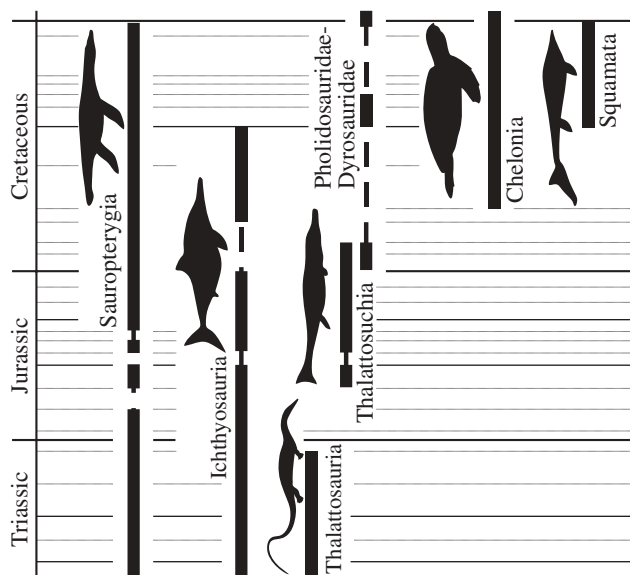* Authors for correspondence (rbb27@cam.ac.uk, butler.richard.j@gmail.com).

Figure 1. Stratigraphic distributions of marine tetrapod occurrences. Dashed lines indicate the inferred presence of clades in geological stages for which they have not been sampled. Squamate silhouette adapted from Lindgren *et al.* (2007), others drawn by A.S.S.

into the timing and severity of mass extinctions more generally. Thus, tetrapods allow hypotheses of marine extinction and diversification to be assessed independently from invertebrate-dominated datasets, and provide comparative data for studies of the effects of sampling biases on other vertebrate groups.

Previous studies of Mesozoic marine tetrapod diversity patterns made inferences based primarily on taxon lists and family ranges (Bardet 1992, 1994) or the phylogenetic composition of Jurassic–Cretaceous marine reptile 'guilds' (Bakker 1993). However, our study is the first to quantitatively analyse palaeodiversity of all Mesozoic marine tetrapods at the species level, taking account of uneven sampling and phylogenetic topology. We make use of methodological advances in the study of rock record biases (Smith & McGowan 2007; McGowan & Smith 2008) and extensive systematic revisions of all major groups, resulting in relatively well-constrained taxonomies and phylogenetic hypotheses (e.g. Hirayama 1998; Rieppel 2000; O'Keefe 2001; McGowan & Motani 2003; Bell & Polcyn 2005; Müller *et al.* 2007; Young *et al.* in press).

2. MATERIAL AND METHODS
(a) *Data collection*
We collected data on marine tetrapod species occurrences in 29 stage-level Mesozoic time bins. This resulted in 572 species occurrences by stage, representing 447 nominal species; one of the largest vertebrate palaeodiversity datasets yet compiled (S1 in the electronic supplementary material). Recent reviews of basal sauropterygians (Rieppel 2000), ichthyosaurs (McGowan & Motani 2003), marine chelonians (Hirayama 1997) and metriorhynchoids (Young *et al.* in press) were complemented by data collection from more recent publications, and the primary literature on thalattosaurs, teleosauroids, squamates and plesiosaurians. Because most taxa are limited to a single stage, we counted taxa

sampled in time bins rather than observing patterns of origination/extinction. These taxic diversity estimates (TDE) do not include phylogenetic ghost lineages.

We used an estimate of the total number of fossiliferous marine formations (FMFs) within each time bin (S1 in the electronic supplementary material) as a proxy for temporal variation in research effort, facies diversity and the volume of fossiliferous rock available for palaeontologists to sample (e.g. Raup 1976; Peters & Foote 2001). These were counted from all records of Mesozoic marine fossils, downloaded from the *The Paleobiology Database* (PBDB) (181 829 records; 12 May 2009). An alternative approach uses subsampling of collections-level occurrence data to standardize sampling between time bins (e.g. Alroy *et al.* 2008). Unfortunately, assembling these data for marine tetrapods would require an immense international databasing project that is beyond the scope of the present study.

