

Geological and anthropogenic controls on the sampling of the terrestrial fossil record: a case study from the Dinosauria

P. UPCHURCH^{1*}, P. D. MANNION¹, R. B. J. BENSON²,
R. J. BUTLER³ & M. T. CARRANO⁴

¹*Department of Earth Sciences, University College London, Gower Street,
London, WC1E 6BT, UK*

²*Department of Earth Sciences, University of Cambridge, Downing Street,
Cambridge CB2 3EQ, UK*

³*Bayerische Staatssammlung für Paläontologie und Geologie, Richard-Wagner-Straße 10,
80333 Munich, Germany*

⁴*Department of Paleobiology, National Museum of Natural History, Smithsonian Institution,
P.O. Box 37012, Washington, DC 20013-7012, USA*

*Corresponding author (e-mail: p.upchurch@ucl.ac.uk)

Abstract: Dinosaurs provide excellent opportunities to examine the impact of sampling biases on the palaeodiversity of terrestrial organisms. The stratigraphical and geographical ranges of 847 dinosaurian species are analysed for palaeodiversity patterns and compared to several sampling metrics. The observed diversity of dinosaurs, Theropoda, Sauropodomorpha and Ornithischia, are positively correlated with sampling at global and regional scales. Sampling metrics for the same region correlate with each other, suggesting that different metrics often capture the same signal. Regional sampling metrics perform well as explanations for regional diversity patterns, but correlations with global diversity are weaker. Residual diversity estimates indicate that sauropodomorphs diversified during the Late Triassic, but major increases in the diversity of theropods and ornithischians did not occur until the Early Jurassic. Diversity increased during the Jurassic, but many groups underwent extinction during the Late Jurassic or at the Jurassic/Cretaceous boundary. Although a recovery occurred during the Cretaceous, only sauropodomorphs display a long-term upward trend. The Campanian–Maastrichtian diversity ‘peak’ is largely a sampling artefact. There is little evidence for a gradualistic decrease in diversity prior to the end-Cretaceous mass extinction (except for ornithischians), and when such decreases do occur they are small relative to those experienced earlier in dinosaur evolution.

Supplementary material: The full data set and details of analyses are available at www.geolsoc.org.uk/SUP18487

The same materials (in the form of an Excel workbook) are also available from the first author on request.

Fluctuations in taxic diversity through time form a key aspect of evolutionary history. The reconstruction and analysis of such fluctuations depends, almost exclusively, on data from the fossil record, either directly in the form of counts of fossil species through time, or indirectly when insights into past diversity are obtained from time-calibrated morphological or molecular phylogenies. Palaeontologists universally acknowledge the fact that the fossil record is incomplete, even if debate continues over the extent of this incompleteness and the evenness of sampling over different temporal and spatial scales (e.g. Smith & McGowan 2007; Alroy *et al.* 2008; Butler *et al.* 2009, 2011; Dyke *et al.* 2009; Benson *et al.* 2010; Heads 2010; Goswami &

Upchurch 2010; Sahney *et al.* 2010; Mannion *et al.* 2011; Benton *et al.* 2011). Concerns about the impact of uneven sampling of the fossil record on estimates of diversity arose soon after the first large-scale quantitative studies of diversity were carried out in the 1960s and 1970s (e.g. Raup 1972, 1976). These concerns have persisted to the current time and, during the past decade, have resulted in numerous studies of the impact of uneven sampling on diversity reconstruction, and the development of several methods for measuring and ‘removing’ the effects of sampling (Peters & Foote 2001; Smith 2001; Smith *et al.* 2001; Crampton *et al.* 2003; Peters 2005; Upchurch & Barrett 2005; Smith & McGowan 2007; Uhen & Pyenson 2007; Alroy

et al. 2008; Fröbisch 2008; McGowan & Smith 2008; Barrett *et al.* 2009; Butler *et al.* 2009, 2011; Wall *et al.* 2009; Benson *et al.* 2010; Mannion *et al.* 2011). Today, there are three broad approaches to the latter (Mannion *et al.* 2011): (1) rarefaction or subsampling approaches (e.g. Fastovsky *et al.* 2004; Alroy *et al.* 2008); (2) phylogenetic corrections ('ghost ranges') (Norell & Novacek 1992*a, b*; Lane *et al.* 2005; Upchurch & Barrett 2005; Lloyd *et al.* 2008); and (3) sampling metrics or proxies coupled to 'residual' diversity estimation (Smith & McGowan 2007; Fröbisch 2008; McGowan & Smith 2008; Barrett *et al.* 2009; Butler *et al.* 2009, 2011; Benson *et al.* 2010; Mannion *et al.* 2011). Recently, this situation has become more complex because several authors have proposed that the sampling of the fossil record, and the genuine diversity of taxa, might be correlated because they are both controlled by a third 'common cause' factor such as sea-level (Peters & Foote 2001; Smith *et al.* 2001; Peters 2005, 2006; Benton & Emerson 2007; Lloyd *et al.* 2008; Wall *et al.* 2009). This creates a dilemma for anyone wishing to reconstruct the diversity of a particular group of organisms: the decision to correct for apparent uneven sampling might only be justified when the absence of a common cause can be confirmed (see discussion in Butler *et al.* 2011).

The majority of studies of palaeodiversity patterns and fossil record sampling have been based on marine invertebrates (e.g. Crampton *et al.* 2003; Smith & McGowan 2007; Alroy *et al.* 2008; McGowan & Smith 2008; Wall *et al.* 2009). There have also been several studies of marine vertebrate diversity (Uhen & Pyenson 2007; Marx 2009; Benson *et al.* 2010; Marx & Uhen 2010; Benson & Butler 2011). Although 'sampling-corrected' studies of terrestrial organismal diversity (typically vertebrates) remain scarce, a number of analyses have appeared recently, focusing on Permo-Triassic vertebrates (Benton *et al.* 2004), anomodont synapsids (Fröbisch 2008), dinosaurs (Sereno 1997, 1999; Weishampel & Jianu 2000; Fastovsky *et al.* 2004; Carrano 2005, 2008*a*; Upchurch & Barrett 2005; Wang & Dodson 2006; Lloyd *et al.* 2008; Barrett *et al.* 2009; Butler *et al.* 2011; Mannion *et al.* 2011), pterosaurs (Butler *et al.* 2009) and Cenozoic mammals (Alroy 2000). These analyses have varied in their conclusions with regard to the impact of sampling in the terrestrial realm. Typically, sampling seems to have had significant effects on observed diversity, but it is usually possible to tease apart such effects from genuine diversity change (Wang & Dodson 2006; Lloyd *et al.* 2008; Barrett *et al.* 2009; Mannion *et al.* 2011). However, there are occasions, such as the study of pterosaur diversity by Butler *et al.* (2009), where the apparent peaks and troughs in diversity are

almost entirely controlled by sampling, especially the presence or absence of Lagerstätten (though see Dyke *et al.* 2009 for a contrary view). At present there remains considerable work still to be carried out on a number of issues, including reconstruction of diversity patterns for a wider range of organisms (e.g. terrestrial plants), tests of common cause hypotheses, and evaluations of the relative performances of competing methods for estimating and correcting for sampling biases.

All of these areas are being explored at present, but one aspect of this field has received little attention. Virtually all previous studies have been either 'global' (in the sense that they deal with nearly all available data for a particular clade), or regional (e.g. the study of North American Cenozoic mammals by Alroy (2000) and New Zealand molluscs by Crampton *et al.* (2003)). To date, few studies have looked at the relationships between global and regional diversity and global and regional sampling. This subject represents an important set of issues. One danger with global analyses of diversity is that they can obscure local or regional variations that might be highly significant for our understanding of the patterns and processes relating to radiations and extinctions. Diversity changes could occur simultaneously across the globe (e.g. a sudden mass extinction) or alternatively could display a different pattern in each region (e.g. a group might originate in one area and then subsequently disperse to other regions or along latitudinal gradients). There are also dangers associated with regional studies of diversity. If regional variations in extinction tempo and selectivity have occurred, then extrapolation from a single regional pattern to the global scale will oversimplify the diversity pattern and will make it difficult to obtain an accurate understanding of evolutionary history and the causes of extinction events. Such concerns have been raised with regard to the end-Cretaceous mass extinction because, until recently, much of our understanding of the tempo and selectivity of this event was dominated by data from North America (Alroy *et al.* 2001; Jackson & Johnson 2001; Wang & Dodson 2006). Problems also occur when comparing fossil record sampling at global and regional scales (e.g. see the discussion of competing explanations for correlations between diversity and sampling metrics in Smith & McGowan (2007) and the 'Discussion' here). Thus, without detailed examination of the relationships between global and regional diversity and sampling, our understanding of both genuine diversity change and artefacts created by sampling biases must remain incomplete.

In this paper, we use a recently compiled dinosaur data set to address the following questions

that touch upon global v. regional diversity and sampling patterns in the terrestrial realm:

- (i) At a global scale, is the observed diversity of dinosaurs (and the clades Theropoda, Sauropodomorpha and Ornithischia) correlated with sampling metrics (numbers of dinosaur-bearing formations and collections)?
- (ii) Do different regions display similar or different observed diversity patterns through time?
- (iii) Are similarities or differences in region-to-region diversity patterns created by genuine evolutionary events or do they reflect artefacts generated by sampling?
- (iv) How well do regional sampling metrics predict observed diversity at regional and global scales?

Although dinosaurs form the focus of our analyses, the abundance and ecological dominance of this group during the Mesozoic means that this study has the potential to produce insights into the more general relationships between global and regional diversity and sampling in the terrestrial realm.

Methods and materials

Data set

The Dinosauria provides a particularly suitable group for a case study of palaeodiversity and sampling in the terrestrial realm. This group produced the dominant elements of global terrestrial faunas from the Late Triassic to the end of the Cretaceous, a time span of approximately 160 million years (Weishampel *et al.* 2004; Lloyd *et al.* 2008). Their fossil record has received considerable attention from taxonomists and phylogeneticists, accurate geographical and stratigraphical data are available, and most of the major clades of dinosaurs were globally distributed (Upchurch *et al.* 2002; Weishampel *et al.* 2004; The Paleobiology Database (PaleoDB)). Moreover, previous studies suggest that dinosaur evolutionary history exhibits some important diversity events, including a gradual rise in diversity throughout much of the Mesozoic, and extinctions at the Triassic/Jurassic, Jurassic/Cretaceous and Cretaceous/Palaeogene boundaries (Fastovsky *et al.* 2004; Wang & Dodson 2006; Lloyd *et al.* 2008; Barrett *et al.* 2009; Mannion & Upchurch 2010a; Mannion *et al.* 2011).

Data on the stratigraphical occurrences and geological settings of dinosaurian species were taken from Butler *et al.* (2011), which draws heavily upon the sauropodomorph data set of Mannion *et al.* (2011) and information on theropods and ornithischians in the PaleoDB (data compiled primarily by MT Carrano [Carrano 2008b]). We have also included the sauropod species *Haplocanthosaurus*

delfsi (now recognized by us as probably not congeneric with the type species *H. priscus*), as well as Mesozoic birds (which were omitted by Butler *et al.* 2011) based on data downloaded from the PaleoDB. All species were reviewed in order to eliminate synonyms, *nomina dubia* and *nomina nuda*. The final data set comprises information on 847 valid dinosaurian species, and can be regarded as up-to-date as of March 2010. This represents the largest species-level data set currently available on terrestrial Mesozoic animals. In order to generate data for regional analyses, this data set has been partitioned into five geographical areas based on the continents of Europe, North America, South America, Africa and Asia. Indo-Madagascan, Australasian and Antarctic taxa were excluded from the regional (but not the global analyses) because these areas have yielded very small numbers (4.3%) of dinosaurian species. The number of dinosaur species present in each of the regions is shown in Table 1.

Two global sampling metrics (numbers of dinosaur-bearing formations (DBFs) and collections (DBC)) have been derived from the PaleoDB (downloaded March 2010). These estimated numbers of DBFs and DBCs include all formations and collections where dinosaur remains have been recovered, irrespective of the ability of systematists to assign the remains to a lower-level taxonomic group or particular species. DBC bin counts range from 53 to 1486 and DBF bin counts range from 22 to 172 (see Table 1 and supplementary material). Regional sampling metrics were generated by calculating the numbers of DBCs and DBFs for a given region (i.e. each of the five continental areas specified above – see Table 1). In addition, we have utilized the numbers of terrestrial collections of fossils and geological units for North America presented by Peters & Heim (2010) and the outcrop area for

Table 1. Summary of the number of dinosaurian species, dinosaur-bearing collections (DBC) and dinosaur-bearing formations (DBFs) in each of the five geographical regions used in the regional analyses (and the Indo-Madagascar, Australasia and Antarctica regions ('Other areas') that contributed to the global analyses)

Region	No. species	No. DBCs	No. DBFs
Africa	58	431	61
Asia	308	1125	224
Europe	116	1079	200
North America	224	2449	204
South America	104	397	86
Other areas	37	242	45
Total	847	5723	820

terrestrial sedimentary rock for western Europe estimated by Smith & McGowan (2007). These various global and regional sampling metrics potentially capture information on two key aspects of fossil record sampling. Numbers of terrestrial units, numbers of DBFs and outcrop area potentially measure geological controls on sampling, whereas numbers of terrestrial collections and numbers of DBCs attempt to estimate anthropogenic controls relating to variation in collecting effort by palaeontologists. There are numerous possible alternatives to DBFs in the sense that we could count all vertebrate-bearing formations, or all terrestrial formations irrespective of whether or not they preserve fossils, or we could limit the taxonomic scope (e.g. by comparing sauropodomorph diversity with the number of sauropodomorph-bearing formations). Upchurch & Barrett (2005), and especially Barrett *et al.* (2009), summarized several arguments for using the number of geological formations as a sampling proxy in general, and DBFs in particular:

- (i) Peters (2005) demonstrated that the number of geological formations generally correlates strongly with other measures of fossil record sampling such as numbers of sedimentary rock sections and estimates of total rock volume.
- (ii) Peters & Foote (2001) suggested that the number of formations captures variability in the range of habitats present during each time bin.
- (iii) There may be little benefit in counting geological formations that do not preserve any fossils (or fossils of the groups under investigation): ecological surveys of extant organisms generally do not devote major search effort in environments that are not inhabited by the target organisms (Barrett *et al.* 2009). However, it is important to establish thoroughly that a formation genuinely lacks fossils as opposed to lacking reports of fossils merely because collecting effort has been insufficient to date.
- (iv) Smith & McGowan (2007) have shown that taking into account unfossiliferous formations does not significantly affect the relationship between rock outcrop area and diversity.
- (v) Dinosaurs had a global distribution and their fossils are recovered from all terrestrial facies. Therefore, absence of dinosaurs in a particular formation is more likely to arise from taphonomic factors (dinosaurs were not preserved), or anthropogenic factors (the formation has not been sampled for dinosaurs) than their genuine absence. If the former explanation (taphonomy) is correct, then it is not appropriate to include non-dinosaur-bearing formations within our sampling proxies, because such

formations do not offer a genuine opportunity to sample dinosaurs. However, if it is the case that certain formations have not been inspected for dinosaurs then these might be capable of preserving dinosaur material and ideally should be added into the list of DBFs. Since we cannot be sure which of these explanations is correct (and in reality both are likely to be true), then some arbitrary cut-off point must be established in order to determine which formations are included/excluded in the sampling proxy. Here, we define a DBF as any formation that has produced dinosaur material of any type, even if the specimens are so fragmentary that they can only be identified as 'Dinosauria indet'. In future analyses, it would be interesting to evaluate the relationships between observed dinosaur diversity and a wider sampling proxy such as all terrestrial vertebrate-bearing formations.

Therefore, we follow Upchurch & Barrett (2005), Wang & Dodson (2006), Barrett *et al.* (2009), Butler *et al.* (2011) and Mannion *et al.* (2011) in using DBFs as a proxy for geological controls on observed dinosaur diversity.

Methods

Choice of time bins. Butler *et al.* (2011) based the time bins for their study on Standard European stages and the absolute dates provided by Gradstein *et al.* (2005). These authors examined the effects of variable time bin duration by assessing the statistical correlation between bin length and taxic diversity, and bin length and geological sampling. They also removed the influence of bin duration by calculating partial correlations for pairwise comparisons for parameters such as taxic diversity counts and sampling metrics. The results of this work demonstrated that diversity of dinosaurian species does not correlate with time bin duration (see also Mannion *et al.* 2011).

Fastovsky *et al.* (2004), Barrett *et al.* (2009) and Mannion *et al.* (2011) analysed dinosaur diversity using substage time bins. However, here we do not attempt this level of temporal resolution. The stratigraphical ages of most dinosaurian genera and species can only be determined accurately to stage level at best (and there are several instances where a taxon's age cannot be estimated more precisely than epoch level). In the current data set, for example, only 26% of 212 sauropodomorph species can be dated with reasonable confidence to the substage level (see also Wang & Dodson 2006). Previous analyses that have employed substages as time bins have had to make the assumption that many dinosaurian taxa known from a given stage (e.g.

the Kimmeridgian) occurred throughout the entirety of that stage (i.e. such taxa are assigned to the early, middle and late Kimmeridgian time bins). Given that the majority of taxa cannot be dated with more precision than the stage level, this means that genuine fluctuations in diversity from one substage to the next within a single stage are likely to be overwhelmed by the ‘noise’ generated by the large number of taxa whose stratigraphical ranges have been ‘smeared’ across the entire stage. Moreover, this poor temporal resolution is likely to have the effect of increasing the similarity between scores of adjacent time bins, increasing temporal autocorrelation and potentially leading to artificial inflation of correlation coefficients. Thus, although substage level temporal resolution is important, especially when examining key events such as the lead up to the end-Cretaceous mass extinction (see below), considerable caution should be exercised when interpreting the results. Since the current study is largely based on the data set of Butler *et al.* (2011), we therefore simply use Standard European stages as time bins.

