CrossMark

# Mesozoic mammaliaform diversity: The effect of sampling corrections on reconstructions of evolutionary dynamics

Elis Newham [a],*, Roger Benson [b], Paul Upchurch [c], Anjali Goswami [d,e]

[a] Department of Earth Sciences, Wills Memorial Building, Queens Road, Bristol BS8 1RJ, United Kingdom
[b] Department of Earth Sciences, South Parks Road, Oxford OX1 3AN, United Kingdom
[c] Department of Earth Sciences, University College London, Gower Street, London WC1E 6BT, United Kingdom
[d] Department of Genetics, Evolution and Environment, University College London, Darwin Building, Gower Street, London WC1E 6BT, United Kingdom
[e] Department of Earth Sciences, University College London, Darwin Building, Gower Street, London WC1E 6BT, United Kingdom

## ARTICLE INFO

## ABSTRACT

Recent years have witnessed an explosion of new fossil discoveries and analyses documenting the unappreciated ecological and morphological diversity of Mesozoic Mammaliaformes. In contrast, the taxonomic diversity dynamics through the first 165 million years of mammal evolution have not yet been rigorously analysed, leaving patterns of diversification during this important period open to conjecture. Here, we present a comprehensive statistical analysis of global mammaliaform diversity spanning from the Late Triassic appearance of mammaliaforms (~230 million years ago [hereafter, mya]) to the end Cretaceous mass extinction (66 mya). We analysed 691 occurrences representing 367 genera and 550 species in standard time bins of approximately 10 million years in duration. Significant correlations between diversity and sampling proxies suggest sampling biases in the mammaliaform fossil record. Shareholder quorum subsampling and model-based approaches were used to mediate these biases. After applying these methods, the following patterns were supported: low standing diversity during the Late Triassic–Early Jurassic evolution of early Mammaliaformes (e.g., morganucodonts) was followed by high standing diversity during the Late Jurassic due primarily to the diversification of Eutriconodonta, Multituberculata and Cladotheria. This peak was followed by a fall in diversity during the middle of the Hauterivian–Barremian interval, suggesting that extinctions typically associated with the Jurassic–Cretaceous boundary may instead have occurred later, during the Early Cretaceous. Standing diversity recovered through several fluctuations during the 'mid' Cretaceous (approximately Barremian to Albian), leading to a second peak in the Campanian that reflects the diversification of key clades, including therians. Analyses of geographically restricted datasets illustrate a significant spatial heterogeneity in sampling, with several intervals dominated by North American occurrences. Uneven sampling effort and geographic heterogeneities in the fossil record are significant factors affecting reconstructions of Mesozoic mammaliaform diversity, and correcting these biases can markedly alter observed patterns and their interpretation.

## 1. Introduction

For decades, the focus of synapsid palaeontology has been on the therapsid response and recovery to the end Permian mass extinction, and the mammalian radiation following the K/Pg mass extinction, which heralded the "Age of Mammals". However, the fossil record of Mesozoic mammaliaforms spans ~2.5 times the duration of the comparatively well-studied record of Cenozoic mammals. Within Mammaliaformes, 11 major clades or functional grades (following Kielan-Jaworowska et al., 2004, and detailed below) formed an ecologically diverse Mesozoic assemblage from the Carnian (Late Triassic, ca 235–229 mya) onwards (Luo, 2007a). Recent fossil discoveries highlight a complex evolutionary history for Mesozoic Mammaliaformes (e.g. Luo et al., 2011), with the traditional scenario of a linear acquisition of mammalian characters being challenged by multiple evolutionary origins of key morphological features such as the tribosphenic molar (Luo et al., 2001) and middle ear ossicles (Luo et al., 2011). Moreover, in contrast to common depictions of early mammaliaforms as small terrestrial and scansorial insectivores, new fossils demonstrate that Mesozoic mammaliaforms invaded a variety of ecological niches, from semi-aquatic to gliding forms, and even dog-sized forms that preyed on juvenile dinosaurs (Hu et al., 2005; Luo and Wible, 2005; Ji et al., 2006; Meng et al.,

* Corresponding author. Tel.: +44 7914930434.
  E-mail address: elis.newham@googlemail.com (E. Newham).

**Table 1**
Mesozoic stages represented by ~10 myr intervals.

| Interval | Stages represented | Duration (myr) |
| --- | --- | --- |
| Triassic 3 | Carnian | 237.0–228.4 |
| Triassic 4 | Norian–Rhaetian | 228.4–201.3 |
| Jurassic 1 | Hettangian–Sinemurian | 201.3–190.8 |
| Jurassic 2 | Pliensbachian | 190.8–182.7 |
| Jurassic 3 | Toarcian–Aalenian | 182.7–170.3 |
| Jurassic 4 | Bajocian–Bathonian | 170.3–166.1 |
| Jurassic 5 | Callovian–Oxfordian | 166.1–157.3 |
| Jurassic 6 | Kimmeridgian–Tithonian | 157.3–145.0 |
| Cretaceous 1 | Berriasian–Valanginian | 145.0–133.9 |
| Cretaceous 2 | Hauterivian–Barremian | 133.9–126.3 |
| Cretaceous 3 | Aptian–Albian | 126.3–100.5 |
| Cretaceous 4 | Cenomanian | 100.5–93.9 |
| Cretaceous 5 | Turonian–Santonian | 93.9–83.6 |
| Cretaceous 6 | Campanian | 83.6–72.1 |
| Cretaceous 7 | Maastrichtian | 72.1–66.0 |

2006; Luo, 2007a). Despite the great attention paid in recent years to this previously unappreciated morphological and ecological diversity of early mammaliaforms, and a series of recent quantitative studies of taxonomic diversity in more basal synapsids (Brocklehurst and Fröbisch, 2014; Brocklehurst et al., 2013; Fröbisch, 2013), there has been little rigorous analysis of mammaliaform diversity dynamics prior to the K/Pg mass extinction (Rose, 2006). Previous approaches have been either broad and qualitative assessments of subclades (Luo, 2007b) or geographically restricted to the North American record (Alroy, 2009) and more specific localities (Wilson, 2005, 2013).