(b) *Correlation between rock record estimates and taxon occurrences*
We used Pearson's product–moment correlation (ρ), Spearman's rank correlation (r_s) and Kendall's (τ) coefficient to test for correlation between estimates of rock availability (FMF) and taxic diversity. Significance thresholds were adjusted for multiple comparisons within families following the false discovery rate procedure of Benjamini & Hochberg (1995). Tests of correlation within Triassic, Jurassic and Cretaceous data subsets were treated as a family of three comparisons; the Triassic–Jurassic, and the Cretaceous data were treated as a family of two comparisons (which did not alter the results for the Cretaceous data); the Mesozoic (i.e. total) data did not require correction for multiple comparisons. The Jarque-Bera test confirmed that all data were normally distributed after $\log_{10}$ transformation (S1 in the electronic supplementary material). Short-term autocorrelation was only detected in our Cretaceous data and first difference transformation of these data only slightly weakened statistical correlations (S2 in the electronic supplementary material).

(c) *Correction of raw diversity data for FMF*
Species occurrence data were corrected for our proxy of geological sampling intensity (FMF) using the method of Smith & McGowan (2007; see also Barrett *et al.* 2009; Butler *et al.* 2009). This method calculates a modelled diversity estimate (MDE), which represents the diversity expected if observed diversity biases are solely the result of sampling intensity. The model was constructed by independently rank-ordering taxic diversity and FMF. The ordered data were then paired-off and $\log_{10}$-transformed, applying the function $[f(\text{TDE})=\log_{10}(\text{TDE}+1)]$. A least-squares regression line was calculated for this re-ordered data, representing a relationship in which FMF accurately predicted taxic diversity. The equation of this line was then used to calculate predicted diversity for each time interval based on FMF; this is the MDE. The difference between taxic diversity and modelled diversity is the residual diversity, not explained solely by variation in FMF. High positive or negative residual values are most likely explained by either exceptional sampling events (e.g. Lagerstätten) or genuine changes in palaeodiversity. Because Triassic and Jurassic modelled diversities may be distorted by Lagerstätten effects, a diversity model and residual diversity were also calculated separately for the Cretaceous data (figure 2b).
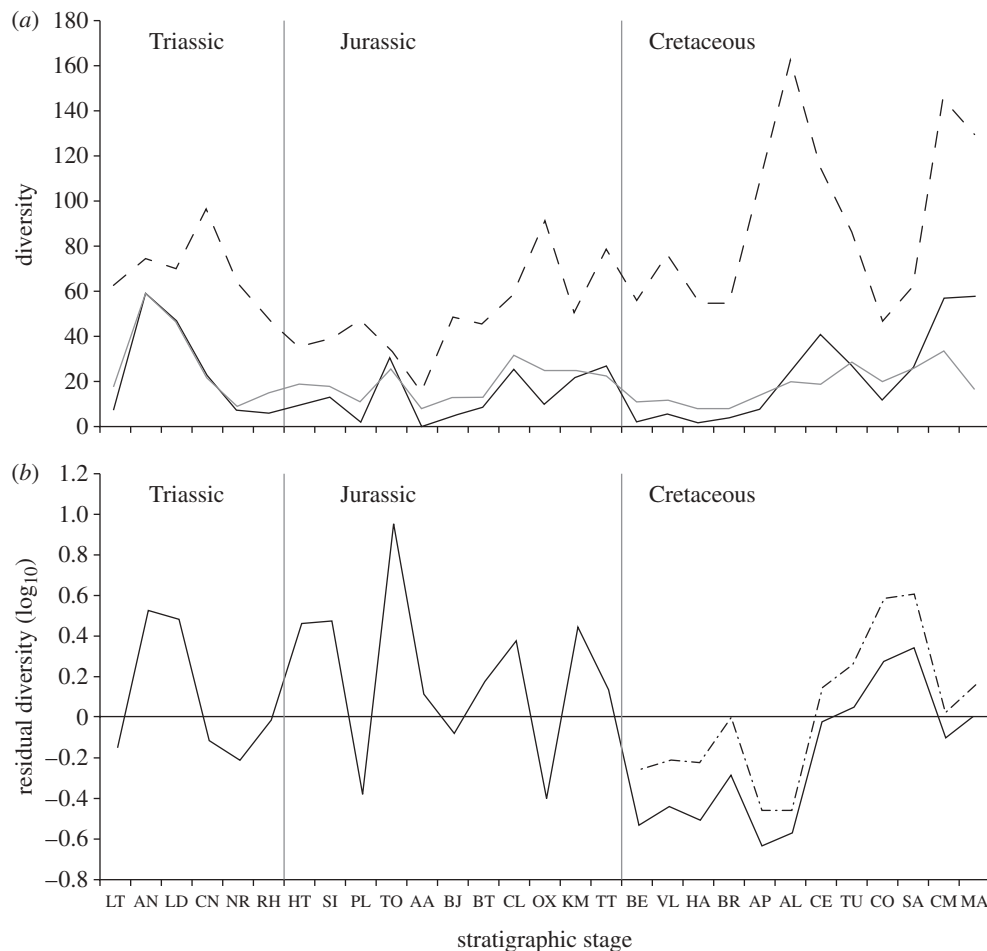
Figure 2. Plots of diversity and FMFs through time. (*a*) Uncorrected (raw) data. (*b*) Residual data after correction for a model in which FMF predicts diversity (see text for an explanation). Dashed line shows FMF, black line shows TDE, grey line shows PDE and dashed-dot line shows TDE corrected for a model based solely on Cretaceous data.