Transformation of the data and statistical comparisons. In order to deal with potentially spurious or inflated correlations caused by trend and temporal autocorrelation in time series data, we applied the method of generalized differencing, which incorporates detrending and differencing but modulates the differences by the strength of the correlation between successive data points (McKinney 1990; Alroy 2000). A full description of this approach is provided by Benson & Butler (2011). All of the data series are detrended; however, only series that show a serial correlation at a time lag of one stage are differenced (with the strength of ‘differencing’ modulated by observed autocorrelation). Comparisons between the detrended data series were made using the non-parametric Spearman’s rank and Kendall’s tau, as implemented in PAST (Hammer *et al.* 2001). Statistical significance was determined using an alpha value of 0.05. However, our analyses form sets or ‘families’ in which the same data (or overlapping portions of the data set) are analysed several times (e.g. the 16 analyses that include comparisons of global DBFs with other data series). For each such family of analyses, the *P*-value used to determine statistical significance has been adjusted for multiple comparisons using the Bonferroni correction (Rice 1989; Waite 2000; Mannion & Upchurch 2010*b*; Butler *et al.* 2011). For example, the global DBF data series is utilized 16 times in our analyses (see supplementary material), so the *P*-value for statistical significance for analyses involving global DBFs is $0.05/16 = 0.0031$. Because each pairwise comparison involves two data series (i.e. members of two families of analyses), the most

stringent *P*-value is used: for example, global dinosaur diversity is analysed 20 times (giving a *P*-value cut-off of 0.0025), so this more stringent *P*-value is applied to the comparison of global dinosaur diversity with global DBFs. However, there are logical inconsistencies inherent to the Bonferroni correction. For example, one researcher carrying out two analyses should apply a *P*-value of 0.025, whereas two researchers each running a single analysis should apply a *P*-value of 0.05. Moreover, we have good reason to suspect a correlation between sampling and diversity and are not simply applying numerous tests of correlation on the off chance that statistically significant relationships are detected. This means that the risk of detecting spurious correlations by making multiple comparisons is not of primary concern. Here, therefore, we discuss the results of our analyses in terms of those that pass or fail when a *P*-value of 0.05 is used, and then comment on the sensitivity of their statistical significance to the application of the Bonferroni correction.

‘Correction’ of taxic diversity and sampling estimates. Differenced taxic diversity counts were ‘corrected’ for both of the global geological sampling metrics (DBC and DBFs) using the residuals method of Smith & McGowan (2007). This approach first reorganizes the data in two data series so that each has its values ranked from low to high. The relationship between the two data series is then expressed as a regression equation, which allows the estimation of the diversity score that would be predicted if observed diversity is entirely controlled by sampling. Residual diversity values are then calculated by subtracting the predicted diversity from the observed diversity (i.e. residuals represent the amount of diversity that cannot be explained by sampling).

Long-term trends. In order to explore putative long-term trends in dinosaur diversity, we have applied a non-parametric runs test using PAST (Hammer *et al.* 2001), which tests for non-random trends in time series data (Hammer & Harper 2006; see also Mannion & Upchurch 2010*a*).

Analyses and results

The results of all analyses (labelled A to Q) are presented in the supplementary material (see also Table 2 for a key to the A–Q analysis labels).

Global diversity patterns

Global dinosaur diversity and sampling. Analyses A1 and A2 compare transformed observed global dinosaur diversity with global DBCs and DBFs

Table 2. Key to analysis labels

Label	Description of analysis
A	Tests of correlation between the observed diversity of dinosaurs, sauropodomorphs, theropods and ornithischians with global DBCs and DBFs
B	Tests of correlation between the observed 'Rest of the world' dinosaur diversity and regional diversity
C	Tests of correlation between the observed dinosaur diversity in two regions
D	Tests of correlation between global sampling and regional sampling
E	Tests of correlation between sampling metrics in two regions
F	Tests of correlation between different sampling metrics for the same region (e.g. DBCs and the numbers of terrestrial units for North America)
G	Tests of correlation between global observed dinosaur diversity and regional sampling metrics
H	Tests of correlation between regional observed dinosaur diversity and a regional sampling metric
I	Tests of correlation between regional observed sauropodomorph diversity and a regional sampling metric
J	Tests of correlation between regional observed theropod diversity and a regional sampling metric
K	Tests of correlation between regional observed ornithischian diversity and a regional sampling metric
L	Tests of correlation between regional observed dinosaur diversity and global sampling metrics
M	Tests of correlation between the DBC-based RDEs of the three dinosaurian subclades
N	Tests of correlation between 'rest of the world' residual diversity estimates and a regional RDE
O	Tests of correlation between the DBC-based RDEs for two regions
P	Tests of correlation between the DBF-based RDEs of the three dinosaurian subclades
Q	Runs tests of long-term trend in the RDEs of Dinosauria, Sauropodomorpha Theropoda and Ornithischia

respectively. There are strong positive correlations between the two sampling metrics and observed global dinosaur diversity (see Fig. 1 for plots of the untransformed data series). Support for these correlations persists even when the Bonferroni correction is applied.

Analyses A3–A8 compare the transformed global observed diversity of sauropodomorphs, theropods and ornithischians with global DBCs and DBFs (see Fig. 2a–c for plots of the untransformed

data series). All analyses demonstrate the presence of strong positive correlations between the observed global diversity of each clade and both sampling metrics, except for the non-significant relationship between sauropodomorph diversity and global DBFs. Application of the Bonferroni correction slightly weakens the support for these positive correlations: all results remain statistically significant except that the comparison of sauropodomorph diversity with global DBCs fails narrowly.

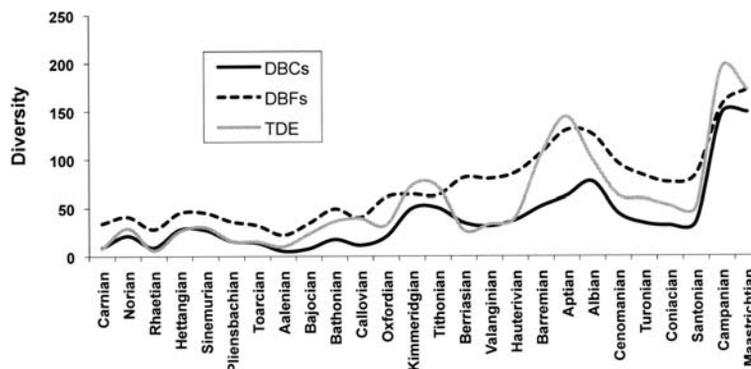


Fig. 1. Graphs showing the fluctuations in observed dinosaur species diversity (TDE), the number of dinosaur-bearing collections (DBC) and the number of dinosaur-bearing formations (DBF). N.B. the number of DBCs has been divided by 10 so it can be plotted on the same y-axis as the number of dinosaur species and the numbers of DBFs.

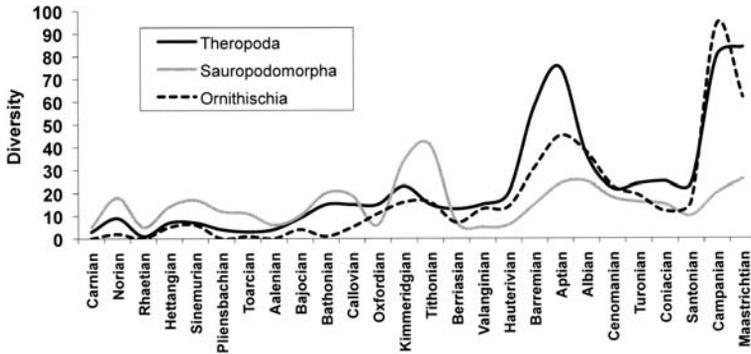


Fig. 2. Graphs of global observed diversity for Ornithischia, Sauropodomorpha and Theropoda.

Global v. regional diversity patterns and regional variations

Global v. regional observed diversity. In Analyses B1–B5, ‘global’ diversity is compared with regional diversity by calculating a ‘rest of the world’ diversity (e.g. the diversity counts for Africa + Asia + Europe + North America) and comparing this with the fifth region (in this case South America). Analyses B1–B4 do not produce statistically significant correlations, although B2 and B3 are relatively narrow fails (P -values are just above 0.05). Analysis B5 (‘rest of the world’ dinosaur diversity compared with African dinosaur diversity) produces statistically significant positive correlations (Fig. 3), but support for this result disappears once the Bonferroni correction is applied.

Region-to-region comparisons of observed diversity. Analyses C1–C10 are a set of pairwise comparisons between the transformed observed diversities of each of the five regions (e.g. European dinosaur diversity v. Asian dinosaur diversity). With

the exception of North American v. South American dinosaur diversity (analysis C2, see Fig. 4 for plots of the untransformed data), all comparisons fail to produce statistically significant results. Moreover, the positive correlation between the dinosaur diversities of North and South America is not supported after the Bonferroni correction is applied.

Comparisons of sampling metrics at global and regional scales

Global sampling v. regional sampling. Analyses D1–D4 compare global sampling metrics (i.e. DBCs and DBFs) with regional sampling metrics (i.e. the North American terrestrial collections and units of Peters & Heim (2010), the western European rock outcrop area of Smith & McGowan (2007), and regional DBC and DBF counts). D5–D14 use ‘rest of the world’ sampling metrics (e.g. the DBCs of Africa + Asia + Europe + North America) to represent global sampling in comparisons with a fifth region (i.e. South America in this example). Six of these comparisons produce

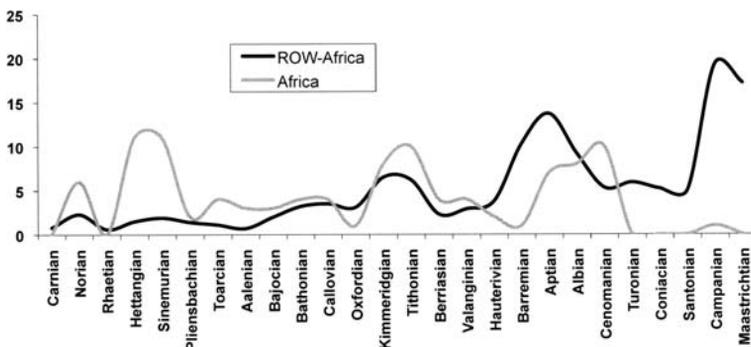


Fig. 3. Graphs of ‘rest of the world’ observed dinosaur diversity (minus Africa) and African dinosaur diversity (see analysis B5). N.B. ‘Rest of the world’ observed diversity has been divided by 10 so that it can be plotted on the same y-axis as African observed diversity.

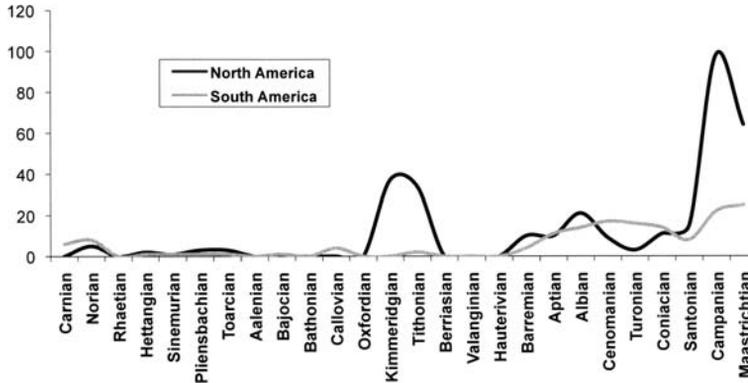


Fig. 4. Graphs of observed dinosaur diversity in North and South America.

statistically significant positive correlations between global (or ‘rest of the world’) and regional sampling metrics. These statistical ‘passes’ include: D1, DBCs v. North American terrestrial collections; D5, rest of the world DBCs v. North American DBCs; D9, rest of the world DBCs v. African DBCs; D11, rest of the world DBFs v. European DBFs; D12, rest of the world DBFs v. South American DBFs (narrow failure according to the Spearman’s rank test, but a pass according to Kendall’s tau); and D14, rest of the world DBFs v. African DBFs. Support for these correlations disappears for D1, D5, D11 and D12 when the Bonferroni correction is applied, but D9 (Fig. 5) and D14 remain statistically significant. Thus, African DBCs and DBFs are apparently positively correlated with ‘rest of the world’ DBCs and DBFs.

Region-to-region comparisons of sampling metrics. Analyses E1–E20 represent pairwise comparisons between sampling metrics from different regions (e.g. North American DBFs v. European

DBFs). Two analyses (E4 and E7) produce statistically significant positive correlations, and E10 and E17 fail narrowly. All four of these analyses are based on comparisons with African DBCs or DBFs. The statistical significance of analyses E4 and E7 disappears when the Bonferroni correction is applied.

Comparisons of different sampling metrics within the same region. Analyses F1–F4 compare different types of sampling metric within the same region (e.g. European DBCs v. western European terrestrial rock outcrop area). All four of these analyses demonstrate the presence of significant positive correlations between the different types of regional metric. For example, analysis F4 demonstrates that European DBFs are positively correlated with the western European rock outcrop area estimates of Smith & McGowan (2007) (Fig. 6a). Support for the significance of F1, F2 and F4 disappears when the Bonferroni correction is applied (F1 and F2 only fail very narrowly), but persists for F3 (North

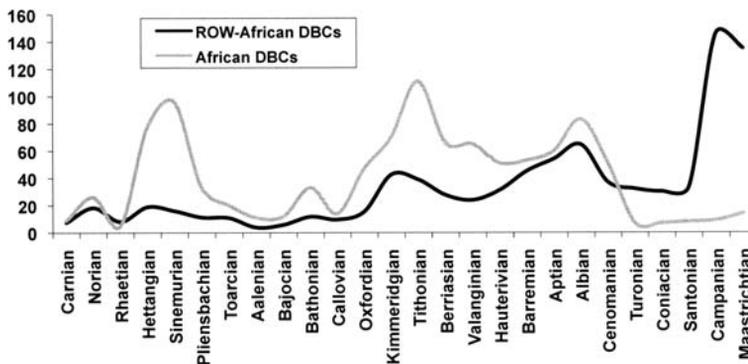


Fig. 5. Graphs of ‘rest of the world’ (minus Africa) DBCs and African DBCs. N.B. ‘Rest of the world’ DBCs have been divided by 10 so that they can be plotted on the same y-axis as African DBCs.

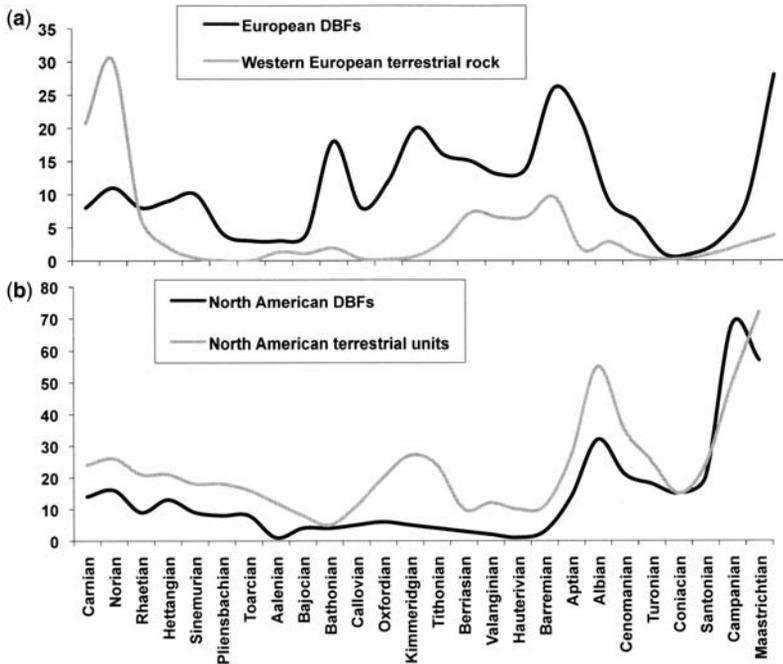


Fig. 6. Graphs of: (a) European DBFs and western European terrestrial rock outcrop area; and (b) North American DBFs and North American terrestrial units. N.B. western European terrestrial rock outcrop area has been divided by 10 so that it can be plotted on the same y-axis as European DBFs.

American DBFs v. the North American terrestrial units of Peters & Heim (2010)) (Fig. 6b).

Regional sampling metrics and observed diversity

Regional sampling v. global diversity. Analyses G1–G13 compare various regional sampling metrics (e.g. North American DBFs) with global dinosaur diversity. Six of these analyses (G1, G4–6, G12, G13) support significant positive correlations, and G2, G3 and G10 fail narrowly. Application of the Bonferroni correction results in the failure of analyses G4–G6, but the statistical significance of analyses G1, G12 and G13 persists. Thus, there is strong support for positive correlations between sampling metrics comprising North American DBCs, African DBCs, or African DBFs and global dinosaur diversity (Fig. 7).