Reconstructing diversity dynamics over deep time is a core theme of palaeobiology (Jablonski, 1999; Raup, 1972; Valentine, 1985). Although the potential effects of geological and anthropogenic biases on accurate taxon counts have been discussed for decades (Raup et al., 1975), it is only more recently that substantial efforts have been made to correct these biases (Alroy, 2000, 2008, 2010; Alroy et al., 2001, 2008; Behrensmeyer et al., 2005; Peters and Foote, 2001; Smith and McGowan, 2007; Smith et al., 2012). A growing number of studies have focussed in particular on biases introduced by differences in outcropping rock area (Crampton et al., 2003; Smith and McGowan, 2007), preservation potential of fossil organisms (Smith, 2001), or evenness and fairness of sampling during standard intervals (Alroy, 2010; Alroy et al., 2001, 2008). These studies suggest that many features of observed diversity curves could be artefacts of changes in fossil preservation, geological sampling, or anthropogenic sampling rather than true biotic signals (e.g. Smith, 2007; Smith et al., 2012). Complex Earth system interactions such as sea level change may drive both sedimentation and ancient biodiversity in the marine realm, suggesting that covariation of fossil taxon counts and potentially biasing factors is not always causal (Peters, 2005; Benson and Butler, 2011; Hannisdal and Peters, 2011; but see Smith and Benson, 2013). However, terrestrial processes may be simpler, with factors such as rock area and collection effort directly biasing taxon counts (e.g. Benson and Upchurch, 2013; Benson et al., 2013; Butler et al., 2011a,b; Upchurch et al., 2011). Here, we present the first quantitative investigation of the global taxonomic palaeodiversity of Mesozoic Mammaliaformes, applying robust sampling-correction approaches to account for geological and anthropogenic biases and reassessing diversity dynamics in early mammal evolution.

## 2. Methods and materials

### 2.1. Datasets

#### 2.1.1. Mammaliaform taxa

We have attempted to maximise coverage of Mesozoic mammaliaform occurrence data in the Palaeobiology Database (Alroy et al., 1998), with an extensive literature review and comparison with data in Kielan-Jaworowska et al. (2004). Mammaliaformes was considered as a monophyletic clade, consisting of all descendants of the most recent common ancestor of Morganucodonta and crown Mammalia (Luo et al., 2002; Rowe, 1988; Zhou et al., 2013). Morganucodonts, docodonts and kuehneotherids are successively more closely related outgroups of crown Mammalia, consistent with current phylogenetic studies and nomenclature. The 'crown' Mammalia is defined as the common ancestor of extant monotremes and extant therians and all fossil groups cladistically nested within this clade. Mammalia comprises eutriconodonts, australosphenidans, multituberculates, spalacotheroids, stem cladotherians, stem boreosphenidans, eutherians and metatherians (Kielan-Jaworowska et al., 2004). Stem allotherians, formerly considered to be related to multituberculates, were recently removed from the crown Mammalia (Zhou et al., 2013).

The updated dataset was downloaded from the Paleobiology Database (paleobiodb.org) on 20/02/2013. We then removed ichnofossils and collections with poorly constrained stratigraphic ages (spanning more than two time bins). This filtered version of the dataset comprises 691 occurrences of 367 genera and 550 species, ranging from the Carnian (237–227 mya; Late Triassic) to the Maastrichtian (72.1–66.0 mya; Late Cretaceous). In order to determine the effects of a geographically heterogeneous fossil record on global diversity patterns, two regional datasets were also created via further filtering of the global dataset. These are: (1) a North American dataset consisting of occurrences from the USA, Canada, Mexico and Greenland (288 occurrences of 208 genera and 302 species); and (2) an Asian dataset comprising occurrences from the Russian Federation, China, Japan, Kazakhstan, Kyrgyzstan, Mongolia, Tajikistan and Uzbekistan (255 occurrences of 137 genera and 191 species). These datasets are all provided in the electronic online supplement, along with other data and scripts used to perform our analyses.

#### 2.1.2. Time bins

Occurrences were assigned to standardised time bins, approximately 10 myr (million years) in duration (Table 1) (Alroy, 2000; Alroy et al., 2008). Whereas previous authors have included only Tithonian occurrences in the 'Jurassic 6' time bin, we include both Kimmeridgian and Tithonian occurrences, as 60% of unique occurrences within these stages spanned both intervals (i.e. were defined as 'Kimmeridgian–Tithonian'). The same approach was applied to the 'Cretaceous 3' interval, which has been previously assigned Aptian occurrences only: here it includes both Aptian and Albian occurrences because 24% of occurrences within these stages spanned both intervals.