Phylogenetic diversity estimates (PDE) incorporate information from tree structure by adding counts of ghost lineages to counts of taxa included in phylogenetic trees (Norell 1992). They can be used to partially correct for incomplete sampling. However, these should be interpreted cautiously as the correction is still dependent on fossil sampling and only extends backwards in time (Wagner 2000). However, phylogenetic diversity may give a more reliable indication of mass extinction events as lineages that range through geological stages are counted as species occurrences rather than absence. When phylogenetic diversity exceeds taxic diversity it is likely that taxic diversity is an underestimate of relative palaeodiversity, but this can also arise from phylogenetic error (Wagner 2000). If phylogenetic diversity is lower than taxic diversity within a given time bin, this indicates that many taxa of that age have not yet been incorporated into phylogenetic analyses. The cladograms used to compute phylogenetic diversity are included in S1 in the electronic supplementary material.

3. RESULTS
(a) *Comparison between the rock record and taxonomic diversity*
Correspondence between the trend lines of taxic diversity and FMFs is supported by statistically significant correlation (table 1). Both lines follow an approximate pattern (figure 2*a*) in which overall values decline through the Triassic and Early Jurassic, reaching a low level at the

beginning of the Middle Jurassic (Aalenian). They then increase into the late Middle Jurassic (Callovian), and remain approximately constant until the end-Jurassic. The Early Cretaceous is characterized by low taxic diversity, while FMF remains at Late Jurassic levels. Both trends rise to a peak around the Albian–Cenomanian. They then decline until the Coniacian before increasing to a Campanian–Maastrichtian high. Correspondence between taxic diversity and FMF is particularly marked during the Cretaceous (table 1). By contrast, Triassic and Jurassic data are punctuated by deviations, high values of either trend that are not present in the other: Taxic diversity is high in the Anisian–Ladinian, Sinemurian, Toarcian, Callovian and Kimmeridgian (figure 2*a*); FMF is high in the Carnian–Norian, Pliensbachian, Bajocian and Oxfordian. Accordingly, the combined Triassic–Jurassic data correlate less strongly than those for the Cretaceous. This weakness is driven by deviations concentrated in the Jurassic, and correlations of the Jurassic data are weakest (table 1). Although the Triassic data are not significantly correlated, this may result partly from the small number of Triassic time bins ($n = 6$). Indeed, the non-parametric tests (Spearman's r_s and Kendall's τ) recover higher correlation coefficients for the Triassic than for the Cretaceous.