Regional sampling metrics v. regional observed diversity. Analyses H1–H14 compare regional dinosaur diversity (e.g. European dinosaur diversity) with regional sampling metrics taken from the same region (in the case of the current example, these are European DBCs, European DBFs, and western European rock outcrop area). Nine of these

analyses (H1, H2, H4–H6, H11–H14) demonstrate the presence of significant positive correlations (see Fig. 8a, b for exemplar plots of the untransformed data series), and four others (H3, H7, H9 and H10) fail the tests only narrowly. With the exception of analysis H2, the significance of these correlations persists even after the application of the Bonferroni correction. These results therefore suggest that in all regions, except perhaps South America, regional sampling metrics are strongly positively correlated with regional diversity counts.

Regional sampling metrics and regional diversity have also been compared for each of the three major clades of dinosaurs (Fig. 9a–c): sauropodomorphs (analyses I1–I14); theropods (analyses J1–J14); and ornithischians (analyses K1–K14). For sauropodomorphs, eight analyses (I1, I4–I8, I13, I14) demonstrate significant positive correlations between regional sampling and regional diversity, and analyses I2 and I9–I11 fail the tests only narrowly. Support for three of these positive correlations (I1, I5, I6) persists even when the Bonferroni correction is applied, and three of the new failures (I4, I8, I13) fail only narrowly. Nine of the theropod analyses (J1, J2, J4–J6, J10–J14) produce significant positive correlations, and one analysis (J7) fails the statistical tests narrowly.

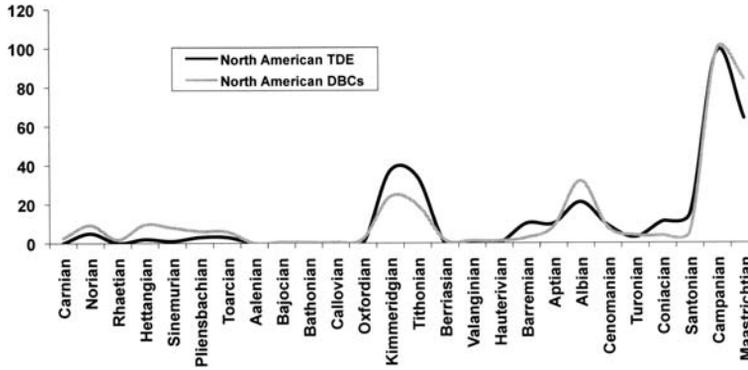


Fig. 7. Graphs of North American DBCs and observed dinosaur diversity. N.B. North American DBCs have been divided by 10 so that they can be plotted on the same y-axis as North American observed dinosaur diversity.

Application of the Bonferroni correction produces seven remaining significant results (J1, J5, J6, J10–J14). For ornithischians, eight analyses yield significant positive correlations (K1–4, K6, K11, K13, K14) and one (K5) fails the tests narrowly. Three of these analyses (K1, K4, K13) continue to pass the statistical tests (and K2 and K14 only fail narrowly) even after the application of the

Bonferroni correction. In general, these results demonstrate the existence of strong positive correlations between sampling and the diversity of saur- opodomorphs, theropods and ornithischians at the regional scale. The strongest correlations typically occur in North America, Europe and Africa, with comparisons involving Asia and South America failing more frequently.

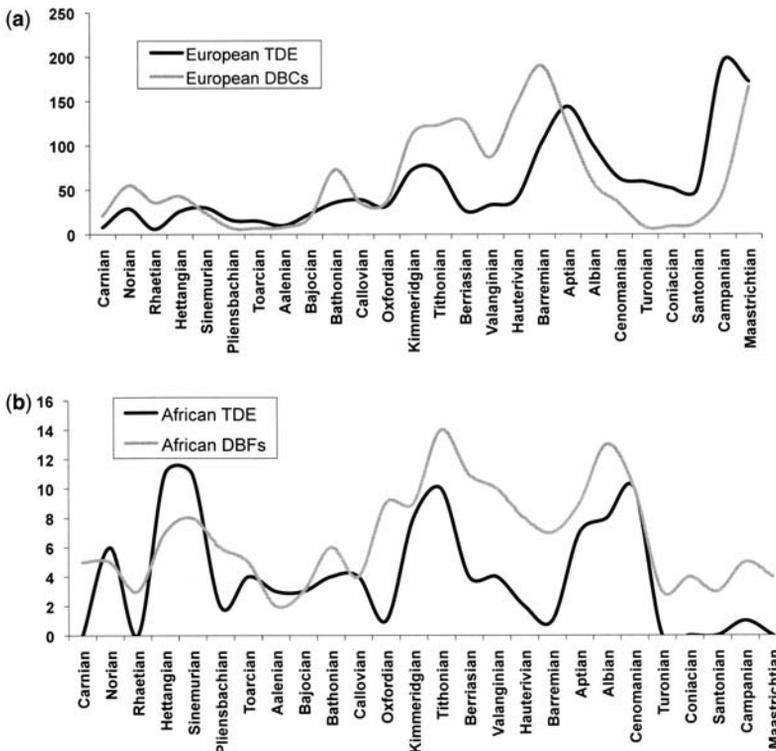


Fig. 8. Graphs of: (a) European observed dinosaur diversity and European DBCs; and (b) African observed dinosaur diversity and African DBFs.

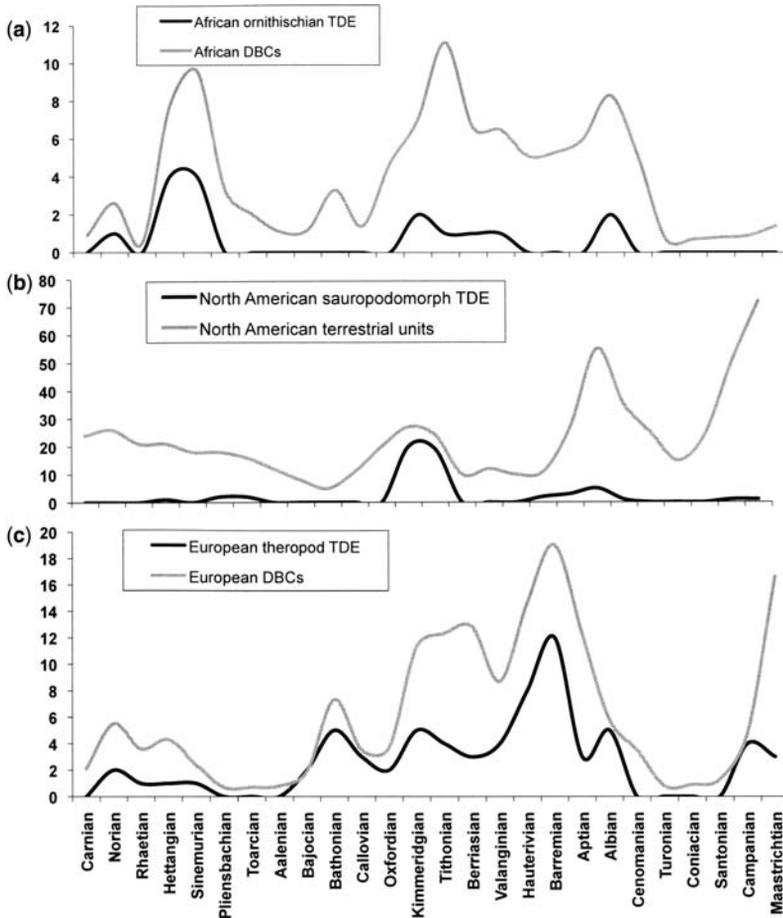


Fig. 9. Graphs of: (a) African ornithischian observed diversity and African DBCs; (b) North American sauropodomorph observed diversity v. North American terrestrial units; and (c) European theropod observed diversity and European DBCs.

Global sampling and regional diversity

Analyses L1–L10 compare global sampling (DBC and DBF) with dinosaur diversity in each of the five regions. Four of these analyses (L1, L4, L8, L9) demonstrate the presence of positive correlations between global sampling and regional diversity (Fig. 10), and two others (L3, L6) fail the tests only narrowly. With the exception of analysis L1, all of these significant results disappear when the Bonferroni correction is applied.

Residual diversity at global and regional scales

Global residuals for dinosaurs and three major clades. Figures 11 and 12a–c show the residual diversity estimates (RDEs) for dinosaurs, theropods,

sauropodomorphs and ornithischians, calculated using DBCs. Analyses M1–M3 make pairwise comparisons between the RDEs of the three subclasses: those of theropods and ornithischians are positively correlated, whereas there are no correlations with the RDEs of sauropodomorphs. The statistical significance of the correlation between theropod and ornithischian RDEs disappears when the Bonferroni correction is applied.

Figure 13a–e show the residual diversity estimates for dinosaur diversity in each of the five geographical regions. Analyses N1–N5 compare the global (rest of the world) dinosaur RDEs with the RDEs for each of these regions. All of these analyses fail the correlation tests; one fails marginally (N5).

Analyses O1–O10 are a series of pairwise comparisons between the RDEs of the five geographical

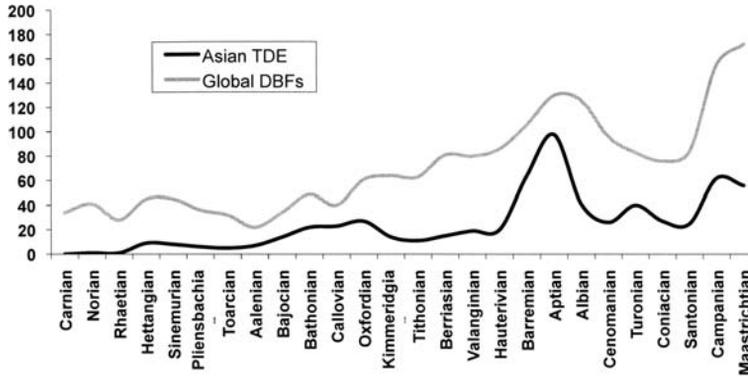


Fig. 10. Graphs of Asian observed dinosaur diversity and global DBFs.

regions (e.g. O1, North American RDE v. European RDE). All of these analyses fail the correlation tests.

In order to enable direct comparisons of our results with those of Barrett *et al.* (2009), we have also calculated the RDEs for theropods, sauropodomorphs and ornithischians using DBFs (see Fig. 14a–c and the supplementary material). Analyses P1–P3 carry out pairwise correlation tests between these three RDEs. The results indicate that theropod and ornithischian RDEs, and theropod and sauropodomorph RDEs, are strongly positively correlated, but a correlation between sauropodomorph and ornithischian RDEs cannot be detected.

Long-term diversity trends

In order to examine the possible presence of long-term trends in dinosaur diversity during the

Mesozoic, we have applied a runs test to the global DBC-based RDEs for dinosaurs (analyses Q1–Q3), DBC-based RDEs of sauropodomorphs (Q4–Q6), theropods (Q7–Q9) and ornithischians (Q10–Q12), and DBF-based RDEs for these three clades (Q13–Q21). For each set of analyses, long-term trends have been investigated for the Mesozoic as a whole, the Late Triassic–Jurassic inclusive, and the Cretaceous. Analysis Q1 suggests that there is a long-term trend of increasing RDE values through the Mesozoic, but Q2 and Q3 indicate that this results from the combination of a weakly supported upward trend during the Triassic and Jurassic, and no discernible trend during the Cretaceous. The DBC-based residuals produce different results for each of the three clades considered. Ornithischians display no detectable long-term trend in RDEs, whereas the increase in RDEs for theropods during

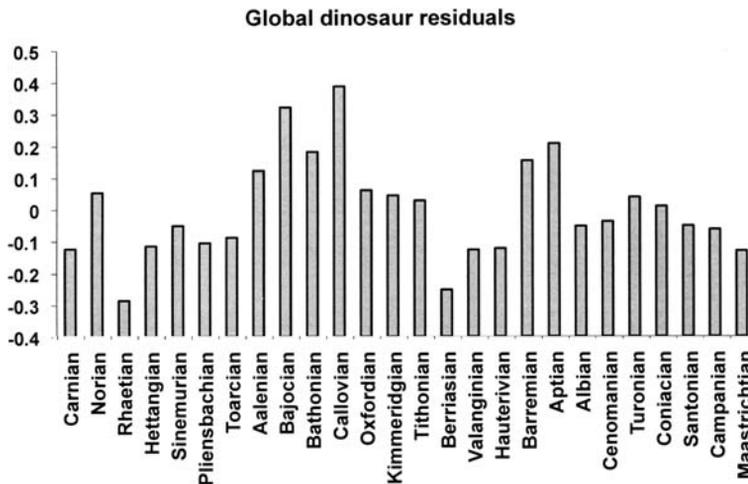


Fig. 11. Residual diversity estimates of global dinosaur diversity based on DBCs.

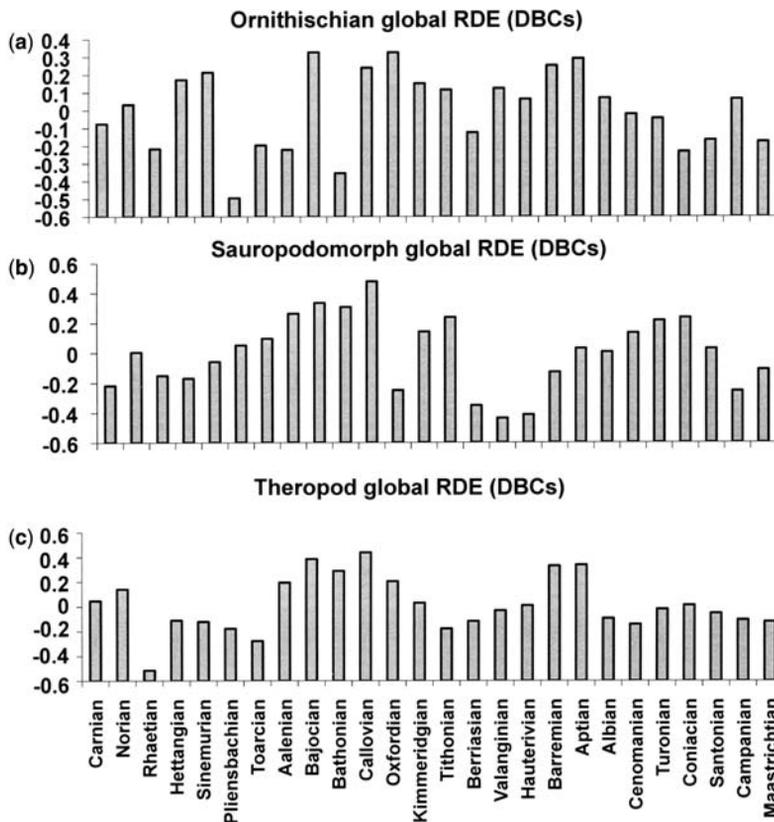


Fig. 12. Residual diversity estimates based on DBCs for: (a) global ornithischian diversity; (b) global sauropodomorph diversity; and (c) global theropod diversity.

the Mesozoic is apparently generated by an upward trend during the Triassic–Jurassic and no trend in the Cretaceous. Sauropodomorphs are interesting because they display evidence for upward trends in RDE values for both the Triassic–Jurassic and Cretaceous, but these combine together to produce only a weakly supported trend for the Mesozoic as a whole. This can be explained as the result of combining two temporally distinct upward trends that are separated by the decrease in sauropod diversity at the Jurassic/Cretaceous boundary (see below). None of the DBF-based RDEs show any evidence for an upward diversity trend during the Mesozoic or in either of the shorter time slices.

Discussion

Diversity and sampling

Several previous studies (e.g. Wang & Dodson 2006; Lloyd *et al.* 2008; Barrett *et al.* 2009; Butler *et al.* 2011; Mannion *et al.* 2011), have

demonstrated that global dinosaur diversity (or that of subclades such as Theropoda) correlates positively with global sampling metrics such as the numbers of DBCs, DBFs and/or dinosaur-bearing localities. This conclusion is supported by the results of analyses A1–A8 here. One exception to this is the diversity of sauropodomorphs *v.* the number of DBFs (analysis A4). Barrett *et al.* (2009) found a weak, non-significant, negative correlation (based on Spearman's rank and Kendall's tau tests) between phylogenetic diversity estimates for sauropodomorphs and the number of DBFs, whereas the same study found a non-significant positive correlation between observed sauropodomorph taxic diversity and DBFs. Mannion *et al.* (2011) and the current study, found no statistical support for either a negative or positive correlation for the Mesozoic as a whole, but Mannion *et al.* (2011) did find a positive correlation for the Cretaceous by itself.