### 2.2. Analytical approach

Recent developments in the quantification of vertebrate palaeodiversity curves have increased confidence in reconstructions of ancient diversity patterns by attempting to correct observed taxon counts for sampling biases (e.g. Alroy, 2000; Barrett et al., 2009; Fröbisch, 2008; Irmis and Whiteside, 2011; Lloyd, 2012; Lloyd et al., 2008; Mannion et al., 2011). These methods fall broadly into two classes: (1) model-based approaches that first attempt to quantify the available record using sampling proxies, then construct models of expected diversity to use as a correction factor for observed diversity (Lloyd, 2012; Peters and Foote, 2001; Smith and McGowan, 2007); and (2) sampling standardisation (or subsampling) approaches that simulate an equal, or fair sample of ancient diversity among time intervals based directly on occurrence data, making fewer assumptions about drivers of the record (Alroy, 2010; Alroy et al., 2001, 2008; Kowalewski, 2002). We used both methods to assess the temporal pattern of mammaliaform diversity through the Mesozoic at both the species and genus levels. The application of bias correction approaches that treat the data very differently highlights the most robust temporal trends when their results are congruent (Mannion et al., 2011, 2012; Smith et al., 2012).
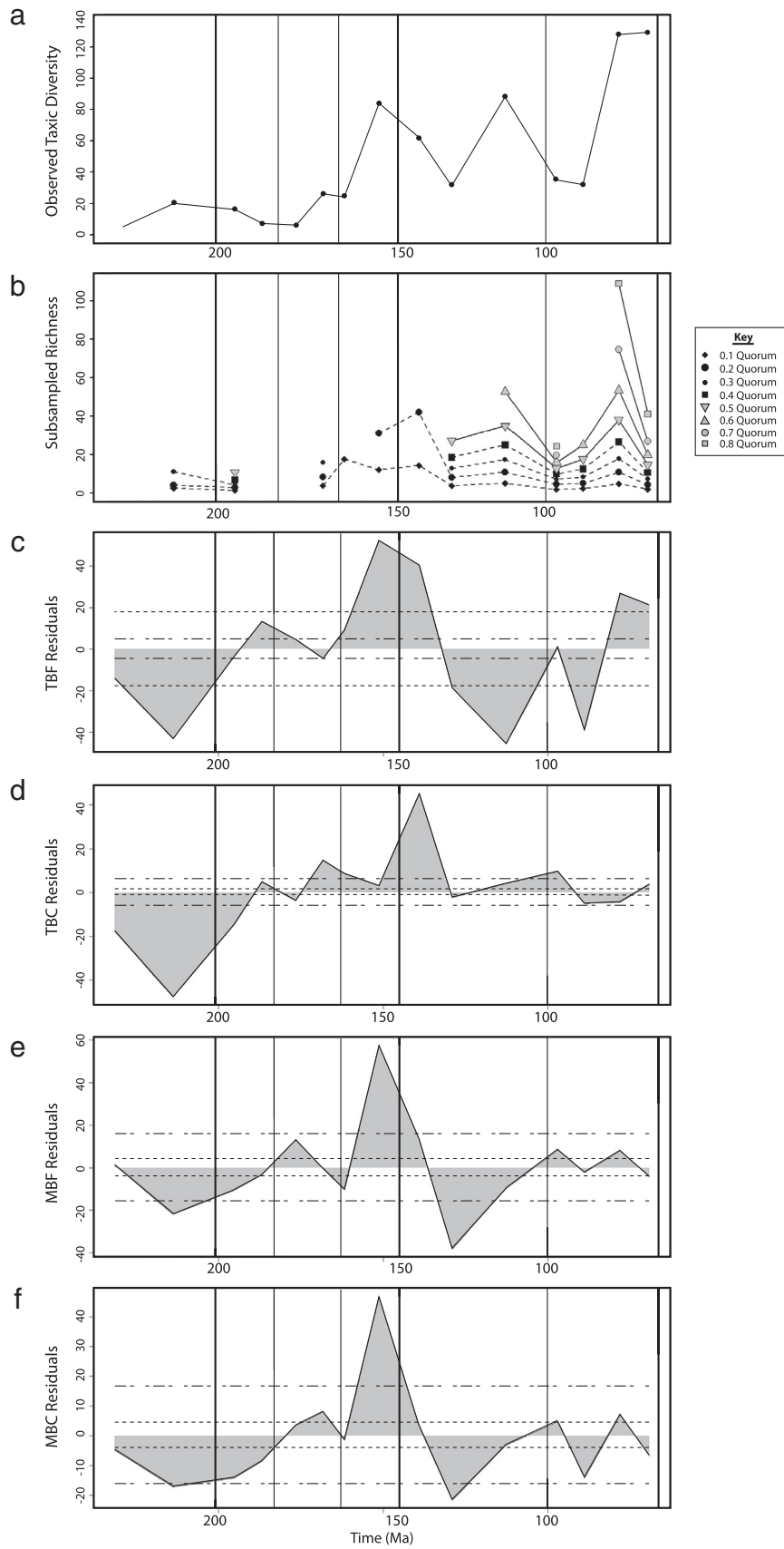
**Table 2**
Correlations between un-transformed time series. STDE: species taxonomic diversity estimate; GTDE: generic taxonomic diversity estimate; TBFs: tetrapod-bearing formations; TBCs: tetrapod-bearing collections; MBFs: mammaliaform-bearing formations; MBCs: mammaliaform-bearing collections.

| Correlation | Spearman's $r$ | p-Value | Adjusted p-value | Kendall's $\tau$ | p-Value | Adjusted p-value significance threshold |
|---|---|---|---|---|---|---|
| MBF/MBC | 0.94 | <0.001 | 3.60E−03 | 0.87 | <0.001 | 0.003571429 |
| MBC/GTDE | 0.9 | <0.001 | 7.20E−03 | 0.77 | <0.001 | 0.007142857 |
| TBF/TBC | 0.7 | <0.001 | 0.01 | 0.71 | <0.001 | 0.017857143 |
| TBC/GTDE | 0.75 | <0.001 | 0.01 | 0.6 | 2.05E−03 | 0.042857143 |
| TBF/MBC | 0.85 | <0.001 | 0.02 | 0.68 | <0.001 | 0.025 |
| TBC/MBC | 0.82 | <0.001 | 0.02 | 0.66 | <0.001 | 0.028571429 |
| MBF/GTDE | 0.81 | <0.001 | 0.03 | 0.65 | <0.001 | 0.035714286 |
| TBF/MBF | 0.8 | <0.001 | 0.03 | 0.81 | <0.001 | 0.010714286 |
| TBF/GTDE | 0.8 | <0.001 | 0.03 | 0.61 | 1.46E−03 | 0.039285714 |
| MBF/STDE | 0.79 | <0.001 | 0.04 | 0.65 | <0.001 | 0.035714286 |
| TBF/STDE | 0.79 | <0.001 | 0.04 | 0.58 | 2.54E−03 | 0.046428571 |
| TBC/STDE | 0.75 | 1.37E−03 | 0.04 | 0.5933 | 2.05E−03 | 0.042857143 |
| TBC/MBF | 0.73 | 0.0019095 | 0.05 | 0.74923 | <0.001 | 0.032142857 |
| TBC/STDE | 0.73 | 1.91E−03 | 0.05 | 0.5619 | 3.50E−03 | 0.05 |