(b) *Corrected diversity estimates*
The trend of residual taxic diversity shows an oscillating pattern in the Triassic and Jurassic, in which

Table 1. Tests of the correlation between $\log_{10}$FMF versus $\log_{10}$TDE over various time intervals. Identical correlation statistics were obtained for the relationship between MDE and $\log_{10}$TDE because MDE is a linear transformation of $\log_{10}$FMF. Statistically significant correlations are indicated by single (*$p < 0.05$) or double (**significant after correction for multiple comparisons: Benjamini & Hochberg (1995); correction not applied to Mesozoic data) asterisks.

interval	Pearson's (ρ)	ρ^2	Spearman's (r_s)	Kendall's (τ)
Mesozoic ($n = 29$)	0.6022, $p = 0.0055$*	0.36	0.4843, $p = 0.0078$*	0.3300, $p = 0.012$*
Triassic–Jurassic ($n = 17$)	0.6146, $p = 0.0087$**	0.38	0.4143, $p = 0.098$	0.2836, $p = 0.11$
Triassic ($n = 6$)	0.6162, $p = 0.19$	0.38	0.7714, $p = 0.072$	0.6000, $p = 0.091$
Jurassic ($n = 11$)	0.5859, $p = 0.058$	0.34	0.2694, $p = 0.42$	0.2037, $p = 0.38$
Cretaceous ($n = 12$)	0.7059, $p = 0.010$**	0.50	0.6895, $p = 0.013$**	0.5231, $p = 0.018$*

Table 2. Tests of the correlation between $\log_{10}$FMF versus $\log_{10}$PDE over various time intervals. Statistically significant correlations ($p < 0.05$), indicated by asterisks (*), were rendered non-significant by the false discovery rate procedure (Benjamini & Hochberg 1995).

interval	Pearson's (ρ)	Spearman's (r_s)	Kendall's (τ)
Mesozoic ($n = 29$)	0.3507, $p = 0.062$	0.3263, $p = 0.084$	0.1930, $p = 0.14$
Triassic–Jurassic ($n = 17$)	0.5153, $p = 0.034$*	0.4332, $p = 0.082$	0.2707, $p = 0.13$
Triassic ($n = 6$)	0.3898, $p = 0.44$	0.6571, $p = 0.16$	0.4667, $p = 0.19$
Jurassic ($n = 11$)	0.6140, $p = 0.044$*	0.4279, $p = 0.19$	0.2617, $p = 0.26$
Cretaceous ($n = 12$)	0.4907, $p = 0.11$	0.4851, $p = 0.11$	0.4341, $p = 0.049$*

Anisian–Ladinian, Hettangian–Sinemurian, Toarcian, Callovian and Kimmeridgian peaks alternate with stages of low or negative residuals (figure 2b). Residuals are negative in the Early Cretaceous and then positive in the Late Cretaceous until the Campanian and Maastrichtian, which have residual values close to zero.

(c) Phylogenetic diversity estimates

The PDE approximately corresponds to taxic diversity in many time bins (figure 2a). However, it is consistently higher in the Norian–Pliensbachian, Aalenian–Oxfordian, Berriasian–Aptian, and Coniacian, suggesting that relative taxic diversity is underestimated for these intervals. This effect is particularly marked in the Oxfordian. Despite this approximate correspondence with taxic diversity (which correlates with FMF), phylogenetic diversity does not correlate with FMF except in the Jurassic and Cretaceous, when marginally significant correlations are recovered from some tests prior to correction for multiple comparisons (table 2). Lack of correlation presumably arises from the inclusion of inferred, non-sampled data in the PDE.