Approximately half of our other analyses that have compared sampling with diversity also

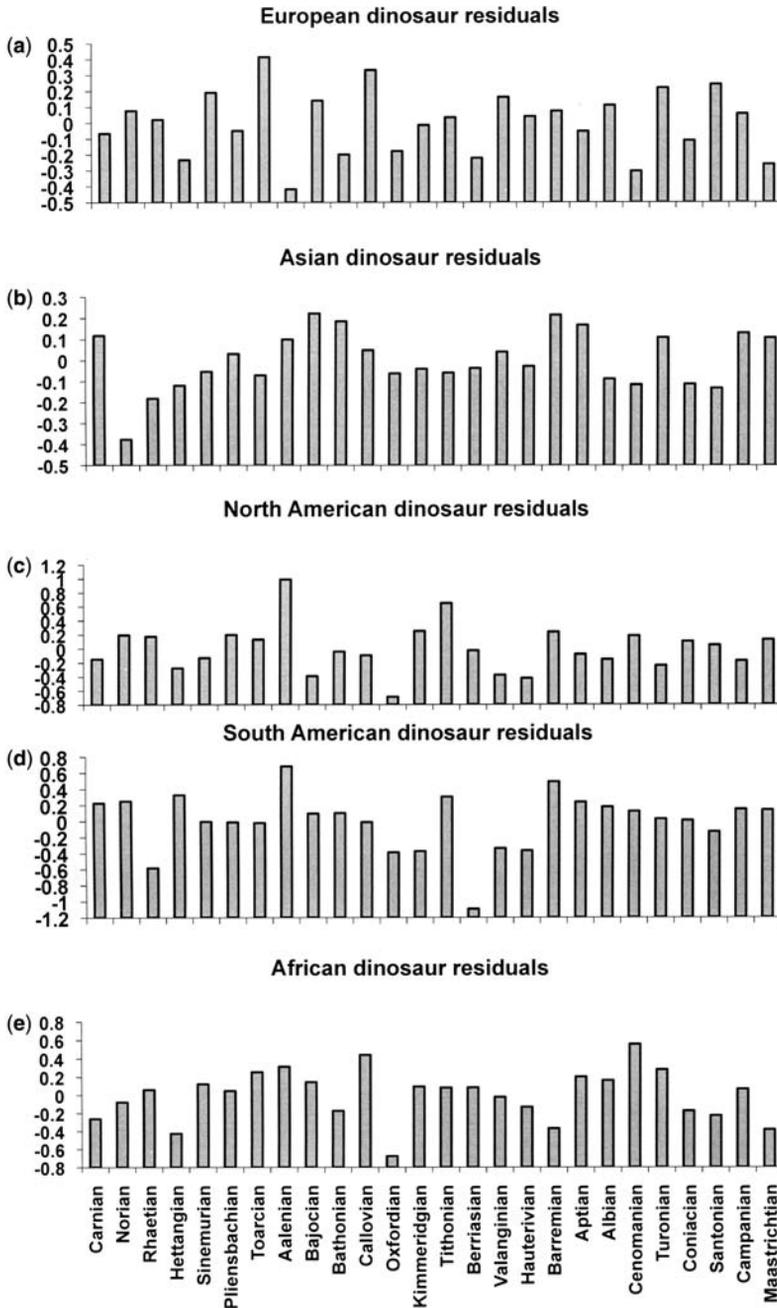


Fig. 13. Regional residual diversity estimates for dinosaurs based on DBCs: (a) Europe; (b) Asia; (c) North America; (d) South America; and (e) Africa.

demonstrate the presence of strong positive correlations between these parameters. These analyses include comparisons of global dinosaur diversity and regional sampling (G analyses – six passes

out of 13 comparisons), regional dinosaur diversity with regional sampling (H analyses – nine passes out of 14 comparisons), regional sauropodomorph, theropod and ornithischian diversity v. regional

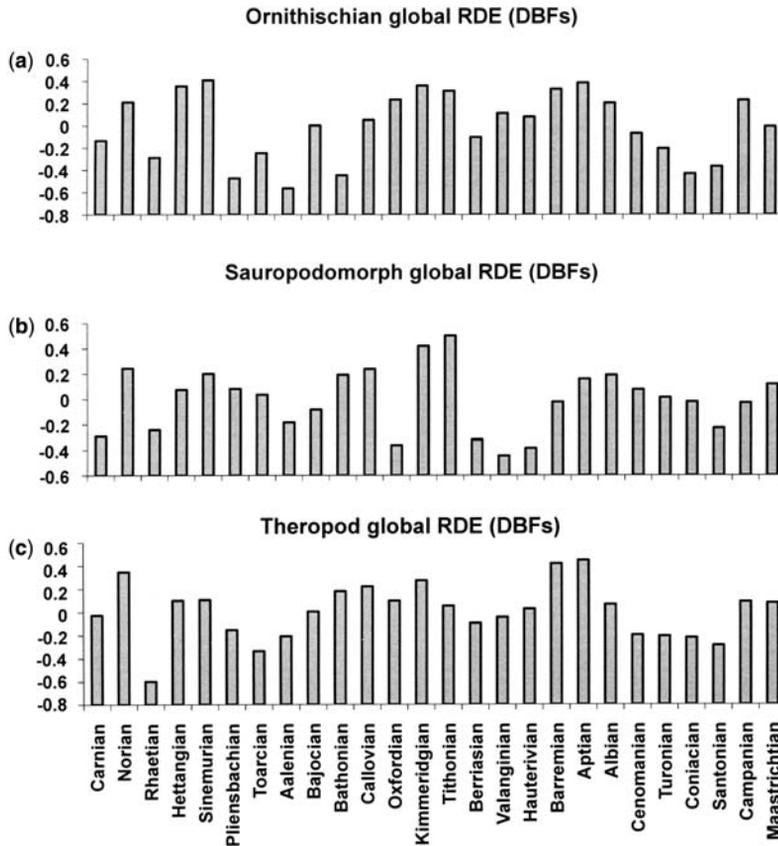


Fig. 14. Residual diversity estimates based on DBFs for: (a) Ornithischia; (b) Sauropodomorpha; and (c) Theropoda.

sampling (I, J and K analyses – 25 passes out of 42 comparisons), and regional dinosaur diversity v. global sampling (L analyses – four passes out of 10 comparisons). There are three principle explanations for these strong positive correlations between the sampling metrics and observed diversity at the various geographical and phylogenetic scales:

- (i) *Common cause.* A third factor, such as sea-level, might control both sampling (e.g. by determining how much fossiliferous sediment is preserved or destroyed during each time bin) and genuine diversity (e.g. via the effects of sea-level on the fragmentation/connection of land areas and/or increases/decreases in land area) (Sepkoski 1976; Peters & Foote 2001; Smith *et al.* 2001; Peters 2005, 2006; Benton & Emerson 2007; Lloyd *et al.* 2008; Wall *et al.* 2009; Butler *et al.* 2011; Mannion *et al.* 2011). Butler *et al.* (2011) presented a series of analyses that demonstrate that

apparent correlations between sea-level or land area and observed dinosaur diversity disappear once data series are detrended and the effects of temporal autocorrelation are taken into account. In contrast, positive correlations between dinosaur diversity and sampling metrics persist even after data transformation has removed the effects of long-term trends. Here, therefore, we do not regard sea-level common cause as a convincing explanation for positive correlations between the sampling metrics and observed dinosaur diversity. At present, we are not aware of any other hypothesized common cause mechanisms that could explain positive correlations between sampling metrics and observed dinosaur diversity.

- (ii) *Circular reasoning or redundancy.* A second possibility is that DBCs and DBFs correlate with dinosaur diversity because of circular reasoning. If dinosaurian taxa have played a significant role in the recognition and definition of geological units (e.g. formations), then the

discovery of new dinosaurs might prompt the naming of new geological formations, and a tendency for DBFs and dinosaur diversity to be positively correlated would be automatically built into these data series. Similarly, it could be argued that those time periods when dinosaurs were genuinely more diverse and abundant will tend to produce outcrops with richer fossil assemblages. Such outcrops might in turn tend to attract disproportionate attention from palaeontological expeditions because of the perceived increased probability of a successful and productive field season. Moreover, a genuine decrease in diversity is likely to place constraints on the number of DBCs: palaeontologists may still search intensively, but must report few DBCs if specimens are rare. If these phenomena occur, then a tendency for observed dinosaur diversity to be positively correlated with the number of DBCs might also be generated automatically. This problem has also been termed 'redundancy' (see Benton *et al.* 2011).

In previous analyses and those carried out here, however, the dangers of circular reasoning have been minimized by counting all DBCs and DBFs irrespective of the quality or quantity of material recovered from each collection or formation. This means that a formation that has only produced a few scraps of indeterminate dinosaur material is still considered to be 'an opportunity to observe' dinosaur diversity (see Upchurch & Barrett 2005), and carries as much weight in the analyses as a geological unit (such as the Late Jurassic Morrison Formation of the USA) that has produced hundreds of skeletons belonging to tens of diagnosable dinosaurian species (Weishampel *et al.* 2004). At present, quantitative data on the extent to which dinosaur discoveries influence the naming of geological units is not available. However, we suspect that the decision to divide sedimentary sequences into separate geological formations is usually based on abiotic evidence (e.g. boundaries between formations are recognized on the basis of changes in facies), and the recognition and naming of geological units often precedes the discovery of dinosaurs in them. When biotic evidence is used to distinguish between formations, this may include the presence/absence of certain dinosaurian taxa, but is frequently based on non-dinosaurs such as ostracods, shelled invertebrates, fish and other tetrapods. For example, the recognition of the Dinosaur Park Formation (Alberta, Canada) and the 'Dinosaur beds' of Malawi, as distinct geological units, seems to result from the presence of

dinosaur fossils. On the other hand, the Cabao Formation in Libya was named prior (approximately in 1963) to the discovery of dinosaur remains in the 1980s' (Le Loeuff *et al.* 2010): this formation therefore seems to have been distinguished from other geological units on the basis of sedimentological differences rather than the presence of certain types of dinosaur. Moreover, any circular relationships between DBCs/DBFs and observed diversity should be at least partially disrupted when comparisons are made between these metrics and clades within Dinosauria: after all, a formation that has been recognized on the basis of the occurrence of a distinctive ornithischian dinosaur is not linked (in terms of circular reasoning at least) to the observed diversity of sauropodomorphs. Yet, analyses A3–A8 demonstrate that the observed diversities of the three main clades of dinosaurs also correlate with numbers of DBCs and DBFs (with the exception of sauropodomorphs v. DBFs – see above), and evidence for these positive correlations also occurs when these clades are compared with regional sampling metrics. These results suggest that there is more to the relationships between these data series than a simple circularity based on counts of geological units or collections defined by the presence of dinosaurs.

Finally, it should be noted that the results of analyses F1–F4 indicate that there are positive correlations between different sampling metrics in North America (i.e. DBFs and DBCs v. the terrestrial rock units and collections of Peters & Heim 2010) and in Europe (i.e. DBFs and DBCs v. the western European outcrop area estimates of Smith & McGowan 2007). Importantly, the Peters & Heim (2010), and especially the Smith & McGowan (2007) sampling metrics are not defined on the basis that they are 'opportunities to observe dinosaur diversity' (e.g. sedimentary rock outcrop area is counted in square km). Thus, aside from an impact on sampling rates, there is no reason to expect a correlation between these sampling metrics and the number of dinosaur taxa present in these sediments. Agreement among sampling metrics, each based on a different criterion, suggests that they are collectively 'homing in' on some form of sampling signal, and correlations cannot be explained merely in terms of their definition based on the presence/absence of dinosaurs. We suggest therefore that while circular reasoning cannot be completely ruled out as a contributory factor to the occurrence of positive correlations between sampling metrics and observed dinosaur

diversity, it is unlikely that this phenomenon entirely (or even largely) accounts for these correlations.

- (iii) *Sampling influences observed diversity.* In agreement with previous analyses, such as Fastovsky *et al.* (2004), Wang & Dodson (2006), Barrett *et al.* (2009), Butler *et al.* (2011) and Mannion *et al.* (2011), current evidence indicates that observed dinosaur diversity is strongly controlled by sampling, suggesting that sampling regimes in the terrestrial realm have an important impact on observed diversity patterns (as is also the case in the marine realm (e.g. Smith & McGowan 2007; Benson *et al.* 2010)). The fact that positive correlations between sampling metrics and observed diversity persist even when data are transformed and detrended, indicates that short-term fluctuations in sampling are probably responsible for short-term changes in observed diversity (Butler *et al.* 2011).

Common cause, circular reasoning and sampling are not mutually exclusive phenomena, and it is possible that all three factors have played some role in shaping observed dinosaur diversity. At present, however, the evidence supports a dominant role for sampling, and indicates that it is legitimate to attempt to reduce or remove the effects of uneven sampling via techniques such as subsampling or residual diversity estimation.

Dinosaur diversity at global and regional scales

Previous studies of dinosaur diversity have typically focused on patterns at the global scale (Carrano 2005; Upchurch & Barrett 2005; Wang & Dodson 2006; Lloyd *et al.* 2008; Barrett *et al.* 2009; Butler *et al.* 2011; Mannion *et al.* 2011) or have examined a single region such as North America or Africa (e.g. Carrano 2008a; Mannion 2009). At present, therefore, little is known about how dinosaur diversity varied from region to region, and there has been virtually no work on how such regional comparisons might be affected by sampling (though see Fastovsky *et al.* 2004; Mannion & Upchurch 2011). Analyses B1–B5 and C1–C10 examine the relationships of observed dinosaur diversity at global and regional scales. The results (one pass out of 15 comparisons) indicate that there is very little support for positive correlations between either global ('rest of the world') observed diversity and that of a given region, or between the diversities of pairs of regions. Thus, although individual peaks or troughs in observed diversity in two or more regions may coincide occasionally (e.g. Late Jurassic diversity peaks in North America and Africa associated with

the rich dinosaur faunas of the Morrison Formation and Tendaguru beds respectively), there is no compelling evidence that fluctuations in observed dinosaur diversity occurred in a uniform way across the globe during the Mesozoic (though see discussion of the Jurassic/Cretaceous boundary extinction below). However, this lack of uniformity or similarity might be an artefact created by the distorting effects of uneven sampling of the fossil record in each region. For example, simultaneous peaks in genuine diversity might have occurred in all geographical regions, but poor sampling in some regions might obscure the truly global nature of this evolutionary radiation. This issue is addressed by analyses N1–N5 and O1–O10, which examine correlations between residual dinosaur diversities at global and regional scales and between pairs of regions. The results demonstrate that there is little evidence for a truly uniform global pattern of evolutionary radiations among dinosaurs, even after the removal of the effects of sampling.

These results have two important implications. First, they suggest that sampling regimes have not imposed either artificially uniform or artificially disparate patterns on diversity across our five geographical regions. Second, it seems that dinosaur evolution (at least in terms of the radiation and extinction of species) followed different patterns in different regions (or at least that genuine differences outweigh similarities sufficiently often to preclude significant passes of correlation tests). This is perhaps not a particularly surprising result. Given regional differences in both biotic and abiotic conditions that inevitably occur as a result of climatic factors and evolutionary history, there is generally no *a priori* expectation that dinosaur diversity should have changed in a similar manner, at similar rates, in each region. Regional differences in dinosaur evolution could have occurred for many reasons. For example, each continental region may have had distinctive climatic, environmental and/or biotic conditions that provoked unique evolutionary responses among their particular dinosaurian faunas. Furthermore, it is conceivable that certain clades originated in one particular region, generating a spike in diversity in that region alone. Whether or not these regional radiations resulted in subsequent changes in diversity elsewhere in the world would then depend on the complex interactions between palaeogeography (e.g. the creation/destruction of dispersal routes) and palaeoecology (e.g. the ability of dispersing taxa to occupy niches in newly invaded areas). Similarly, there may have been region-specific extinctions: for example, diplocoids in North America apparently died out at the Jurassic/Cretaceous boundary, whereas this clade continued to flourish in South America during the Early Cretaceous. However, one obvious exception

to this 'lack of uniformity' might occur when mass extinction events result in a simultaneous global reduction in diversity, as may have occurred at the Triassic/Jurassic, Jurassic/Cretaceous and Cretaceous/Palaeogene boundaries. These extinction events are discussed at both global and regional scales in 'Dinosaur radiations and extinctions revisited' below.

Sampling metrics at regional and global scales

As noted above, some sampling metrics such as numbers of North American terrestrial units/collections (Peters & Heim 2010) and estimates of sedimentary rock outcrop area (Smith & McGowan 2007; McGowan & Smith 2008), circumvent the potential problem of 'circular reasoning' because they estimate our opportunities to find fossils using units of measurement that are not defined by the presence of the taxa in which we are interested. One limitation with these sampling metrics, however, is that their temporal, and particularly their geographical extents are often limited by the quality and quantity of geological data available. For example, the intensive investigation of European stratigraphy during the past two centuries has led to a generally well-understood, detailed and accurate knowledge of the ages and distributions of sediments in this area, which in turn means that estimates of rock outcrop area per time bin can be made relatively easily and accurately. However, it is no coincidence that such regional sampling metrics have been produced for North America and Europe, but rarely for other parts of the world. In Asia, for example, uncertainties about the ages of rock outcrops are far more prevalent (e.g. the Mengyin Formation, that has yielded the sauropod *Euhelopus*, could be either Late Jurassic or Early Cretaceous in age – see Wilson & Upchurch 2009 and references therein). Moreover, the combination of extensive remote areas and a shorter history of geological research in parts of Asia, Africa and South America, means that many sediments remain to be mapped and/or dated accurately. These circumstances raise the question as to whether detailed and accurate sampling metrics based on data for North America and/or Europe are suitable as sampling metrics for studies of diversity at global scales. A similar issue was raised by Smith & McGowan (2007) in their study of global marine invertebrate diversity and western European rock outcrop area. These authors demonstrated the existence of positive correlations between their regional sampling metric and the global diversity of marine invertebrates. Smith & McGowan (2007) proposed two hypotheses to account for this correlation. One possibility is that a regional sampling metric

captures information on global sampling regimes: for example, fluctuations in rock outcrop area in western Europe could reflect the impact of factors that might control sedimentation at a global scale (such as the eustatic component of sea-level). Alternatively, the marine invertebrate data set might actually be dominated by taxa from North America and western Europe because of the longer history of palaeontological research in these regions. If the latter is correct, then the positive correlation between data series might occur because the analysis is actually comparing a regional sampling metric with a regional (or at least regionally biased) taxonomic count.