### 2.2.1. Model-based correction of uneven sampling

*2.2.1.1. Residuals method.* For the model-based approach, we used several sampling proxies, including counts of fossil-bearing formations (a measure of rock record bias) and fossil collections (a measure of anthropogenic collecting bias). These have previously been shown to correlate with taxon counts of Mesozoic terrestrial vertebrates in a likely causal relationship (Barrett et al., 2009; Butler et al., 2011a, b; Upchurch et al., 2011; Benson et al., 2013; but see Benton et al., 2011, 2013 for a contrary opinion). We use 'narrower' sampling proxies based on the occurrences of all mammaliform specimens, and 'wider' sampling proxies based on the occurrences of all terrestrial tetrapod specimens (See Upchurch and Barrett, 2005 and Benton et al., 2011, 2013 for discussions of the construction of sampling proxies and the issue of 'signal redundancy'). In total, 240 mammaliaform-bearing collections (MBCs), 106 mammaliform-bearing formations (MBFs), 4290 terrestrial tetrapod-bearing collections (TBCs) and 599 terrestrial tetrapod-bearing formations (TBFs) were binned according to ~10 myr intervals as described above for taxon occurrences. Data for TBCs and TBFs were obtained from the *Paleobiology Database* (see above) and, along with MBC and MBF data, are presented in the electronic online supplement.

The 'residuals method' of Smith and McGowan (2007) generates a model in which a fossil record sampling proxy is used to estimate the amount of taxonomic diversity expected ('modelled diversity estimate'; MDE) for that amount of sampling if diversity was equal in all time intervals. All values of sampling and diversity were $log_{10}$-transformed prior to analysis, with a start of 1.0. Next, the transformed values for each data series were independently rank-ordered from lowest to highest. A linear ordinary least squares regression line was then calculated between the ranked sampling proxy and ranked TDE (observed taxonomic diversity) data, representing a relationship in which sampling predicts diversity. This equation is then used to calculate an MDE for each interval. The MDEs are then subtracted from the TDEs of each interval, leaving a residual diversity estimate (RDE). Thus, the residual diversity estimate reflects the amount of observed in-bin diversity unaccounted for by variation in fossil record sampling. Lloyd (2012) refined this method by first assessing which model (linear, logarithmic, exponential, hyperbolic, sigmoidal or polynomial) provides the best fit between ranked sampling proxy and ranked TDE using the sample

size-corrected Akaike Information Criterion (Johnson and Omland, 2004). This method also calculates 95% confidence limits based on 1.96 times the standard error values of the model line fit. This provides more appropriate error estimations than previous residual methods (e.g. those presented by Barrett et al., 2009) because it more effectively reflects significant deviations from the sampling-driven model (Lloyd, 2012). Here, models were created using the R script available from Graeme Lloyd's home page (http://www.graemetlloyd.com). The relationships between each of the four sampling proxies (MBC, MBF, TBC and TBF) and TDE were examined using this approach. We found that a linear relationship between each sampling proxy and TDE was best supported. Thus, all analyses and results described below are based on these linear models.

*2.2.1.2. Correlation tests.* In order to determine the extent to which a sampling proxy is responsible for driving observed TDE, multiple pairwise correlation tests were conducted between the $log_{10}$-transformed data series using Pearson's product moment ($r$) and the non-parametric Spearman's $\rho$ and Kendall's $\tau$ in PAST (Hammer et al., 2001). These tests were performed on both raw data and data transformed by generalised differencing (explained below) using a significance cut-off of $p < 0.05$. We evaluated statistical support for correlations based on both original p-values, and p-values adjusted for the implementation of multiple tests. The latter was carried out using the 'false discovery rate' procedure of Benjamini and Hochberg (1995).

Generalised differencing incorporates detrending and differencing to attempt to correct for temporal autocorrelation within a dataset (i.e. the effects of the statistical overprinting of an interval's diversity onto the interval proceeding it; McKinney, 1990; Alroy, 2000; Benson and Butler, 2011). Autocorrelation tends to result in spurious detection of correlation between data series and thus must be removed prior to statistical comparisons. First, the presence of long-term trend is assessed by $log_{10}$-transforming the data for each stage, and plotting this against the stage midpoint age (in mya). If the regression line provides a significant fit to the data ($p < 0.05$), residuals from the regression represent a statistically detrended data series. Next, autocorrelation is evaluated by regressing each point in the detrended time series against that of the preceding interval. If the regression line is non-significant ($p > 0.05$) then no statistical autocorrelation

**Fig. 1.** Global species level data series through the Mesozoic. (a) Observed mammaliaform species taxonomic diversity estimate (STDE) through the Mesozoic. (b) SQS subsampled taxonomic richness. (c–f) Residual time series; (c) terrestrial tetrapod-bearing formations (TBFs) vs STDE residuals. (d) Terrestrial tetrapod-bearing collections (TBCs) vs STDE residuals. (e) Mammaliaform-bearing formations (MBF) vs STDE residuals. (f) Mammaliform-bearing collections (MBCs) vs STDE residuals. Vertical lines represent the Triassic/Jurassic boundary, Early/Mid Jurassic boundary, Mid/Late Jurassic boundary, Jurassic/Cretaceous boundary, Early/Mid Cretaceous boundary respectively. Dashed lines mark standard error 95% confidence intervals. Dash-dot lines mark standard deviation 95% confidence intervals.