4. DISCUSSION AND CONCLUSIONS

(a) The nature of geological megabias

The significant correlation between taxic diversity and FMF suggests that a portion of the observed pattern of taxic diversity might be explained by temporal variation in the quantity of fossiliferous rock and range of facies sampled by palaeontologists (ρ^2, the proportion of variation in taxic diversity explained by FMF is given in table 1). If this is the case, caution is required when testing macroevolutionary hypotheses using uncorrected taxic diversity. Alternatively, sea level change may drive both FMF and taxic diversity (e.g. Peters 2005), in which case uncorrected taxic diversity may reflect genuine patterns. The relative importance of such effects is still under debate (Benton & Emerson 2007; Smith 2007;

Wall *et al.* 2009) and differentiating between direct causation and 'common cause' is not possible given our data. However, the magnitude of Lagerstätten effects on our data (below) indicates that at least some aspects of sampling heterogeneity cannot be ignored when interpreting palaeodiversity.

A striking feature of our residual diversity trend is the presence of five Triassic–Jurassic peaks (figure 2b). These are not explained by variation in FMF, and could be interpreted as periods of genuinely high diversity. However, these peaks correspond to a small number of European formations from which marine tetrapods have been intensively sampled for *ca* 200 years, and are thus classified as Lagerstätten effects. These formations have yielded a high proportion of species occurrences from their respective stages: the Anisian (55%) and Ladinian (66%) formations of central Europe, the primarily Sinemurian Lower Lias Group of the UK (100%), the Toarcian Posidonienschiefer Lagerstätte of Germany (52%), the Callovian Peterborough Member of the Oxford Clay Formation (73%) and the primarily Kimmeridgian Kimmeridge Clay Formation of the UK (55%). It is clear that the exceptional sampling of these formations distorts observed diversity patterns to the extent that a significant correlation between taxic diversity and FMF is difficult to recover for Triassic–Jurassic data (table 1). The Triassic data include only one Lagerstätte (Anisian–Ladinian) and, despite the small sample size, which may obstruct statistical detection of true correlation, only marginally non-significant correlations arise from non-parametric tests (table 1). This possible correlation in the Triassic suggests that Lagerstätten effects are temporally localized departures from an underlying correspondence between taxic diversity and FMF across the Mesozoic. The absence of such departures from the Cretaceous record is probably because few European Cretaceous formations yielding well-preserved marine tetrapods have been sampled so extensively.

Previous analyses of the relationship between estimates of geological sampling and vertebrate diversity have illustrated that megabiases strongly influence palaeodiversity signals. However, our results demonstrate that qualitatively different sampling regimes may predominate at different points in Earth's history: marine tetrapod palaeodiversity is primarily influenced by variation in FMF in the Cretaceous, by isolated instances of exceptional sampling in the Jurassic and by a possible combination of these factors in the Triassic. Understanding this temporal heterogeneity in the nature of sampling bias is critical to interpreting patterns in vertebrate diversity through time because a single correctional regime may not be sufficient to unpick the patchwork of geological megabias affecting vertebrate preservation. Crucially, it may be difficult to calculate a single line representing 'true' palaeodiversity across all time bins. We explain residual diversity (figure 2b) using a combination of intensely sampled single formations and genuine fluctuations in biological diversity. The former mechanism is most influential in the Triassic–Jurassic.