Inevitably, there is some regional skew in our dinosaur data set, which is dominated by Northern Hemisphere taxa in general and Asian forms in particular (Table 1). Nevertheless, the data set is truly global in terms of its taxonomic coverage and two of the sampling metrics (DBC and DBF). In addition, the dominance of any given region fluctuates through time, so that no one region is uniformly dominant in the sample throughout the Mesozoic. This means that evidence for positive correlations between regional sampling metrics and global diversity cannot be easily explained as a result of a regional taxonomic bias (especially for those analyses that do not use Asian DBCs and DBFs as the regional sampling metric). Thus, if positive correlations occur, they probably indicate that the regional sampling metric has captured at least part of the sampling signal at the global scale. These issues are examined here via the series of analyses that compare global and regional sampling (D analyses), inter-regional sampling (E and F analyses), regional sampling and global observed diversity (G analyses), global sampling and regional observed diversity (L analyses), and regional sampling v. regional observed diversity (H–K analyses).

Analyses D1–D14 and E1–E20 compare global ('rest of the world') and regional sampling metrics, or carry out pairwise comparisons between sampling metrics for individual regions. Most of these analyses (27 out of 34 comparisons) fail the statistical tests. However, there are some striking examples of positive correlations between data series. Interestingly, most such examples involve African DBCs or DBFs, suggesting that sampling regimes in this region are not independent of those in other regions. In general, however, the results of analyses D and E suggest that regional sampling metrics are not particularly effective at capturing global sampling signals. This may be because there is no single uniform regime that has been imposed on the sampling of the dinosaurian fossil record by an external factor such as sea-level. Even though sea-level will rise and fall in a coherent manner across the globe, this may have different

effects on the sampling of the terrestrial fossil record in different regions, perhaps depending on local conditions such as climate, continental size and configuration, continental shelf slope and area, and also biotic factors such as variations in the extent to which the terrestrial faunas tended to occupy coastal, fluviolacustrine and/or arid inland environments. Similarly, the lack of correlations between 'rest of the world' and regional DBCs also indicates that there is no evidence for a uniform global pattern in collecting effort. This could mean that palaeontologists are attempting to find dinosaur fossils throughout the known stratigraphical and geographical range of this group, without marked biases towards particular time periods (e.g. the Late Cretaceous) or regions (e.g. North America). It is also conceivable, however, that palaeontologists do display biases in their collecting effort, but these are constrained and overwhelmed by geological controls on rock availability.

Comparison of regional sampling with global observed diversity (G analyses) and global sampling with regional diversity (L analyses), yield a surprising number of statistical passes (10 out of 23 comparisons, although only four passes remain after the application of the Bonferroni correction). Not surprisingly, however, significant positive correlations occur far more often (34 out of 56 comparisons) when the observed diversity of a region is compared with a sampling metric for that same region (H–K analyses). The general lack of correlation between global sampling and regional observed diversity, and regional sampling and global observed diversity, may stem from the combination of both regional variation in genuine diversity patterns and regional variation in sampling rates (see above). These results have at least two important implications. First, although regional sampling metrics can be used as proxies for global sampling under some conditions, in general it is better to employ global sampling metrics when dealing with global observed diversity and a regional sampling metric when investigating the observed diversity of that region. Second, global diversity patterns and global sampling regimes should be viewed as generalized patterns that represent summaries across regions and therefore tend to obscure important regional variations. There are thus increased risks of error when extrapolating from a regional diversity pattern to the global scale and when using a regional sampling metric as a proxy for global sampling.

Dinosaur radiations and extinctions revisited

The following sections discuss the evolutionary history of dinosaurs based on our residual diversity estimations and comparisons between these RDEs

and raw taxic diversity. Before doing so, however, some caveats regarding the interpretation of RDEs are warranted. Some deviations from the model's predictions will occur by chance, but the greater the distance between the straight line representing modelled diversity and a given data point representing observed diversity, the higher the probability that the observed diversity value genuinely departs from the model. The assumption that every fluctuation in RDE, however small, is in some way meaningful in macro-evolutionary terms, runs a severe risk of producing over-interpretations of the results. Barrett *et al.* (2009) addressed this issue by generating 95% confidence intervals, implemented based on the distribution of residual diversity, so that only approximately 5% of data points fell outside of the intervals. Thus, by definition, the vast majority of fluctuations in RDEs lie within the 95% confidence limits (*sensu* Barrett *et al.* 2009). If the confidence intervals generated by this approach are interpreted literally, very few deviations from the sampling model can be interpreted in terms of diversity change: however, this approach does not take account of the goodness of model fit, and is probably not appropriate. Moreover, it ignores the size of relative changes between stages. For example, one of the most dramatic changes in the RDEs presented by Barrett *et al.* (2009, fig. 2c) occurs when sauropodomorph diversity crosses the Jurassic/Cretaceous (J/K) boundary, yet this deflection lies entirely within the 95% confidence intervals. Given that a major decrease in sauropodomorph diversity occurs at the J/K boundary according to raw taxic data, phylogenetic diversity estimates, RDEs and rarefaction (Mannion *et al.* 2011), this would seem to be a genuine macro-evolutionary event that would be overlooked if the 95% confidence intervals (*sensu* Barrett *et al.* 2009) were strictly enforced. A more appropriate method for calculating confidence intervals is being formulated (G. T. Lloyd, pers. comm. 2011). For the present, however, we take the view that qualitative shifts in RDEs can still provide useful information on genuine diversity patterns and the impact of sampling, especially when RDEs are compared with raw taxic diversity and 'other sources of sampling-corrected' diversity data (e.g. subsampling curves). Nevertheless, the reader should be aware that many of the fluctuations in RDEs reported below are unlikely to qualify as statistically significant (at least in terms of Barrett *et al.*'s (2009) confidence intervals).

Late Triassic–Early Jurassic. The earliest unequivocal dinosaurs are known from the late Carnian (c. 230 Ma) Ischigualasto Formation of Argentina (e.g. *Herrerasaurus*, *Eoraptor*, *Panphagia*) and the Santa Maria Formation of Brazil (e.g.

Staurikosaurus, *Saturnalia*) (Brusatte *et al.* 2010; Langer *et al.* 2010), although footprints may extend the stratigraphical range of the group into pre-Carnian time (Langer *et al.* 2010, and references therein). Benton (1983, 1994) argued for an end-Carnian extinction event (now likely to be dated within the Norian, because of revisions to the Triassic time-scale (Mundill *et al.* 2010; Martinez *et al.* 2011)), in which the dominant herbivorous groups (rhynchosaurs and dicynodonts) were dramatically depleted in abundance/diversity. Benton suggested that this extinction was more significant for terrestrial vertebrates than the subsequent end-Triassic event (see below). He also hypothesized that the Carnian–Norian event opened ecological space that allowed sauropodomorphs to radiate opportunistically (see also Brusatte *et al.* 2008*a, b*), in contrast to the classical scenario involving a long-term competitive replacement of synapsids, rhynchosaurs, and crurotarsans by dinosaurs (e.g. Charig 1984). Dinosaurs became widespread in the Norian and Rhaetian, particularly sauropodomorphs (e.g. lower Elliot Formation, southern Africa; Los Colorados Formation, Argentina; Löwenstein Formation, Germany). One exception is North America (Chinle Group) where sauropodomorphs are absent, and theropods are the only dinosaur group known (Nesbitt *et al.* 2007; Brusatte *et al.* 2010; Langer *et al.* 2010) (N.B., several Late Triassic body fossil remains (Long & Murray 1995; Hunt *et al.* 1998; Harris *et al.* 2002) and tracks (Wilson 2005) from North America have been attributed to basal sauropodomorphs, but all of these were rejected by Nesbitt *et al.* (2007) on the basis of a lack of sauropodomorph synapomorphies, with most reinterpreted as indeterminate archosauriforms). Ornithischians are extremely scarce throughout the Late Triassic and may have been geographically restricted to southern Gondwana (Butler *et al.* 2007; Irmis *et al.* 2007; Nesbitt *et al.* 2007). A peak in dinosaurian diversity during the Late Triassic, especially the Norian, has been reported by most analyses, including those based on raw taxic data (e.g. Dodson 1990; Haubold 1990; and the current study – see Fig. 1), phylogenetic diversity estimates (PDEs) (e.g. Sereno 1997, 1999) and those that have employed more sophisticated subsampling and residual methods (e.g. Lloyd *et al.* 2008; Barrett *et al.* 2009; Mannion *et al.* 2011).

Throughout the Late Triassic and Early Jurassic, observed taxic diversity seems to track sampling very closely (Fig. 1), suggesting that the Norian peak is at least partly an artefact. However, global RDEs for dinosaurs (Fig. 11) also show a peak in the Norian relative to the Carnian and Rhaetian (N.B., it should be noted that this results from lower than expected diversity in the latter two stages rather than elevated levels of diversity in

the Norian: the latter has a RDE score just above zero). Examination of RDEs for the five geographical regions (Fig. 13*a–e*) indicates that this pattern of a Norian peak bracketed by Carnian and Rhaetian troughs is far from being a global phenomenon. European RDEs come closest to the ‘global pattern’, whereas Norian Asian diversity is much lower than expected and is considerably lower than that of either the Carnian or Rhaetian. This Asian ‘pattern’, however, is spurious because it is based on inadequate data (only one out of 37 Late Triassic dinosaur species is known from Asia). The ‘global’ pattern is apparently created by adding together disparate regional patterns in which Norian diversity is relatively higher everywhere except Asia; Carnian diversity is lower in North America, Africa and Europe; and Rhaetian diversity is lower in Asia and South America. In terms of the three main dinosaurian clades (Figs 12 & 14), the sauropodomorph RDEs conform most closely to the ‘general pattern’, although Norian peaks also occur in most RDEs for theropods and ornithischians as well. Given that sauropodomorphs contribute 59% of Late Triassic dinosaurian species, it is not surprising that a peak in the diversity of this clade should produce a peak in Norian dinosaur RDEs as a whole. Thus, the Norian peak in sauropodomorph diversity is consistent with Benton’s (1983, 1994) proposal that this clade radiated opportunistically after a Carnian/Norian boundary extinction of non-dinosaurian herbivores. However, claims regarding the severity of this extinction event and the dynamics of subsequent ecological replacements should be viewed with caution given new data on the stratigraphical ranges of many groups (Langer *et al.* 2010), the observation that this peak is perhaps exaggerated by sampling biases, and the apparent regional variation in the timing and magnitude of the Late Triassic peak in RDEs. The recent recalibration of the Triassic time-scale has made the Norian the longest stage in the Mesozoic, with a greater disparity in temporal length relative to its neighbours than almost any other. It has also resulted in repositioning many formerly ‘Carnian’ species as Norian ones, and enhanced the existence of a diversity ‘peak’. It is likely that the ability to place taxa more precisely within the Norian would reveal that many are not contemporaneous with one another (as is the case for many Campanian dinosaurs, Carrano 2008*b*). Finally, the drop in diversity during the Rhaetian should be treated with caution because of the uncertainties surrounding the dating of many Late Triassic deposits.

The Triassic/Jurassic (T/J) boundary mass extinction event has been identified as one of the ‘big five’ Phanerozoic mass extinctions, with estimates suggesting that up to 30% of ‘families’

became extinct (Benton 1995). Extinctions occurred in both terrestrial and marine realms across a breadth of taxonomic groups, and included the final loss of several primitive groups of crurotarsan or more basal archosauriforms (e.g. phytosaurs, rauisuchians, aetosaurs). Several previous studies (e.g. Lloyd *et al.* 2008; Barrett *et al.* 2009) have suggested that dinosaur diversity increased in the wake of this end-Triassic extinction. For example, Lloyd *et al.* (2008, fig. 2b) found evidence for an increase in dinosaur taxonomic diversity across the T/J boundary when subsampling approaches were used. They also found that the Early and Middle Jurassic marked a period of significantly elevated levels of diversification. Here, a Hettangian–Sinemurian peak in diversity, similar in magnitude to that witnessed in the Norian, occurs in the raw taxic diversity data (Fig. 1). This apparent peak in diversity corresponds closely with a peak in the number of DBCs (Fig. 1), suggesting that it may be artefactual. Genuine increases in diversity during the earliest Jurassic are supported by some sampling-corrected diversity estimates, but the details vary from clade to clade and depend on the analytical method utilized (see below). Our global RDEs for dinosaurs (Fig. 11) indicate a drop in diversity in the Rhaetian (but see above), followed by a relative recovery in the Early Jurassic (although diversity remains depressed relative to many other time bins). This pattern occurs in Asia and South America, whereas there are extinction events among dinosaurs at the T/J boundary in Africa, Europe and North America (Fig. 13a–e). This complex pattern may partly reflect the diversity of sauropodomorphs in each area during the Late Triassic–Early Jurassic transition. In regions where sauropodomorphs were scarce in the Late Triassic (e.g. Asia), Early Jurassic radiations among theropods and/or ornithischians would produce an apparent increase in dinosaur diversity after the T/J boundary. In contrast, regions where sauropodomorphs were diverse in the Late Triassic (i.e. Africa and Europe) may give the impression of an overall dinosaur diversity decrease at the T/J boundary because the losses experienced by this clade (see below) were not entirely compensated for by new Early Jurassic theropods and ornithischians. However, it is not clear why North America, which has no Late Triassic sauropodomorphs, displays an apparent decrease in diversity across the T/J boundary.

Much of the putative Early Jurassic dinosaurian radiation has been attributed to diversification events among theropods and ornithischians, with sauropodomorphs supposedly displaying few negative or positive effects of the T/J extinction event. For example, the immediate aftermath of the end-Triassic extinction in eastern North America is

marked by a notable increase in theropod body size (based upon the ichnological record: Olsen *et al.* 2002), which may reflect a global theropod radiation following the extinction of carnivorous crurotarsan lineages (Brusatte *et al.* 2010). However, Tanner *et al.* (2004, p. 113) noted that larger-bodied Late Triassic theropod body fossils are known from Europe (*Lilliensternus*) and even North America (*Gojirasaurus*), and Barrett *et al.* (2009) argued for a decline in theropod diversity, from the Late Triassic to Early Jurassic, based on declining RDE values. Both our DBC-based RDEs (Fig. 12c) and DBF-based RDEs (Fig. 14c) suggest that theropods underwent an increase in diversity (relative to the Rhaetian) in the two earliest stages of the Jurassic, although diversity remained relatively depressed during the Pliensbachian and Toarcian and did not start to increase markedly until the Aalenian. The fact that our DBF-based RDEs support this pattern, in contrast to Barrett *et al.*'s (2009) DBF-based RDEs, suggests that the disagreement between the latter authors and ourselves (plus Lloyd *et al.* 2008) stems from differences in data sets rather than methodological approach. One possibility is that the Lloyd *et al.* (2008) data set and the updated taxic data set used here are more similar to each other than they are to the data set of Barrett *et al.* (2009), which was based on data in Weishampel *et al.* (2004): some of the same gaps in the dinosaurian fossil record may have been filled since 2004, either by ghost range reconstruction (Lloyd *et al.* 2008) or via the addition of taxa discovered during the past 6–8 years (the current data set; see Fig. 15). It should also be noted that, outside of Europe, it is difficult to date most Early Jurassic dinosaurian taxa accurately to the stage level, so interpretations of diversity changes at this time should be treated with caution.

An Early Jurassic global radiation in the distribution, diversity and abundance of ornithischian dinosaurs has been attributed to an expansion into vacant ecological space (Butler *et al.* 2007). Barrett *et al.* (2009) also found that although residual diversity for ornithischians is negative in the Late Triassic, it displays a small positive peak in the earliest Jurassic. Both our DBC-based (Fig. 12a) and DBF-based (Fig. 14a) RDEs support the conclusion that ornithischians increased in diversity during at least the first two stages of the Jurassic, relative to the Late Triassic.