**Table 3**
Correlations between un-transformed time series and ~10 myr interval mid-point ages. STDE: species taxonomic diversity estimate; GTDE: generic taxonomic diversity estimate; TBFs: tetrapod-bearing formations; TBCs: tetrapod-bearing collections; MBFs: mammaliaform-bearing formations; MBCs: mammaliaform-bearing collections.

|  | TBF | TBC | MBF | MBC | STDE | GTDE |
|---|---|---|---|---|---|---|
| $r^2$ | 0.38 | 0.26 | 0.5 | 0.54 | 0.64 | 0.66 |
| p-Value | 0.02 | 0.05 | 3.00E−03 | 1.79E−03 | <0.001 | <0.001 |
| Adjusted p-value significance threshold | 0.04 | 0.05 | 0.03 | 0.03 | 0.02 | 0.01 |

**Table 4**
The significance of autocorrelation in each time series. STDE: species taxonomic diversity estimate; GTDE: generic taxonomic diversity estimate; TBFs: tetrapod-bearing formations; TBCs: tetrapod-bearing collections; MBFs: mammaliaform-bearing formations; MBCs: mammaliaform-bearing collections.

|  | TBF | TBC | MBF | MBC | STDE | GTDE |
|---|---|---|---|---|---|---|
| $r^2$ | 0.05 | 0.02 | 0.02 | 0.05 | 0.01 | 3.30E−03 |
| p | 0.46 | 0.6 | 0.59 | 0.43 | 0.76 | 0.85 |
| Adjusted p-value significance threshold | 0.01 | 0.02 | 0.025 | 0.03 | 0.04 | 0.05 |

is present and no further action is needed prior to analysis. If autocorrelation is present then, generalised differencing is implemented using the equation:

$$x_{gd} = x_t - ax_{t-1}$$

where a is the slope of the regression line (autocorrelation coefficient), $x_{gd}$ is the differences data series, $x_t$ is the original (or detrended) data series, and $x_{t-1}$ is the data series at a lag of 1.

### 2.2.2. Subsampling approaches to the correction of uneven sampling

We used a subsampling approach, Shareholder Quorum Subsampling (SQS) (Alroy, 2010) to assess statistical unevenness in the coverage of observed diversity and provide a subsampled taxonomic richness for each interval at both the generic and species levels. SQS is designed to simulate a fair sample of original diversity by subsampling to a pre-specified frequency of statistical coverage, as opposed to traditional rarefaction, which corrects for uneven sampling of intervals by selecting the same number of samples from each interval. Alroy (2010) argued that the uniform subsampling provided by rarefaction can overcompensate for bias by concentrating sampling in intervals with the poorest coverage, which serves to decrease the fidelity of genuine diversity fluctuations at lower sampling levels. Alroy (2009, 2010) thus developed the SQS method to minimise this overcompensation in

time-series by viewing taxa as 'shareholders' whose frequency becomes their 'share' within a time interval. SQS concentrates sampling in well-represented intervals, by using Good's (1953) u coverage estimator to estimate the relative proportions of 'true' diversity represented in each interval. By randomly drawing fossil collections within a time bin until taxa that have been sampled at least once have summed a total of shares (sum frequencies) that meets a pre-determined quorum, SQS bases subsampled richness on relative frequencies of taxa rather than counts (Alroy, 2009, 2010). Because each interval is represented by the same proportion of taxa, the amount of taxa forming this 'quorum' is allowed to change between intervals relative to their coverage (Alroy, 2009, 2010).

Here, we utilised the R script version of SQS available from John Alroy's home page (http://bio.mq.edu.au/~jalroy/SQS.html) implemented in R version 2.14.1. Occurrences from collections spanning two time bins were assumed to occur in both bins, whereas collections spanning more than two bins were deleted prior to analysis. Then occurrences within specific collections were programmed to only be read once, to make sure only unique combinations of collection/taxa were included in the dataset and minimise spurious multiple records (Alroy, 2010).

## 3. Results

### 3.1. Observed Mesozoic mammaliaform taxonomic diversity

We first calculated observed in-bin species counts (species taxonomic diversity estimate; STDE) and generic counts (generic taxonomic diversity estimate; GTDE) (Table 1). The uncorrected mammaliaform fossil record displays an apparent long-term increase in diversity through the Mesozoic, punctuated by four peaks occurring in the Late Triassic (Triassic 4), Late Jurassic (Jurassic 6), early Late Cretaceous (Cretaceous 3) and Late Cretaceous (Cretaceous 6–7) (Fig. 1a). These peaks are separated by troughs during the Callovian–Oxfordian (Jurassic 5), the Early Cretaceous (Cretaceous 2) and Early–Late Cretaceous (Cretaceous 4–5).

### 3.2. Influence of sampling on mammaliaform diversity

All pairwise comparisons of the untransformed time series (STDE and GTDE against each of the four sampling proxies) were significantly positively correlated (p < 0.05) (Table 2). Regressions of all data series except TBC also produced significant correlations with interval mid-point age (Table 3), demonstrating the presence of trend. Therefore most of the data were detrended during generalised differencing. Following generalised differencing, none of the autocorrelation coefficients were statistically significant (Table 4), so further screening of the data series was not required. Following generalised differencing,

**Table 5**
Correlations between generalised differenced time series. STDE: species taxonomic diversity estimate; GTDE: generic taxonomic diversity estimate; TBFs: tetrapod-bearing formations; TBCs: tetrapod-bearing collections; MBFs: mammaliaform-bearing formations; MBCs: mammaliaform-bearing collections.