(b) *Marine reptile diversification and extinctions*

The uncorrected taxic diversity trend line (figure 2a) shows detailed similarity to that used by Bardet (1992, 1994) to suggest marine tetrapod extinction events at the ends of the Ladinian, Tithonian, Cenomanian and Maastrichtian. Bakker (1993) independently proposed end-Tithonian and mid-Cretaceous extinctions. However, residual diversity corrected for variation in FMF shows a different pattern (figure 2b). A high taxonomic and morphological (e.g. Rieppel 2000; McGowan & Motani 2003) diversity of marine tetrapods was already present by the early Middle Triassic (Anisian), less than 6 Ma after the Permian–Triassic boundary. This suggests either a missing Permian record for at least some clades, or a rapid radiation of marine tetrapods after the end-Permian mass extinction. The high diversity of the Anisian–Ladinian is followed by negative residual diversity through the Late Triassic, despite an increase in marine fossil sampling in the early Late Triassic (Carnian). This is corroborated by a decline in phylogenetic diversity, and suggests an extinction event after the Ladinian, encompassing the demise of nothosauroids (Ladinian), pachypleurosaurs and thalattosaurs (Carnian). However, it is difficult to constrain the timing and severity of this event because of pronounced Lagerstätten effects, which distort observed and residual palaeodiversity (figure 2). Subsequent Triassic–Jurassic fluctuations in residual diversity (figure 2b) are impossible to distinguish from the presence or absence of intensive sampling but high phylogenetic diversity suggests that taxic diversity is underestimated during the intervals of apparent low diversity. There is no evidence that an end-Triassic extinction event affected marine reptiles. This conflicts with the controversial evidence for a major mass extinction among marine invertebrates, terrestrial vertebrates and plants (e.g. Tanner *et al.* 2004).

A sharp end-Tithonian (terminal Jurassic) drop in taxic diversity results in pronounced negative residual diversity, from which marine reptiles did not recover until the early Late Cretaceous. This corresponds to a well-documented decline in the diversity of thalattosuchian crocodiles (Young *et al.* in press) and ichthyosaurs (e.g. McGowan & Motani 2003). This result complements hypotheses of a wider extinction event coinciding with a major end-Jurassic reorganization of terrestrial dinosaur faunas (Bakker 1978; Upchurch & Barrett 2005). However, recognition of a marine tetrapod extinction here is complicated by evident differences in sampling regimes between the Jurassic and the Cretaceous. The earliest Late Cretaceous (Cenomanian) is marked by residual diversity close to zero, increasing to high positive values in subsequent stages. This is not consistent with an end-Cenomanian extinction event (Bardet 1992, 1994), instead suggesting progressive diversification driven by the radiation of marine chelonians (despite the final extinction of ichthyosaurs), but obscured by the low numbers of Cenomanian–Santonian marine formations. It is also possible that the observed end-Cenomanian marine invertebrate extinction event is an artefact of geological megabias (Smith *et al.* 2001).

Diversity declined in the latest Cretaceous (Campanian–Maastrichtian). This does not represent an 'edge effect' (e.g. Signor & Lipps 1982) because it results from the in-bin taxon counts and not from range-through data. Barrett *et al.* (2009) recovered a similar result for theropod and ornithischian dinosaurs. In combination, these results suggest that global ecosystems, marine and terrestrial, underwent a progressive decline in vertebrate biodiversity during the terminal stages of the Cretaceous.

Our data provide evidence for significant extinction events in the early Late Triassic, terminal Jurassic and terminal Cretaceous, but not at the end of the Cenomanian. The end-Triassic extinction event, one of the 'big five' mass extinctions, had little effect on marine tetrapods. This is surprising as many taxa were at the apices of marine food chains, and thus might be expected to be highly susceptible to major ecosystem changes. The apparent long-term decline in marine tetrapod diversity prior to the Cretaceous/Palaeogene boundary contradicts prevailing wisdom that the end-Cretaceous extinction event was geologically rapid and cataclysmic (marine tetrapods: Bardet 1992, 1994; Ross 2009). These patterns provide important new insights into the evolutionary history of Mesozoic marine communities.

REFERENCES

Alroy, J. *et al.* 2008 Phanerozoic trends in the global diversity of marine invertebrates. *Science* **321**, 97–100. (doi:10.1126/science.1156963)

Bakker, R. T. 1978 Dinosaur feeding behaviour and the origin of flowering plants. *Nature* **274**, 661–663. (doi:10.1038/274661a0)

Bakker, R. T. 1993 Plesiosaur extinction cycles—events that mark the beginning, middle and end of the Cretaceous. *Geol. Ass. Can. Sp. Pap.* **39**, 641–664.

Bardet, N. 1992 Evolution et extinction des reptiles marins au cours du Mésozoique. *Paleovertebrata* **24**, 177–283.