Brusatte *et al.* (2010) and Langer *et al.* (2010) have argued that the end-Triassic extinction appears to have had little impact on sauropodomorphs. Moreover, Barrett *et al.* (2009) found that there is substantial positive residual diversity for sauropodomorphs from the Norian through the Early Jurassic. Our DBC-based RDEs (Fig. 12b)

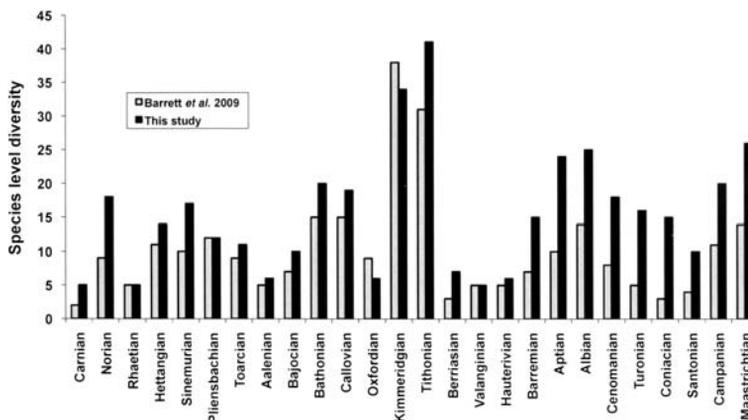


Fig. 15. Comparison of the number of sauropodomorph species per Standard European Stage in the Barrett *et al.* (2009) data set (open columns) and the data set used here (filled columns). N.B. the Barrett *et al.* data on sauropodomorph diversity was obtained from Upchurch *et al.* (2004), the latter being accurate only for taxa named prior to 2003.

support the view that sauropodomorphs experienced little diversity change at the T/J boundary. In the Early and Middle Jurassic, however, sauropodomorph diversity increases steadily to reach a peak in the Callovian: this could reflect the initial radiation of Eusauropoda and the origination of a number of neosauropod lineages (see below). DBF-based RDEs (Fig. 14b) also suggest that sauropodomorphs experienced an increase in diversity during the early stages of the Jurassic, but like theropods and ornithischians, Pliensbachian and Toarcian diversity is depressed relative to this peak and again in the Bathonian–Callovian. Thus, accurate reconstruction of diversity change among sauropodomorphs across the T/J boundary is difficult because the results depend on which analytical approach and/or sampling proxy are used. Sauropodomorph diversity was no doubt negatively affected by the loss of some basal forms at the end of the Triassic, but this seems to have been more than compensated for by their invasion of areas such as North America in the Early Jurassic, coupled with the onset of the eusauropod radiation.

Middle Jurassic to the J/K boundary. Raw taxic dinosaur diversity increases steadily from the beginning of the Middle Jurassic (Aalenian) to the end of the Jurassic, with a dip in the Oxfordian and dramatic decrease at the Jurassic/Cretaceous (J/K) boundary (Fig. 1). Sampling (in the form of DBCs) shows approximately the same trends during the Middle and Late Jurassic, suggesting that the observed diversity pattern may be artefactual. However, the various RDEs also confirm the occurrence of Middle and Late Jurassic diversity increases and a decrease in the earliest Cretaceous (Figs 11–14). The precise pattern of increases and

decreases varies depending on which RDEs are examined. Global DBC-based RDEs (Fig. 11) indicate that dinosaur diversity peaked in the Middle Jurassic (Bajocian–Callovian) and then gradually declined in the Late Jurassic, before a more dramatic extinction at the J/K boundary. The J/K boundary drop in RDEs is the third largest of 12 stage-to-stage decreases in our data set (surpassed in magnitude only by the drops in diversity during the Rhaetian and Oxfordian). This extinction is most clearly observed in Europe, North America and South America (Fig. 13a, c, d). In contrast, Asia and Africa show the Late Jurassic decline and Oxfordian extinction respectively, but there is little evidence for an extinction at the J/K boundary (Fig. 13b, e). DBC-based RDEs suggest that theropods and ornithischians underwent gradual reductions in diversity during the Late Jurassic followed by extinctions among the latter clade at the J/K boundary (Fig. 12a, c), whereas sauropodomorphs experienced peaks in diversity during the late Middle Jurassic (Bathonian and Callovian) and late Late Jurassic (Kimmeridgian and Tithonian), with more dramatic decreases in the Oxfordian and especially at the J/K boundary (Fig. 12b: as also noted by Upchurch & Barrett 2005; Mannion *et al.* 2011). DBF-based RDEs for theropods, sauropodomorphs and ornithischians suggest a more uniform increase in diversity towards the end of the Jurassic, and all display losses at the J/K boundary (Fig. 14) (although the losses among theropods are relatively small). Analysis of the theropod data without Mesozoic birds (Fig. 16) indicates a more profound decrease in diversity during the Late Jurassic and across the J/K boundary (e.g. the drop in DBC-based RDE values for non-avian dinosaurs becomes the largest of 11 stage-to-stage decreases).

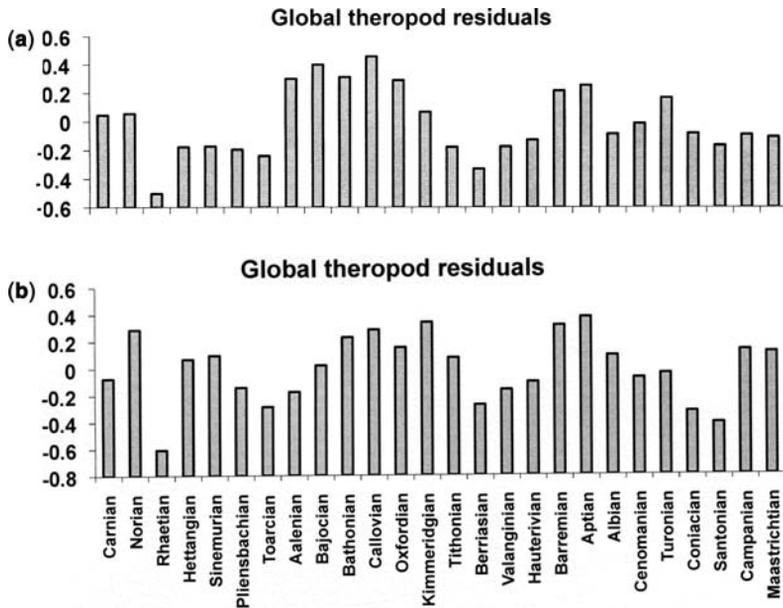


Fig. 16. Residual diversity estimates for non-avian theropods: (a) based on DBCs; (b) based on DBFs.

This implies some selectivity in the extinction whereby larger bodied non-avian theropods (and other large dinosaurs such as sauropods) were adversely affected and smaller volant forms remained unscathed or even diversified. In general, these results are in broad agreement with the DBF-based RDEs of Barrett *et al.* (2009), which identified marked drops in diversity in all three clades in the Oxfordian, decreases in sauropodomorph and ornithischian diversity at the J/K boundary, and a more gradual decrease in theropod diversity through the Late Jurassic and into the Early Cretaceous.

Raup & Sepkoski (1986) identified the J/K boundary event as one of eight major extinctions that have occurred during the last 250 million years. Subsequently, however, most studies have regarded this event as only a minor extinction (e.g. Hallam & Wignall 1997; Bambach 2006; Arens & West 2008). Recent studies of dinosaur diversity have either not commented on a J/K extinction, or have argued that this event strongly affected sauropods (Upchurch & Barrett 2005; Mannion *et al.* 2011), but probably had little impact on the diversity of ornithischians and theropods (e.g. Barrett *et al.* 2009). However, Orcutt *et al.*'s (2007) work on terrestrial tetrapods, and Benson *et al.*'s (2010) study of marine reptiles (see also Bakker 1993; Bardet 1994), suggests that the J/K boundary extinction event may have been more significant and widespread than previously realized. Both our raw taxic

data and RDEs for dinosaurs support this view and indicate that, while sauropodomorphs and stegosaurs (Bakker 1978; Galton & Upchurch 2004; Maidment *et al.* 2008) might have been particularly strongly affected, many other dinosaur clades probably also declined in diversity at this time.

Cretaceous diversity and the end-Cretaceous extinction. After a period of apparent lowered diversity in the Early Cretaceous, raw taxic diversity of dinosaurs displays prominent peaks in the Aptian–Albian and the Campanian–Maastrichtian, with an apparent trough in the Cenomanian–Santonian (Fig. 1). However, this pattern of peak-trough-peak follows the sampling of DBFs and DBCs very closely. The various RDEs generally agree that diversity increased throughout much of the Early Cretaceous, with peaks in the Barremian–Albian (Figs 11, 12a, c, 13 & 14), although in Africa the mid-Cretaceous peak occurs in the Cenomanian (Fig. 13e). One exception to this is the DBC-based RDEs for sauropodomorphs (Fig. 12b), which shows a steady increase in diversity from the beginning of the Cretaceous to a peak in the Coniacian (see also the results of runs test Q6), followed by a decline into the Campanian and a rise in the Maastrichtian.

Patterns in dinosaur diversity during the early Late Cretaceous are more difficult to elucidate. In general, DBC-based RDEs display moderately lower diversity during this time compared to the

Barremian–Aptian, with some results suggesting that a peak in diversity may have occurred during the Turonian (e.g. global dinosaur RDEs, European RDEs, Asian RDEs). DBF-based RDEs show a similar pattern, but lack the Turonian peak and indicate a more profound decrease in diversity in the Cenomanian–Santonian, especially during the Coniacian–Santonian (Fig. 14). Thus, the very low raw taxic diversity levels throughout the early Late Cretaceous are mainly the result of poorer sampling (particularly in the case of the Cenomanian and Turonian), but a genuine drop in diversity seems to have occurred in the Coniacian and Santonian.

The end-Cretaceous or Cretaceous/Palaeogene (K/Pg) mass extinction has received more attention from palaeobiologists than almost any other event in Earth history. Whether non-avian dinosaurs went extinct abruptly at the K/Pg boundary, or the final extinction was preceded by a slow decline during the Campanian and/or Maastrichtian, remains controversial. Many studies based on raw taxic data have concluded that dinosaurs underwent a gradual terminal decline prior to the extinction (e.g. Sloan *et al.* 1986; Sarjeant & Currie 2001; Sullivan 2006), although Sheehan *et al.* (1991) and Pearson *et al.* (2002) argued for a sudden extinction. Most recent studies, in which sampling biases have been addressed, have concluded that there is little evidence for a gradual decline (e.g. Fastovsky *et al.* 2004; Wang & Dodson 2006; Carrano 2008a; Lloyd *et al.* 2008; Mannion *et al.* 2011). For example, Carrano (2008a) carried out a regional study based on North American dinosaurs: he concluded that fluctuations in observed diversity during the Campanian and Maastrichtian largely reflect ecological differences between geological units and sampling issues, rather than long-term macro-evolutionary changes. He also noted that lumping all Campanian dinosaurs together into a single diversity count obscured the fact that many species were not contemporaneous, but members of successive faunas. One exception to this recent trend is the work of Barrett *et al.* (2009): these authors argued that their DBF-based RDEs indicated gradual declines in dinosaur diversity in the lead up to the K/Pg boundary (although detailed inspection of their results indicates that sauropodomorph diversity does not display a clear decrease, and theropods display a far less dramatic decline relative to ornithischians, across the Campanian/Maastrichtian boundary).

Our global RDEs for dinosaurs suggest that the diversity of this group in the latest Cretaceous was generally similar to that seen in the Late Jurassic (i.e. lower than in the Middle Jurassic and Barremian–Aptian, but higher than in the earliest Cretaceous), with a small decrease from the Campanian to the Maastrichtian (Fig. 11). This decrease

from the Campanian to the Maastrichtian is consistent with the conclusions of Barrett *et al.* (2009). However, the RDEs of the major clades within Dinosauria indicate a more complex pattern. Both DBC-based and DBF-based RDEs display a marked decrease for ornithischians (Figs 12a & 14a), virtually no decrease for theropods (Figs 12c & 14c), and an increase for sauropodomorphs (Figs 12b & 14b), from the Campanian to the Maastrichtian. Moreover, the fluctuations in RDEs for dinosaurs as a whole (and for the three major clades) in the latest Cretaceous are no greater than those observed during other periods of their evolutionary history, and indeed are often smaller. For example, the decrease in DBC-based RDEs from the Campanian to the Maastrichtian is only the sixth largest out of 12 stage-to-stage decreases during the Mesozoic.

The final extinction of the non-avian dinosaurs is also more complicated than expected when examined at the regional scale. The RDEs for Europe and Africa display marked declines from the Campanian to the Maastrichtian (Fig. 13a, e). However, in South America and Asia, dinosaurian RDEs remain fairly stable from the Campanian to the Maastrichtian (Fig. 13b, d), and the RDE for North America actually displays an increase in the Maastrichtian (Fig. 13c). The latter pattern might be regarded as particularly significant, given the fact that North America possesses the best sampled record of latest Cretaceous dinosaurs.

In summary, analyses of Late Cretaceous dinosaur diversity that take sampling biases into account, generally agree that there is little support for a gradual decline leading up to the K/Pg boundary. This conclusion has been reached by both regional and global studies, and through the application of rarefaction (Fastovsky *et al.* 2004; Mannion *et al.* 2011), abundance-based coverage estimation (Wang & Dodson 2006), and sampling metrics based on DBCs, DBFs and the numbers of dinosaur-bearing localities (Lloyd *et al.* 2008; the current study). Until the current work, the disagreement between previous studies and Barrett *et al.* (2009) could have been attributed to the latter's unique use of DBF-based residuals. However, the same approach has been used here and failed to support clear pre-K/Pg declines in theropods and sauropodomorphs (although Ornithischia may still conform to the gradualistic extinction scenario). Failure to find support for consistent Campanian–Maastrichtian drops in DBF-based RDEs (as well as our DBC-based RDEs) may reflect the effects of an influx of new data on the Late Cretaceous since 2004. It is interesting to note, for example, that both DBC-based and DBF-based RDEs for sauropodomorphs (Figs 12 & 14) indicate an increase in the diversity of this clade from the

Campanian to the Maastrichtian, which potentially reflects the discovery of numerous new Late Cretaceous sauropods in the past eight years: compare the numbers of sauropod taxa in the Campanian and Maastrichtian available to Barrett *et al.* (2009) and the current study (Fig. 15). Finally, one caveat should be noted: the current study has examined dinosaur diversity using stage-level time bins, rather than the substages employed by Barrett *et al.* (2009). As such, we can comment on diversity fluctuations from the Campanian to the Maastrichtian, but have no information on changes within the Maastrichtian. If, for example, dinosaur diversity was relatively high in the early Maastrichtian, and then underwent a gradual decline during the middle and late Maastrichtian, our data would fail to capture this pattern and would produce a diversity estimate that is averaged across this stage. Thus, our results contradict a long-term gradual decline in dinosaur diversity during the 15 million years of the Campanian–Maastrichtian, but this still leaves open the possibility that a gradual decline occurred globally over the final 3–4 million years prior to the K/Pg boundary (but see Fastovsky *et al.* 2004).

Long-term trends during the Mesozoic. The raw taxic diversity of dinosaurs as a whole, and theropods and ornithischians separately, generally increases throughout the Jurassic and Cretaceous, culminating in a large peak in the Campanian–Maastrichtian (Figs 1 & 2). Sauropodomorphs follow a slightly different pattern: essentially, their observed diversity increases during the Jurassic, suffers a serious set-back at the J/K boundary (see above), and then recovers during the Cretaceous. This apparent general increase in raw taxic diversity raises the question as to whether this represents a genuine evolutionary phenomenon (i.e. dinosaur diversification rates increased through time, perhaps conforming to an exponential growth curve), or is an artefact of sampling (e.g. the tendency for younger sedimentary deposits to be preserved in greater abundance than older ones). The analyses of Fastovsky *et al.* (2004) and Wang & Dodson (2006) concluded that a general increase in diversity can still be detected even after uneven sampling rates are taken into account. However, this pattern of long-term growth in diversity is not supported by other recent studies and the analyses presented here. Lloyd *et al.* (2008) noted that the Campanian peak in observed diversity is somewhat reduced when ghost range data and sampling metrics are considered. The results of our runs tests (analyses Q) provide very little support for the existence of a persistent upward trend in dinosaur diversity. For example, none of the DBF-based RDEs, or any of the RDEs for ornithischians, demonstrate

the presence of a trend of increasing diversity during the Mesozoic or the Triassic–Jurassic and Cretaceous time slices (see runs tests Q10–Q21). Although the runs test for DBC-based RDEs for all Mesozoic dinosaurs (Q1) does support the presence of such an upward trend, time-sliced analyses (Q2, Q3) show that this result is the product of combining increasing diversity through the Triassic–Jurassic with no detectable trend in the Cretaceous. It appears that theropods and sauropodomorphs may have experienced an upward trend in diversity during the former time slice, whereas in the Cretaceous only sauropods show such a trend (Q4–Q9). Importantly, the Triassic–Jurassic and Cretaceous trends in sauropodomorph diversity are not part of one continuous pattern, rather they are two distinct trends separated by the J/K boundary extinction that reset their diversity to low levels during the Early Cretaceous.

Changes in RDEs, and the results of the runs tests, suggest that the long-term history of dinosaur diversity is better characterised as a series of radiations punctuated by ‘set-backs’, rather than a steady (perhaps exponential) rise towards an end-Cretaceous peak (contrast this model, for example, with Fastovsky *et al.* 2004, fig. 1). Key phases in dinosaur evolution would thus include:

- (i) Divergences into the three main clades, and major radiation of sauropodomorphs, during the Late Triassic (especially the Carnian–Norian).
- (ii) A marked decrease in diversity during the Rhaetian (although this might have been exaggerated by uncertainties in the dating of Late Triassic rocks, perhaps ‘back-smearing’ the effects of a T/J boundary mass extinction event).
- (iii) Major radiations of all three clades during the Early and Middle Jurassic, with theropods and sauropodomorphs displaying statistically detectable upward trends in diversity (see below).
- (iv) A gradual decline in diversity through the Late Jurassic (theropods and ornithischians) combined with a more severe extinction event at the J/K boundary.
- (v) A recovery phase in the Early Cretaceous, culminating in peaks in diversity during the Barremian–Albian.
- (vi) A moderate decline in diversity, reaching a low point in the Coniacian–Santonian.
- (vii) Moderately increased diversity in the Campanian–Maastrichtian, driven especially by the radiation of titanosaur sauropods (only sauropods display a statistically significant upward trend in diversity during the Cretaceous). Despite the appearance of many new forms

of theropod (e.g. Cretaceous birds) and ornithischians (radiations of ceratopsians, hadrosaurs and ankylosaurs), these were apparently insufficient to result in a persistent upward trend in diversity, perhaps because originations were offset by high extinction rates. Consequently, although sauropod diversity during the Campanian–Maastrichtian recovers to levels comparable to those of the Late Jurassic (see also Mannion *et al.* 2011), the end of the Cretaceous does not have abnormally elevated levels of diversity when compared to earlier peaks, such as those that occur during the Middle Jurassic and mid-Cretaceous (Figs 11–14).