| Correlation | Spearman's r | p-Value | Adjusted p-value | Kendall's $\tau$ | p-Value | Adjusted p-value significance threshold |
|---|---|---|---|---|---|---|
| MBF/MBC | 0.94 | <0.001 | 3.60E−03 | 0.83 | <0.001 | 3.57E−03 |
| TBF/TBC | 0.77 | <0.001 | 7.10E−03 | 0.68 | <0.001 | 7.14E−03 |
| TBF/MBF | 0.76 | <0.001 | 0.01 | 0.58 | 2.54E−03 | 0.01 |
| MBC/STDE | 0.73 | 1.81E−03 | 0.02 | 0.58 | 2.54E−03 | 0.01 |
| MBC/GTDE | 0.73 | 2.23E−03 | 0.02 | 0.56 | 3.50E−03 | 0.02 |
| TBC/MBC | 0.75 | 1.39E−03 | 0.01 | 0.54 | 4.79E−03 | 0.02 |
| MBF/GTDE | 0.67 | 6.13E−03 | 0.03 | 0.54 | 4.79E−03 | 0.03 |
| TBF/MBC | 0.68 | 5.76E−03 | 0.03 | 0.52 | 6.49E−03 | 0.03 |
| TBC/MBF | 0.68 | 0.005764 | 0.03 | 0.52 | 6.49E−03 | 0.03 |
| TBF/MBC | 0.68 | 5.76E−03 | 0.03 | 0.52 | 6.49E−03 | 0.04 |
| TBC/STDE | 0.66 | 6.91E−03 | 0.04 | 0.51 | 8.72E−03 | 0.04 |
| TBC/GTDE | 0.62 | 1.41E−02 | 0.04 | 0.49 | 1.16E−02 | 0.04 |
| TBF/STDE | 0.5 | 5.77E−02 | 0.05 | 0.41 | 3.33E−02 | 0.05 |
| TBF/GTDE | 0.49 | 6.64E−02 | 0.05 | 0.39 | 4.25E−02 | 0.05 |
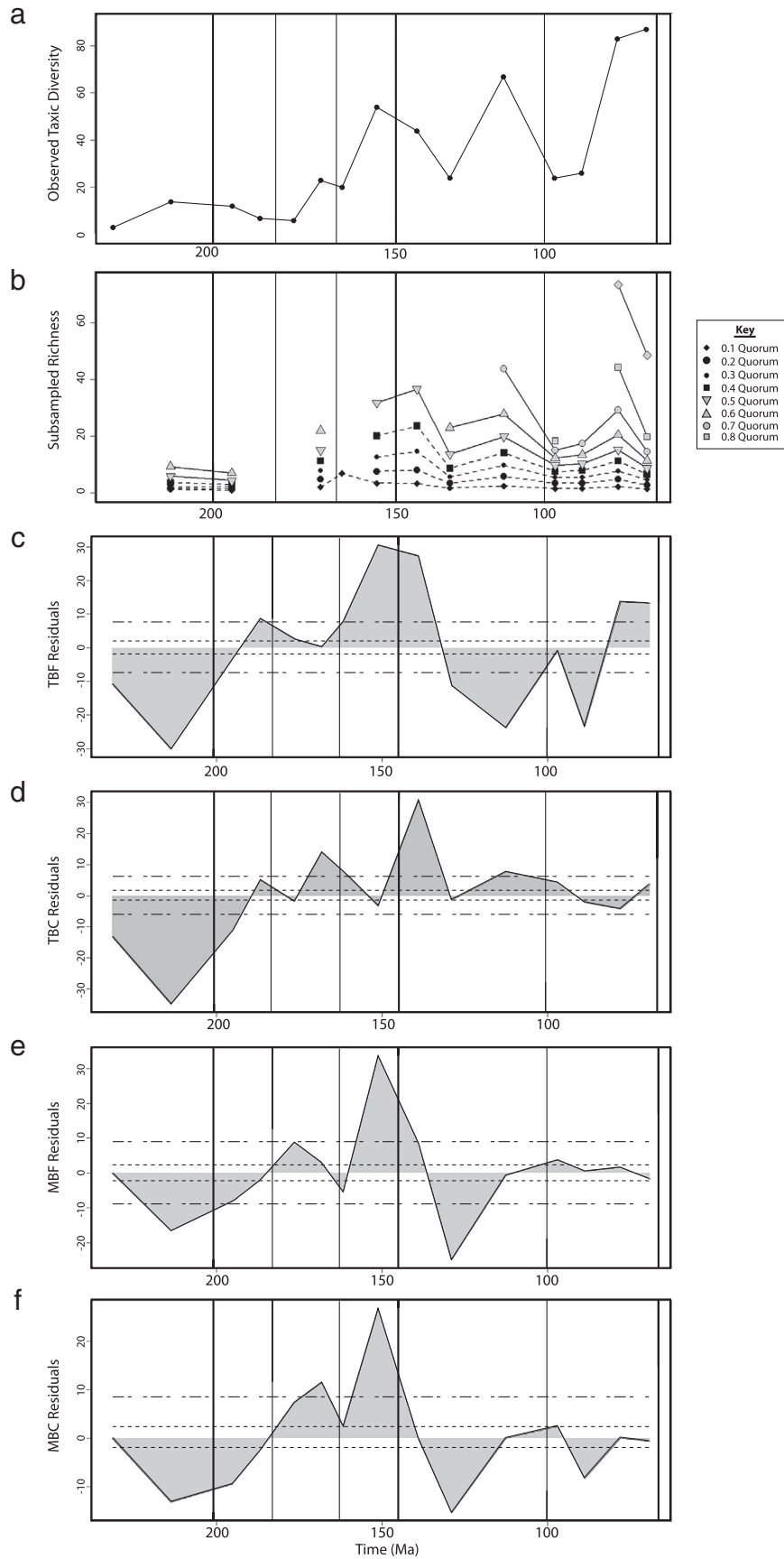
**Fig. 2.** Global generic data series. (a) Observed mammaliaform generic taxonomic diversity estimate (GTDE) (b) SQS subsampled taxonomic richness. (c–f) Residual time series; (c) terrestrial tetrapod-bearing formations (TBFs) vs GTDE residuals. (d) Terrestrial tetrapod-bearing collections (TBCs) vs GTDE residuals. (e) Mammaliaform-bearing formations (MBFs) vs GTDE residuals. (f) Mammaliaform-bearing collections (MBCs) vs GTDE residuals. Vertical lines represent the Triassic/Jurassic boundary, Early/Mid Jurassic boundary, Mid/Late Jurassic boundary, Jurassic/Cretaceous boundary, and Early/Mid Cretaceous boundary respectively. Dashed lines mark standard error 95% confidence intervals. Dash-dot lines mark standard deviation 95% confidence intervals.

pairwise comparisons between every data series except TBF/Species TDE and TBF/Genus TDE produced significant correlations (Table 5). Once false discovery rate was accounted for, all correlations were significant except for Spearman's *r* correlations between generalised differenced TBF/STDE and TBF/GTDE time series (Tables 2–5). Thus, observed mammaliaform species and genus-level diversities are largely controlled by fossil record sampling, necessitating application of corrections via the residuals and SQS methods outlined above.