Bardet, N. 1994 Extinction events among Mesozoic marine reptiles. *Hist. Biol.* **7**, 313–324. (doi:10.1080/10292389 409380462)

Barrett, P. M., McGowan, A. J. & Page, V. 2009 Dinosaur diversity and the rock record. *Proc. R. Soc. B* **276**, 2667–2674. (doi:10.1098/rspb.2009.0352)

Bell Jr, G. L. & Polcyn, M. J. 2005 *Dallasaurus turneri*, a new primitive mosasauroid from the Middle Turonian of Texas and comments on the phylogeny of Mosasauridae (Squamata). *Geol. Mijnbouw NJG* **84**, 177–194.

Benjamini, Y. & Hochberg, Y. 1995 Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B* **57**, 289–300.

Benton, M. J. & Emerson, B. J. 2007 How did life become so diverse? The dynamics of diversification according to the fossil record and molecular phylogenetics. *Palaeontology* **50**, 23–40. (doi:10.1111/j.1475-4983.2006.00612.x)

Butler, R. J., Barrett, P. M., Nowbath, S. & Upchurch, P. 2009 Estimating the effects of sampling biases on pterosaur diversity patterns: implications for hypotheses of bird/pterosaur competitive replacement. *Palaeobiology* **35**, 432–446. (doi:10.1666/0094-8373-35.3.432)

Crampton, J. S., Beu, A. G., Cooper, R. A., Jones, C. M., Marshall, B. & Maxwell, P. A. 2003 Estimating the rock volume bias in paleobiodiversity studies. *Science* **301**, 358–360. (doi:10.1126/science.1085075)

Fröbisch, J. 2008 Global taxonomic diversity of anomodonts (Tetrapoda, Therapsida) and the terrestrial rock record across the Permian-Triassic boundary. *PLoS ONE* **3**, 1–14. (doi:10.1371/journal.pone.0003733)

Hirayama, R. 1997 Distribution and diversity of Cretaceous chelonioids. In *Ancient marine reptiles* (eds J. M. Callaway & E. L. Nicholls), pp. 225–241. San Diego, CA: Academic Press.

Hirayama, R. 1998 Oldest known sea turtle. *Nature* **392**, 705–708. (doi:10.1038/33669)

Lindgren, J., Jagt, J. W. M. & Caldwell, M. W. 2007 A fishy mosasaur: the axial skeleton of *Plotosaurus* (Reptilia, Squamata) reassessed. *Lethaia* **40**, 153–160. (doi:10.1111/j.1502-3931.2007.00009.x)

Lloyd, G. T., Davis, K. E., Pisani, D., Tarver, J. E., Ruta, M., Sakamoto, M., Hone, D. W. E., Jennings, R. & Benton, M. J. 2008 Dinosaurs and the Cretaceous Terrestrial Revolution. *Proc. R. Soc. B* **275**, 2483–2490. (doi:10.1098/rspb.2008.0715)

Marx, F. G. 2008 Marine mammals through time: when less is more in studying palaeodiversity. *Proc. R. Soc. B* **276**, 887–892. (doi:10.1098/rspb.2008.1473)

Massare, J. A. 1987 Tooth morphology and prey preference of Mesozoic marine reptiles. *J. Vertebr. Paleontol.* **7**, 121–137.

McGowan, C. & Motani, R. 2003 *Handbuch Paläoherp.* **8**, 1–175.

McGowan, A. J. & Smith, A. B. 2008 Are global Phanerozoic marine diversity curves truly global? A study of the relationship between regional rock records and global Phanerozoic marine diversity. *Paleobiology* **34**, 80–103. (doi:10.1666/07019.1)

Müller, J., Renesto, S. & Evans, S. E. 2007 The marine diapsid reptile *Endennasaurus* from the Upper Triassic of Italy. *Palaeontology* **48**, 15–30. (doi:10.1111/j.1475-4983.2004.00434.x)

Norell, M. A. 1992 The effect of phylogeny on temporal diversity and evolutionary tempo. In *Extinction and phylogeny* (eds M. J. Novacek & Q. D. Wheeler), pp. 89–118. New York, NY: Columbia University Press.