- (viii) The extinction of all non-avian dinosaurs at the Cretaceous/Palaeogene boundary.

The radiation of dinosaurs during the Early and Middle Jurassic could be regarded as one of the most important phases in dinosaur evolution, with significant upward trends in theropod and sauropod diversity, and the first clear radiations of theropods and ornithischians. The proposal that the Early and Middle Jurassic represents a key phase in dinosaurian diversification is consistent with several other recent studies and lines of evidence. For example, ghost ranges indicate that many of the distinct dinosaurian clades originated prior to the late Middle Jurassic (Upchurch & Barrett 2005) and this is reflected in the discovery of early members of clades that were originally found in the Cretaceous but are now known from the Jurassic (e.g. tyrannosaurs (Xu *et al.* 2006; Benson 2008), deinonychosaurs (Hu *et al.* 2009), titanosaurs (Day *et al.* 2002, 2004), and ankylosaurs (Carpenter *et al.* 1998)). Lloyd *et al.* (2008) also used a time-calibrated dinosaurian supertree to argue that diversification rates peaked during the Middle Jurassic but were no greater than background rates during most of the Cretaceous. Finally, Upchurch *et al.* (2002) suggested that the presence of continent-scale vicariance patterns among Late Jurassic and Cretaceous dinosaurs is consistent with the dispersal of members of most clades across Pangaea prior to the onset of its fragmentation in the Middle Jurassic. In short, it seems probable that the true pattern of dinosaurian diversification is ‘bottom heavy’ (i.e. cladogenetic events are concentrated in the early phases of dinosaur evolution), but the impression of a more ‘top heavy’ history (i.e. high diversity concentrated towards the middle and Late Cretaceous) is an artefact generated by better preservation of younger fossiliferous deposits, coupled with the more intensive sampling of Campanian–Maastrichtian rocks because of disproportionate interest in the K/Pg boundary extinction.

Conclusions, caveats and prospects

The results of this study have the following implications for our understanding of dinosaur evolutionary history and the effects of sampling in the terrestrial realm on palaeodiversity patterns:

- (i) The observed taxic diversity of dinosaurs is positively correlated with the numbers of dinosaur-bearing collections and dinosaur-bearing formations. Correlations persist even when the data set is partitioned into the clades Theropoda, Sauropodomorpha and Ornithischia, or into continental regions (i.e. Africa, Asia, Europe, and North and South America). It is therefore difficult to escape the conclusion that many of the fluctuations in observed diversity actually represent artefacts generated by uneven sampling.
- (ii) Sampling regimes in the terrestrial realm did not cause the observed diversity patterns of separate regions to be artificially more or less similar to each other.
- (iii) The diversity histories of dinosaur faunas on different continents apparently varied from region to region. Global environmental change may have imposed some congruence in terms of the timing and magnitude of radiations and extinctions, but this is not reflected in significant positive correlations between the diversity patterns in separate regions.
- (iv) Regional sampling metrics are not particularly effective at capturing global sampling signals in the terrestrial realm. This might reflect regional variations in sampling rates, so that the ‘global signal’ is in fact a summary of disparate regional patterns. In contrast, regional sampling metrics are typically strongly positively correlated with regional observed diversity. These results suggest that, if a regional sampling metric is found to correlate with supposed global diversity, it is important to rule out the possibility that the ‘global’ taxonomic data set is biased in terms of over-representing taxa from the region, which has yielded the sampling metric.
- (v) Different sampling metrics for the same region (i.e. DBFs, DBCs, rock units and collections for North America, western European terrestrial sedimentary rock outcrop area) typically correlate with each other, suggesting that they are ‘homing in’ on approximately the same sampling signal. Nevertheless, it should also be noted that choice of sampling metric could affect the details of residual diversity estimates (e.g. the extent to which sauropod diversity was

relatively depressed during the Cenomanian–Coniacian) depends on whether DBFs or DBCs are used to produce RDEs (compare Figs 12b & 14b).

- (vi) Concerns that correlations between observed diversity and sampling metrics result from circularity, rather than the controlling effects of sampling, are addressed by the observation that numbers of DBFs and DBCs correlate with other sampling metrics such as rock outcrop area and gap-bound packages. Moreover, correlations between observed diversity and sampling metrics persist even when the criterion for recognising a unit of sampling (e.g. counting all formations containing any evidence of the presence of dinosaurs, even trackways or fragmentary body fossils) is far broader than the taxonomic group under investigation (e.g. Theropoda).
- (vii) The three main dinosaurian clades appear to have experienced early bursts in diversification rates at different times. The initial sauropodomorph diversification occurred during the Late Triassic (Norian), followed by strong radiations of theropods and ornithischians in the Early Jurassic.
- (viii) Both the Rhaetian and Oxfordian appear to be times of genuinely low diversity for most dinosaurian clades according to raw TDEs and RDEs. However, both of these stages are affected by dating issues whereby deposits that should be assigned to them may have been incorrectly dated as Norian in the case of the Rhaetian and Kimmeridgian in the case of the Oxfordian (e.g. see Mannion *et al.* 2011).
- (ix) An extinction among dinosaurs apparently occurred at the Jurassic/Cretaceous boundary. Previously this was thought to have mainly affected sauropodomorphs, but all three of the main dinosaurian clades display either a decline in diversity throughout the Late Jurassic and into the Early Cretaceous (non-avian theropods) or a more dramatic decline at the J/K boundary itself (sauropodomorphs and ornithischians). Combined with recent work on marine reptile diversity (Bakker 1993; Bardet 1994; Benson *et al.* 2010), these results suggest that the J/K boundary extinction event was probably more significant than previously acknowledged.
- (x) There is little evidence for a significant gradualistic decline in dinosaur diversity during the latest Cretaceous. Where a decrease in diversity from the Campanian to the Maastrichtian does occur, such declines are no larger (and are typically much smaller) than declines in diversity that were experienced earlier in dinosaur evolution. Regional

residual diversity estimates, and the RDEs for clades within Dinosauria, often display either no decrease, or even an increase in diversity, from the Campanian to Maastrichtian. These observations do not rule out a gradual decline in dinosaur diversity prior to the K/Pg boundary event, but if such a decline did occur, it must have happened during the middle and/or late Maastrichtian, with relatively high diversity levels in the early Maastrichtian.

- (xi) Dinosaur diversity does not display a gradual increase throughout the Mesozoic, culminating in a Campanian–Maastrichtian peak, as has been claimed by several previous studies (e.g. Sereno 1997, 1999; Fastovsky *et al.* 2004; Wang & Dodson 2006). This apparent long-term trend is probably an artefact generated by better sampling available in younger rocks, combined with disproportionate sampling effort regarding Campanian–Maastrichtian rocks by workers interested in the causes of the K/Pg mass extinction. When sampling biases are taken into account, the history of dinosaur diversity is characterised by a series of growth phases (i.e. Carnian–Norian, Early and Middle Jurassic, Early Cretaceous to Barremian–Albian and Campanian) punctuated by gradual decreases and occasional severe extinction events (i.e. the Rhaetian, Oxfordian, J/K boundary and Coniacian–Santonian). This pattern is consistent with the work of Lloyd *et al.* (2008), who noted a statistically significant increase in diversity rates during the Early and Middle Jurassic, followed by background rates of diversification throughout the rest of dinosaurian evolutionary history.

There are a number of caveats that should be borne in mind when analysing data sets of terrestrial vertebrates and when applying methods such as residual diversity estimation. Although statistical comparisons are a vital component of quantitative palaeobiology, they can obscure important details as well as reveal patterns. One reason for this is that which patterns are found or not found depends heavily on how the researcher partitions their data for analysis. For example, here we have used five continental areas in order to examine regional patterns in diversity: it is hoped that these five regions represent genuine geographical units that had some kind of biological reality during the Mesozoic. However, palaeogeographical events mean that, for example, Africa and South America might be considered one area during the Early Jurassic but are clearly two distinct regions during the Late Cretaceous.

Similarly, the search for ‘signal’ in large complex data sets creates a dilemma: analysis of all or most of the data means that ‘patterns’ can be identified with greater statistical rigour, but the ‘general pattern’ may obscure many important secondary signals (e.g. separate clades showing different responses to the same environmental events). However, attempts to locate these secondary signals by partitioning the data set results in fewer data points for analysis, which may make it more difficult to resolve true signal against the backdrop of ‘noise’ (e.g. see the ‘Discussion’ in Mannion & Upchurch 2010*b*). Thus, our choice to look at diversity change among Dinosauria, Theropoda, Sauropodomorpha and Ornithischia, has probably obscured important events that affected smaller clades (e.g. the extinction of many stegosaurs at the J/K boundary and the radiation of ankylosaurs during the Cretaceous).

Another problem is that, while RDEs offer one of the best ways to remove the effects of uneven sampling from raw diversity data, the resulting fluctuations in RDEs may be over-interpreted: many such fluctuations lie within the bounds of statistical ‘noise’. Much of our interpretation and discussion has hinged on whether two data series pass or fail statistical tests. However, two data series may correlate well over part of the time range and not correlate at all over the remainder (as was found by Benson *et al.* (2010) and Mannion *et al.* 2011). Thus, while we have found that there is little evidence for a common global pattern of dinosaurian diversification, this is a general statement and really means that there is not enough agreement among the diversity patterns from each region to allow a statistical pass. In reality, there may still be many peaks and troughs (in either observed diversity or RDEs) that coincide, but these are not quite sufficient to produce positive correlations.

Finally, in this paper, we have discussed increases and decreases in diversity in terms of diversification and extinction events respectively. However, the majority of dinosaur genera and species are point occurrences in terms of their stratigraphical ranges, or have almost certainly had their ranges artificially truncated by poor sampling: thus, it is not possible to calculate meaningful origination and extinction rates. Consequently, the true dynamics of diversity fluctuations during dinosaur evolution are difficult to elucidate using the currently available data. For example, the J/K boundary reduction in standing diversity could reflect a dramatic increase in extinction rates, or it might stem from a decrease in origination rates (with extinction rates remaining largely unchanged) after the Early–Middle Jurassic ‘burst’ of diversification.

Our results and conclusions suggest a number of lines of future study that may reveal important

insights into dinosaur evolution, terrestrial sampling, and the methods we use to deal with sampling and diversity. First, there are several other ways in which the data set could be partitioned (e.g. comparisons of Northern and Southern Hemisphere patterns, time-slicing into Jurassic and Cretaceous subsets, etc.). Second, although the current data set is the largest available for terrestrial Mesozoic organisms, it remains narrow relative to the full spectrum of available tetrapods. This raises questions such as: do any of the observed patterns in dinosaur diversity also occur in other tetrapods? Does the addition of other non-dinosaurian groups reinforce or contradict apparent regional differences in diversity patterns? Does the observed diversity of terrestrial vertebrates as a whole correlate with sampling of the terrestrial fossil record, rock outcrop area, etc.? Finally, the estimation, and especially the correction, of the effects of uneven sampling on observed diversity patterns, has yet to overcome several methodological challenges. In particular, debate still rages over whether or not sampling has produced significant distortions of observed diversity, and even if this has occurred, there is the possibility that our methods for creating sampling-corrected diversity curves may produce additional distortions that result from data transformations. At present, many workers seem to fall into one of two camps – subsampling (e.g. Alroy *et al.* 2008) *v.* residuals (e.g. Smith & McGowan 2007) – but it is often feasible to apply both methods to the same data set. Indeed, on one of the few occasions when this has been implemented (Mannion *et al.* 2011), subsampling and residuals produced very similar reconstructions of diversity fluctuations in sauropodomorphs, which may be reassuring for those palaeobiologists who are more concerned with understanding the evolution of their study organisms than they are with methodological issues.

This paper presents the results of just one case study of diversity and sampling in the terrestrial realm, and it is obviously dangerous to overstate the generality of its conclusions. Nevertheless, it demonstrates the need to take sampling into account when reconstructing the diversity history of terrestrial Mesozoic organisms, and shows that some intriguing, unexpected and thought-provoking conclusions can result. We hope that this study will prompt further analyses by those wishing to either build upon our results or to overturn them. Such studies will play a key role in capturing the complexity of evolutionary patterns at regional and global scales.

This paper is a Paleobiology Database official publication No. 133. PU gratefully acknowledges support from the Abbey International Collaboration Grant scheme (funded

by Santander Bank and administered by University College London), and a Palaeontological Association Research Grant, which enabled collection of data on sauropodomorph dinosaurs. PDM was supported by a University College London NERC studentship (NER/S/A/2006/14347) and a grant from the Jurassic Foundation. RJB is supported by an Alexander von Humboldt Research Fellowship. MTC received support from the Evolution of Terrestrial Ecosystems (ETE) program at the Smithsonian Institution; this is ETE publication 231. We also thank A. Smith and A. J. McGowan for their invitation to talk at the Lyell meeting (part of the 3rd International Palaeontological Congress, London, 2010) and to participate in this edited volume.

References

- ALROY, J. 2000. Successive approximations of diversity curves: ten more years in the library. *Geology*, **28**, 1023–1026, doi: [http://dx.doi.org/10.1130/0091-7613\(2000\)28<1023:SAODCT>2.0.CO;2](http://dx.doi.org/10.1130/0091-7613(2000)28<1023:SAODCT>2.0.CO;2).
- ALROY, J., MARSHALL, C. R. *ET AL.* 2001. Effects of sampling standardization on estimates of Phanerozoic marine diversification. *Proceedings of the National Academy of Sciences, USA*, **98**, 6261–6266.
- ALROY, J., ABERHAN, M. *ET AL.* 2008. Phanerozoic trends in the global diversity of marine invertebrates. *Science*, **321**, 97–100, doi: 10.1126/science.1156963.
- ARENS, N. C. & WEST, I. D. 2008. Press-pulse: a general theory of mass extinction? *Paleobiology*, **34**, 456–471.
- BAKKER, R. T. 1978. Dinosaur feeding behaviour and the origin of flowering plants. *Nature*, **274**, 661–663.
- BAKKER, R. T. 1993. Plesiosaur extinction cycles – events that mark the beginning, middle and end of the Cretaceous. *Geological Association of Canada, Special Papers*, **39**, 641–664.
- BAMBACH, R. K. 2006. Phanerozoic biodiversity: mass extinctions. *Annual Review of Earth and Planetary Science*, **34**, 127–155.
- BARDET, N. 1994. Extinction events among Mesozoic marine reptiles. *Historical Biology*, **7**, 313–324.
- BARRETT, P. M., MCGOWAN, A. J. & PAGE, V. 2009. Dinosaur diversity and the rock record. *Proceedings of the Royal Society, B*, **276**, 2667–2674, doi: 10.1098/rspb.2009.0352
- BENSON, R. B. J. 2008. New information on *Stokesosaurus*, a tyrannosauroid (Dinosauria: Theropoda) from North America and the United Kingdom. *Journal of Vertebrate Paleontology*, **28**, 732–750.
- BENSON, R. B. J., BUTLER, R. J., LINDGREN, J. & SMITH, A. S. 2010. Mesozoic marine tetrapod diversity: mass extinctions and temporal heterogeneity in geological megabiases affecting vertebrates. *Proceedings of the Royal Society, B*, **277**, 829–834, doi: 10.1098/rspb.2009.1845
- BENSON, R. B. J. & BUTLER, R. J. 2011. Uncovering the diversification history of marine tetrapods: ecology influences the effect of geological sampling biases. In: MCGOWAN, A. J. & SMITH, A. B. (eds) *Comparing the Geological and Fossil Records: Implications for Biodiversity Studies*. Geological Society, London, Special Publications, **358**, 191–208.
- BENTON, M. J. 1983. Dinosaur success in the Triassic: a noncompetitive ecological model. *Quarterly Review of Biology*, **58**, 29–55.
- BENTON, M. J. 1994. Late Triassic to Middle Jurassic extinctions among continental tetrapods: testing the pattern. In: FRASER, N. C. & SUES, H.-D. (eds) *In the Shadow of the Dinosaurs*. Cambridge University Press, Cambridge, 366–397.
- BENTON, M. J. 1995. Diversification and extinction in the history of life. *Science*, **268**, 52–58.
- 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.
- BENTON, M. J., TVERDOKHLEBOV, V. P. & SURKOV, M. V. 2004. Ecosystem remodelling among vertebrates at the Permian–Triassic boundary in Russia. *Nature*, **432**, 97–100.
- BENTON, M. J., DUNHILL, A. M., LLOYD, G. T. & MARX, F. G. 2011. Assessing the quality of the fossil record: insights from vertebrates. In: MCGOWAN, A. J. & SMITH, A. B. (eds) *Comparing the Geological and Fossil Records: Implications for Biodiversity Studies*. Geological Society, London, Special Publications, **358**, 63–94.
- BRUSATTE, S. L., BENTON, M. J., RUTA, M. & LLOYD, G. T. 2008a. Superiority, competition, and opportunism in the evolutionary radiation of dinosaurs. *Science*, **321**, 1485–1488.
- BRUSATTE, S. L., BENTON, M. J., RUTA, M. & LLOYD, G. T. 2008b. The first 50 mya of dinosaur evolution: macroevolutionary pattern and morphological disparity. *Biology Letters*, **4**, 733–736.
- BRUSATTE, S. L., NESBITT, S. J., IRMIS, R. B., BUTLER, R. J., BENTON, M. J. & NORELL, M. A. 2010. The origin and early radiation of dinosaurs. *Earth-Science Reviews*, **101**, 68–100.
- 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. *Paleobiology*, **35**, 432–446, doi: 10.1666/0094-8373-35.3.432.
- BUTLER, R. J., SMITH, R. M. H. & NORMAN, D. B. 2007. A primitive ornithischian dinosaur from the Late Triassic of South Africa, and the early evolution and diversification of Ornithischia. *Proceedings of the Royal Society, B*, **274**, 2041–2046.
- BUTLER, R. J., BENSON, R. B. J., CARRANO, M. T., MANNION, P. D. & UPCHURCH, P. 2011. Sea level, dinosaur diversity, and sampling biases: investigating the ‘common cause’ hypothesis in the terrestrial realm. *Proceedings of the Royal Society, B*, **278**, 1165–1170, doi: 10.1098/rspb.2010.1754
- CARPENTER, K., MILES, C. & CLOWARD, K. 1998. Skull of a Jurassic ankylosaur (Dinosauria). *Nature*, **393**, 782–783.
- CARRANO, M. T. 2005. The dinosaur fossil record. *Journal of Vertebrate Paleontology*, **25** (Suppl. 3), 42A.
- CARRANO, M. T. 2008a. Patterns of diversity among latest Cretaceous dinosaurs in North America. *Journal of Vertebrate Paleontology*, **28** (Suppl. 3), 61A.
- CARRANO, M. T. 2008b. Taxonomy and classification of non-avian Dinosauria. *Paleobiology Database Online Systematics Archive*, **4**, www.paleodb.org