### 3.3. Sampling-corrected mammaliaform diversity

#### 3.3.1. 'Long-term' (multi-interval) diversity patterns

Three robust signals in Mesozoic mammaliaform diversity were evident among most of our analyses.

*3.3.1.1. Low diversity at the Triassic–Jurassic boundary.* Global residual analyses and SQS both suggest generally low mammaliaform diversity during the Triassic–Jurassic boundary interval, with diversity increasing until the earliest Cretaceous (Figs. 1–2; although note that only one time bin in this interval achieves a quorum level of >0.3 at the species level). Residual analysis of the North American data suggested a similar pattern of mammaliaform diversity during this interval, but there were insufficient data to confirm this regional result using SQS, and there were also insufficient Asian data to draw reliable conclusions (note the zero or low quorum levels for this entire interval; Fig. 4).

*3.3.1.2. Low diversity in the early Early Cretaceous.* A substantial decrease in diversity between Cretaceous 1 (Berriasian–Valanginian) and Cretaceous 2 (Hauterivian–Barremian) was suggested by global residual analysis and SQS (Figs. 1–2; although note the low quorum values in species-level SQS). Regional datasets did not provide sufficient data to examine this at 10 myr resolution using SQS because of the scarcity of Berriasian–Barremian mammaliaform localities in North America (Sames et al., 2010). Nonetheless, the North American species-level data (SQS and residuals) and generic-level data (residuals only) suggests a decline in diversity between the Late Jurassic and late Early Cretaceous, although this was not confirmed by SQS of the generic-level data (Fig. 3). There were insufficient Asian data to draw reliable conclusions about this interval (note the zero or low quorum levels; Fig. 4).

*3.3.1.3. A peak in diversity in the Late Cretaceous.* Residual analyses did not consistently support any pattern following the decrease in Cretaceous 2 (Figs. 1c–f, 2c–f). A third peak in observed diversity during Cretaceous 6–7 in the raw data (Fig. 1a) was not supported as statistically significant in the residual analyses. SQS results were returned at substantial quorum levels (>0.5) for each of the intervals proceeding Cretaceous 6–7, suggesting low standing diversity through Cretaceous 4–5. However, a primarily North American assemblage dominated during Cretaceous 4 (78–85% of global subsampled richness), and a primarily Asian assemblage dominated during Cretaceous 5 (65–71% of global subsampled richness), so 'global' results may be more reflective of local diversity specific to these regions. High SQS subsampled diversity during Cretaceous 6–7 was returned at the most substantial quorum levels of the entire Mesozoic, indicating a genuine Late Cretaceous peak in diversity despite the lack of statistical support for this peak in the residual analyses. SQS results also indicated relatively lower diversity during Cretaceous 7 than in Cretaceous 6, compared to observed levels (Fig. 1b). Asian SQS analyses only produced results at a quorum of 0.2, and so made no significant contribution to global results at higher quorum levels (Fig. 4).

#### 3.3.2. 'Short-term' (interval-to-interval) diversity patterns

Interval-to-interval changes in diversity can offer a more detailed context to long-term trajectories, although caution must be taken when assessing results of data-rich intervals relative to surrounding stages with poorer data quality. The first peak in observed species/ generic TDE occurred during Triassic 4 (Fig. 1a), supported by high generic residuals (Figs. 1c–f, 2c–f). The Triassic 4 signal was not dominated by any particular geographic subsets of the data (Figs. 3, 4), and represents a collection of European/Asian occurrences. Analysis of the Asian data demonstrated one of the highest residual peaks of the entire Mesozoic through this interval, and yet the subset provides <50% of global richness, while North America did not contribute data during this interval. Jurassic 4 presented a second residual peak, and was shown to be another of the few high diversity intervals not to be dominated by a particular geographic subset of the data (Figs. 3, 4). The subsequent fall in diversity during Jurassic 5 observed in the residuals is not statistically significant and so cannot be confidently interpreted as a genuine biological pattern. Finally, residual results support high diversity during Cretaceous 3, bolstered by the reintroduction of North American occurrences (19% of total occurrences) (Figs. 1–3).

As in the residual results, SQS supports Triassic 4 and Jurassic 4 diversity peaks, and the Jurassic 5 fall in diversity is not represented at significant quorum levels (Figs. 1b, 2b). SQS analyses provided further support for a fall in diversity between Cretaceous 1 and Cretaceous 2, with results at significant quorum levels implying a genuine biotic turnover. SQS results at significant quorum levels also supported high diversity through Cretaceous 3 (Fig. 1).

## 4. Discussion

### 4.1. Sampling bias in the mammaliaform fossil record

The significant correlations between observed mammaliaform diversity and multiple fossil sampling proxies suggest that some patterns in apparent diversity may be explained by temporal variation in rock record quality rather than by evolutionary dynamics. As a result, caution is needed when interpreting uncorrected data. Both sets of analyses described here do support discrete series of statistically robust shifts in Mesozoic mammaliaform diversity. However, it is important to note that these results may also be affected by inconsistencies within the dataset itself that cannot be accounted for using our methods. For example, differences in the intensity of fossil collection within single formations or at particular localities may affect the results of our residual analyses. A possible example occurs in Jurassic 6 and Cretaceous 1, which contain the exceedingly well-sampled Morrison and Lulworth formations (respectively), and have apparently high residual diversity. These intervals also have high subsampled diversity, suggesting that the signal might be robust. However, we consider that this result should be interpreted with caution. Identifying points of consistency among the results of multiple methods of bias correction allows us to determine which features of diversity curves are well supported and to reassess previous hypotheses concerning the diversity dynamics of Mesozoic mammaliaforms.