O'Keefe, F. R. 2001 A cladistic analysis and taxonomic revision of the Plesiosauria (Reptilia: Sauropterygia). *Acta Zool. Fenn.* **213**, 1–63.

Peters, S. E. 2005 Geologic constraints on the macroevolutionary history of marine animals. *Proc. Natl Acad. Sci. USA* **102**, 12 326–12 331. (doi:10.1073/pnas.0502616102)

Peters, S. E. & Foote, M. 2001 Biodiversity in the Phanerozoic: a reinterpretation. *Paleobiology* **27**, 583–601. (doi:10.1666/0094-8373(2001)027<0583:BITPAR>2.0.CO;2)

Raup, D. M. 1972 Taxonomic diversity during the Phanerozoic. *Science* **177**, 1065–1071. (doi:10.1126/science.177.4054.1065)

Raup, D. M. 1976 Species diversity in the Phanerozoic: an interpretation. *Paleobiology* **2**, 289–297.

Rieppel, O. 2000 Sauropterygia. *Handbuch Paläoherp.* **12A**, 1–134.

Ross, M. R. 2009 Charting the Late Cretaceous seas: mosasaur richness and morphological diversification. *J. Vertebr. Paleontol.* **29**, 409–416. (doi:10.1671/039.029.0212)

Sepkoski, J. J., Bambach, R. K., Raup, D. M. & Valentine, J. W. 1981 Phanerozoic marine diversity and the fossil record. *Nature* **293**, 435–437. (doi:10.1038/293435a0)

Signor III, P. W. & Lipps, J. H. 1982 Sampling bias, gradual extinction patterns and catastrophes in the fossil record. *Geol. Soc. Amer. Spec. Pap.* **190**, 291–296.

Smith, A. B. 2001 Large-scale heterogeneity of the fossil record: implications for Phanerozoic biodiversity studies. *Phil. Trans. R. Soc. Lond. B* **356**, 351–367. (doi:10.1098/rstb.2000.0768)

Smith, A. B. 2007 Marine diversity through the Phanerozoic: problems and prospects. *J. Geol. Soc. Lond.* **164**, 1–15. (doi:10.1144/0016/76492006-184)

Smith, A. B. & McGowan, A. J. 2007 The shape of the Phanerozoic marine palaeodiversity curve: how much can be predicted from the sedimentary rock record of western Europe? *Palaeontology* **50**, 765–774. (doi:10.1111/j.1475-4983.2007.00693.x)

Smith, A. B., Gale, A. S. & Monks, N. 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<0241:SLCARR>2.0.CO;2)

Tanner, L. H., Lucas, S. G. & Chapman, M. G. 2004 Assessing the record and causes of Late Triassic extinctions. *Earth Sci. Rev.* **65**, 103–139. (doi:10.1016/S0012-8252(03)00082-5)

Upchurch, P. & Barrett, P. M. 2005 Sauropodomorph diversity through time. In *The sauropods: evolution and paleobiology* (eds K. Curry Rogers & J. Wilson), pp. 104–124. Berkeley, CA: University of California Press.

Wagner, P. J. 2000 The quality of the fossil record and the accuracy of phylogenetic inferences about sampling and diversity. *Syst. Biol.* **49**, 65–86. (doi:10.1080/10635150050207393)

Wall, P. D., Ivany, L. C. & Wilkinson, B. H. 2009 Revisiting Raup: exploring the influence of outcrop area on diversity in light of modern sample-standardisation techniques. *Paleobiology* **35**, 146–167. (doi:10.1666/07069.1)

Young, M. T., Brusatte, S. L. & Ruta, M. In press. The evolution of Metriorhynchoidea (Mesoeucrocodylia, Thalattosuchia): an integrated approach using geometric morphometrics, analysis of disparity and biomechanics. *Zool. J. Linn. Soc.*