- CHARIG, A. J. 1984. Competition between therapsids and archosaurs during the Triassic Period: a review and synthesis of current theories. *Symposium of the Zoological Society of London*, **52**, 597–628.
- 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.
- DAY, J. J., UPCHURCH, P., NORMAN, D. B., GALE, A. S. & POWELL, H. P. 2002. Sauropod Trackways: evolution and behavior. *Science*, **296**, 1659.
- DAY, J. J., NORMAN, D. B., GALE, A. S., UPCHURCH, P. & POWELL, H. P. 2004. New Middle Jurassic dinosaur trackways from Oxfordshire, U.K. *Palaeontology*, **47**, 319–348.
- DODSON, P. 1990. Counting dinosaurs: how many kinds were there? *Proceedings of the National Academy of Science, USA*, **87**, 7608–7612.
- DYKE, G. J., MCGOWAN, A. J., NUDDS, R. L. & SMITH, D. 2009. The shape of pterosaur evolution: evidence from the fossil record. *Journal of Evolutionary Biology*, **22**, 890–898.
- FASTOVSKY, D. E., HUANG, Y., HSU, J., MARTIN-MCNAUGHTON, J., SHEEHAN, P. M. & WEISHAMPEL, D. B. 2004. Shape of Mesozoic dinosaur richness. *Geology*, **32**, 877–880.
- 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
- GALTON, P. M. & UPCHURCH, P. 2004. Stegosauria. In: WEISHAMPEL, D. B., DODSON, P. & OSMOLSKA, H. (eds) *The Dinosauria II*. University of California Press, Berkeley and Los Angeles, 343–362.
- GOSWAMI, A. & UPCHURCH, P. 2010. The dating game: a reply to Heads (2010). *Zoologica Scripta*, **39**, 406–409.
- GRADSTEIN, F. M., OGG, J. G. & SMITH, A. G. 2005. *A Geologic Time Scale 2004*. Cambridge University Press, Cambridge.
- HALLAM, A. & WIGNALL, P. B. 1997. *Mass Extinctions and their Aftermath*. Oxford University Press, Oxford.
- HAMMER, Ø. & HARPER, D. A. T. 2006. *Paleontological Data Analysis*. Blackwell, Oxford.
- HAMMER, Ø., HARPER, D. A. T. & RYAN, P. D. 2001. PAST: Palaeontological Statistics software package for education and data analysis. *Palaeontologica Electronica*, **4** (1).
- HARRIS, S. K., HECKERT, A. B., LUCAS, S. G. & HUNT, A. P. 2002. The oldest North American prosauropod, from the Upper Triassic Tecovas Formation of the Chinle Group (Adamanian: Latest Carnian), West Texas. *New Mexico Museum of Natural History & Science Bulletin*, **21**, 249–252.
- HAUBOLD, H. 1990. Dinosaurs and fluctuating sea levels during the Mesozoic. *Historical Biology*, **4**, 75–106.
- HEADS, M. 2010. Evolution and biogeography of primates: a new model based on molecular phylogenetics, vicariance and plate tectonics. *Zoologica Scripta*, **39**, 107–127.
- HU, D., HOU, L., ZHANG, L. & XU, X. 2009. A pre-*Archaeopteryx* troodontid theropod from China with ling feathers on the metatarsus. *Nature*, **461**, 640–643.
- HUNT, A. P., LUCAS, S. G., HECKERT, A. B., SULLIVAN, R. M. & LOCKLEY, M. G. 1998. Late Triassic dinosaurs from the western United States. *Geobios*, **31**, 511–531.
- IRMIS, R. B., PARKER, W. G., NESBITT, S. J. & LIU, J. 2007. Early ornithischian dinosaurs: the Triassic record. *Historical Biology*, **19**, 3–22.
- JACKSON, J. B. C. & JOHNSON, K. G. 2001. Measuring past biodiversity. *Science*, **293**, 2401–2404.
- LANE, A., JANIS, C. M. & SEPKOSKI, J. J., JR. 2005. Estimating paleodiversities: a test of the taxic and phylogenetic methods. *Paleobiology*, **31**, 21–34.
- LANGER, M. C., EZCURRA, M. D., BITTENCOURT, J. S. & NOVAS, F. E. 2010. The origin and early evolution of dinosaurs. *Biological Reviews*, **85**, 55–110.
- LE LOEUFF, J., MÉTAIS, E. ET AL. 2010. An Early Cretaceous vertebrate assemblage from the Cabao Formation of NW Libya. *Geological Magazine*, **147**, 750–759.
- LLOYD, G. T., DAVIS, K. E. ET AL. 2008. Dinosaurs and the Cretaceous terrestrial revolution. *Proceedings of the Royal Society, B*, **275**, 2483–2490, doi: 10.1098/rspb.2008.0715.
- LONG, R. A. & MURRY, P. A. 1995. Late Triassic (Carnian and Norian) tetrapods from the Southwestern United States. *New Mexico Museum of Natural History & Science Bulletin*, **4**, 1–254.
- MAIDMENT, S. C. R., NORMAN, D. B., BARRETT, P. M. & UPCHURCH, P. 2008. Systematics and phylogeny of Stegosauria (Dinosauria, Ornithischia). *Journal of Systematic Palaeontology*, **6**, 367–407.
- MANNION, P. D. 2009. Review and analysis of African sauropodomorph dinosaur diversity. *Palaeontologia Africana*, **44**, 108–111.
- MANNION, P. D. & UPCHURCH, P. 2010a. Completeness metrics and the quality of the sauropodomorph fossil record through geological and historical time. *Paleobiology*, **36**, 283–302, doi: 10.1666/090008.1
- MANNION, P. D. & UPCHURCH, P. 2010b. A quantitative analysis of environmental associations in sauropod dinosaurs. *Paleobiology*, **36**, 253–282, doi: 10.1666/08085.1
- MANNION, P. D. & UPCHURCH, P. 2011. A re-evaluation of the ‘mid-Cretaceous sauropod hiatus’, and the impact of uneven sampling of the fossil record on patterns of regional dinosaur extinction. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **299**, 529–540.
- MANNION, P. D., UPCHURCH, P., CARRANO, M. T. & BARRETT, P. M. 2011. Testing the effect of the rock record on diversity: a multidisciplinary approach to elucidating the generic richness of sauropodomorph dinosaurs through time. *Biological Reviews*, **86**, 157–181, doi: 10.1111/j.1469-185X.2010.00139.x
- MARTINEZ, R. N., SERENO, P. C., ALCOBER, O. A., COLOMBI, C. E., RENNE, P. R., MONTAÑEZ, I. P. & CURRIE, B. S. 2011. A basal dinosaur from the dawn of the dinosaur era in southwestern Pangaea. *Science*, **331**, 206–210.
- MARX, F. G. 2009. Marine mammals through time: when less is more in studying palaeodiversity. *Proceedings of the Royal Society, B*, **276**, 887–892.
- MARX, F. G. & UHEN, M. D. 2010. Climate, critters, and cetaceans: Cenozoic drivers of the evolution of modern whales. *Science*, **327**, 993–996.
- 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.

- McKINNEY, M. L. 1990. Classifying and analysing evolutionary trends. In: McNAMARA, K. J. (ed.) *Evolutionary Trends*. University of Arizona Press, Tucson, 28–58.
- MUNDILL, R., PALFY, J., RENNE, P. R. & BRACK, P. 2010. The Triassic timescale: new constraints and a review of geochronological data. In: LUCAS, S. G. (ed.) *The Triassic Timescale*. Geological Society, London, Special Publications, **334**, 41–60.
- NESBITT, S. J., IRMIS, R. B. & PARKER, W. G. 2007. A critical re-evaluation of the Late Triassic dinosaur taxa of North America. *Journal of Systematic Palaeontology*, **5**, 209–243.
- NORELL, M. A. & NOVACEK, M. J. 1992a. Congruence between superpositional and phylogenetic patterns comparing cladistic patterns with fossil evidence. *Cladistics*, **8**, 319–337.
- NORELL, M. A. & NOVACEK, M. J. 1992b. The fossil record and evolution: comparing cladistic and paleontologic evidence for vertebrate history. *Science*, **255**, 1690–1693.
- OLSEN, P. E., KENT, D. V. ET AL. 2002. Ascent of dinosaurs linked to an iridium anomaly at the Triassic–Jurassic boundary. *Science*, **296**, 1305–1307.
- ORCUTT, J., SAHNEY, S. & LLOYD, G. T. 2007. Tetrapod extinction across the Jurassic–Cretaceous boundary. *Journal of Vertebrate Paleontology*, **27** (Suppl. 3), 126A.
- PEARSON, D. A., SCHAEFER, T., JOHNSON, K. R., NICHOLS, D. J. & HUNTER, J. P. 2002. Vertebrate biostratigraphy of the Hell Creek Formation in southwestern North Dakota and northwestern South Dakota. *Geological Society of America Special Paper*, **361**, 145–167.
- PETERS, S. E. 2005. Geologic constraints on the macro-evolutionary history of marine animals. *Proceedings of the National Academy of Science, USA*, **102**, 12326–12331, doi: 10.1073/pnas.0502616102
- PETERS, S. E. 2006. Genus extinction, origination, and the durations of sedimentary hiatuses. *Paleobiology*, **32**, 387–407, doi: 10.1666/05081.1
- PETERS, S. E. & FOOTE, M. 2001. Biodiversity in the Phanerozoic: a reinterpretation. *Paleobiology*, **27**, 583–601, doi: [http://dx.doi.org/10.1666/0094-8373\(2001\)027<0583:BITPAR>2.0.CO;2](http://dx.doi.org/10.1666/0094-8373(2001)027<0583:BITPAR>2.0.CO;2)
- PETERS, S. E. & HEIM, N. A. 2010. The geological completeness of paleontological sampling in North America. *Paleobiology*, **36**, 61–79.
- 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.
- RAUP, D. M. & SEPKOSKI, J. J., JR. 1986. Periodic extinction of families and genera. *Science*, **231**, 833–836.
- RICE, W. R. 1989. Analyzing tables of statistical tests. *Evolution*, **43**, 223–225.
- SAHNEY, S., BENTON, M. J. & FERRY, P. A. 2010. Links between global taxonomic diversity, ecological diversity and the expansion of vertebrates on land. *Biology Letters*, **6**, 544–547, doi: 10.1098/rsbl.2009.1024
- SARJEANT, W. A. S. & CURRIE, P. J. 2001. The Great Extinction that never happened: the demise of the dinosaurs considered. *Canadian Journal of Earth Sciences*, **38**, 239–247.
- SEPKOSKI, J. J., JR. 1976. Species diversity in the Phanerozoic: species-area effects. *Paleobiology*, **2**, 298–303.
- SERENO, P. C. 1997. The origin and evolution of dinosaurs. *Annual Review of Earth and Planetary Science*, **25**, 435–489, doi: 10.1146/annurev.earth.25.1.435
- SERENO, P. C. 1999. The evolution of dinosaurs. *Science*, **284**, 2137–2147.
- SHEEHAN, P. M., FASTOVSKY, P. M., HOFFMANN, R. G., BERGHAUS, C. B. & GABRIEL, D. L. 1991. Sudden extinction of the dinosaurs: Latest Cretaceous, Upper Great Plains, U.S.A. *Science*, **254**, 835–839.
- SLOAN, R. E., RIGBY, J. K., JR., VAN VALEN, L. M. & GABRIEL, D. L. 1986. Gradual dinosaur extinction and simultaneous ungulate radiation in the Hell Creek Formation. *Science*, **234**, 1173–1175.
- SMITH, A. B. 2001. Large-scale heterogeneity of the fossil record: implications for Phanerozoic biodiversity studies. *Philosophical Transactions of the Royal Society of London, B Biological Sciences*, **356**, 351–367, doi: 10.1098/rstb.2000.0768
- 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: [http://dx.doi.org/10.1666/0094-8373\(2001\)027<0241:SLCARR>2.0.CO;2](http://dx.doi.org/10.1666/0094-8373(2001)027<0241:SLCARR>2.0.CO;2)
- SULLIVAN, R. M. 2006. The shape of Mesozoic dinosaur richness: a reassessment. *New Mexico Museum of Natural History and Science Bulletin*, **35**, 403–405.
- TANNER, L. H., LUCAS, S. G. & CHAPMAN, M. G. 2004. Assessing the record and causes of Late Triassic extinctions. *Earth Science Reviews*, **65**, 103–139.
- UHEN, M. D. & PYENSON, N. D. 2007. Diversity estimates, biases, and historiographic effects: resolving cetacean diversity in the Tertiary. *Palaeontologia Electronica*, **10** (2).
- UPCHURCH, P. & BARRETT, P. M. 2005. A taxic and phylogenetic perspective on sauropod diversity. In: CURRY ROGERS, K. & WILSON, J. A. (eds) *The Sauropods: Evolution and Paleobiology*. University of California Press, Berkeley, 104–124.
- UPCHURCH, P., HUNN, C. A. & NORMAN, D. B. 2002. An analysis of dinosaurian biogeography: evidence for the existence of vicariance and dispersal patterns caused by geological events. *Proceedings of the Royal Society, B*, **269**, 613–622.
- UPCHURCH, P., BARRETT, P. M. & DODSON, P. 2004. Sauropoda. In: WEISHAMPEL, D. B., DODSON, P. & OSMOLSKA, H. (eds) *The Dinosauria II*. University of California Press, Berkeley and Los Angeles, 259–322.
- WAITE, S. 2000. *Statistical Ecology in Practice: A Guide to Analyzing Environmental and Ecological Field Data*. Pearson Education Limited, Harlow.
- 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

- WANG, S. C. & DODSON, P. 2006. Estimating the diversity of dinosaurs. *Proceedings of the National Academy of Science, USA*, **103**, 13601–13605, doi: 10.1073/pnas.0606028103
- WEISHAMPEL, D. B. & JIANU, C.-M. 2000. Planteaters and ghost lineages: dinosaurian herbivory revisited. In: SUES, H.-D. (ed.) *The Evolution of Herbivory in Terrestrial Vertebrates. Perspectives from the Fossil Record*. Cambridge University Press, Cambridge, 123–143.
- WEISHAMPEL, D. B., BARRETT, P. M. *ET AL.* 2004. Dinosaur distribution. In: WEISHAMPEL, D. B., DODSON, P. & OSMOLSKA, H. (eds) *The Dinosauria II*. University of California Press, Berkeley and Los Angeles, 517–606.
- WILSON, J. A. 2005. Integrating ichnofossil and body fossil records to estimate locomotor posture and spatiotemporal distribution of early sauropod dinosaurs: a stratocladistic approach. *Paleobiology*, **31**, 400–423.
- WILSON, J. A. & UPCHURCH, P. 2009. Redescription and reassessment of the phylogenetic affinities of *Euhelopus zdanskyi* (Dinosauria: Sauropoda) from the Early Cretaceous of China. *Journal of Systematic Palaeontology*, **7**, 199–239.
- XU, X., CLARK, J. M. *ET AL.* 2006. A basal tyrannosauroid dinosaur from the Late Jurassic of China. *Nature*, **439**, 715–718.