### 4.2. Is mammaliaform evolution episodic?

Based on an appraisal of Mesozoic mammaliaform phylogeny and diversity, Luo (2007a) described Mesozoic mammalian evolution as successive diversifications of clades or grades whose acmes (in terms

**Fig. 3.** North American data series through the Mesozoic. (a) SQS subsampled mammaliaform species taxonomic richness. (b) Mammaliaform-bearing formations (MBFs) vs STDE residuals. (c) Mammaliaform-bearing collections (MBCs) vs STDE residuals. (d) SQS subsampled mammaliaform generic taxonomic richness. (e) MBF vs GTDE residuals. (f) MBC vs GTDE residuals. Vertical lines represent the Triassic/Jurassic boundary, Early/Mid Jurassic boundary, Mid/Late Jurassic boundary, Jurassic/Cretaceous boundary, and Early/Mid Cretaceous boundary respectively. Dashed lines mark standard error 95% confidence intervals. Dash-dot lines mark standard deviation 95% confidence intervals.
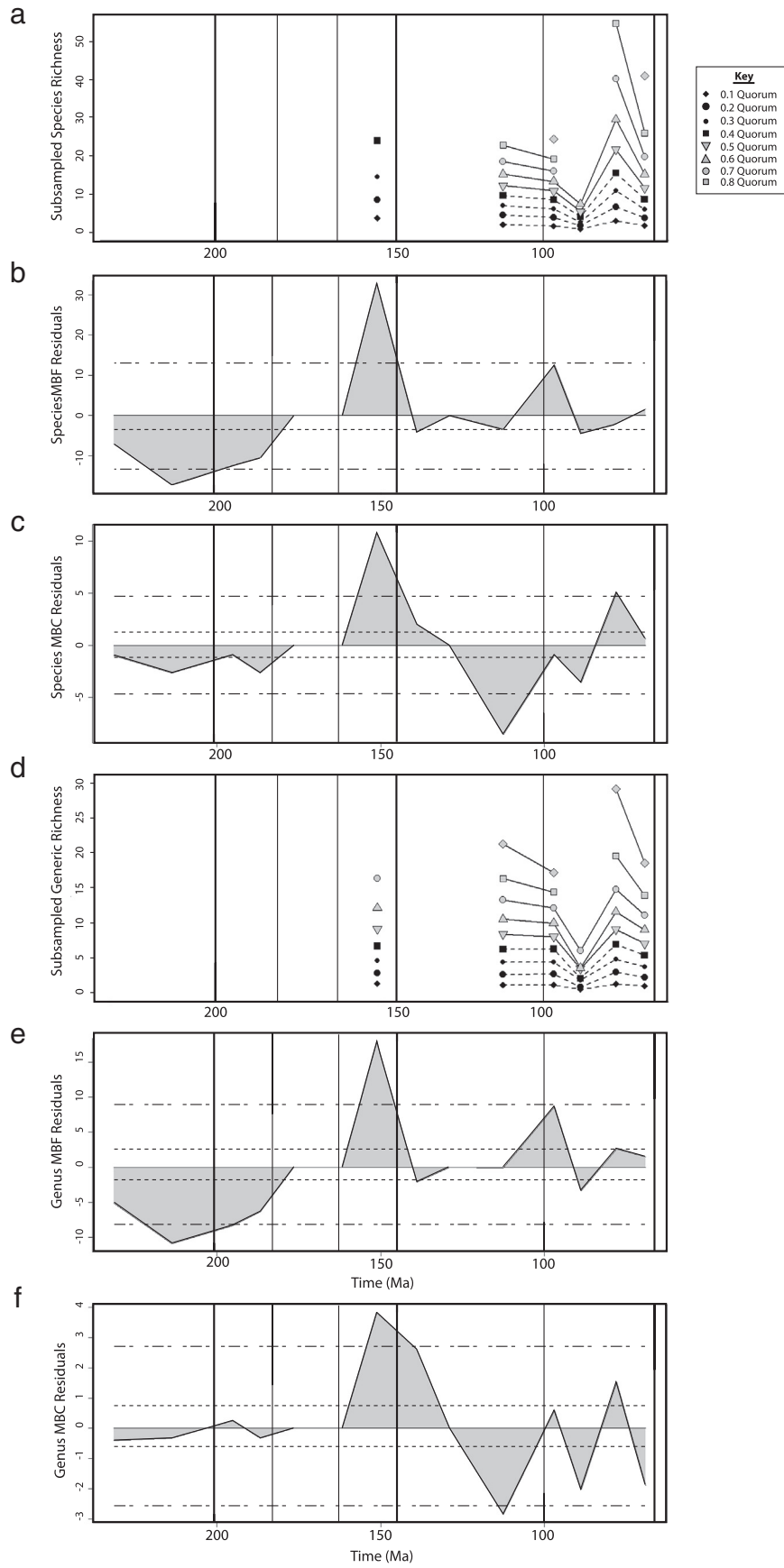
of abundance and diversity) were relatively short-lived. When mapped over geological time, this hypothesis appears as 'waves of diversification' (Luo, 2007a; Fig. 1). Luo (2007a) identified four principal diversification episodes; (1) the Late Triassic–Early Jurassic diversification of early Mammaliaformes; (2) the radiation of the Docodonta through the Middle Jurassic; (3) the diversification of the Eutriconodonta, Multituberculata, Cladotheria and the monotreme stem group (aukstribosphenidans) through the Late Jurassic; and (4) the diversification of tribosphenic stem-group therians during the 'mid' to Late Cretaceous. Luo's (2007a) qualitative appraisal of mammaliaform diversity approximately matches the patterns of richness seen in our data before correction for biases (Fig. 1). While our approach is essentially taxon-blind and thus does not account for the importance of new clade originations or morphological and ecological diversification during specific events or intervals, our analyses suggest that the high taxon counts of many intervals attributed to species diversification of Mammaliaformes (an evolutionary explanation) are more plausibly due to intervals of increased sampling by palaeontologists (a geological or anthropogenic explanation). Nevertheless, once sampling biases are accounted for, our data do robustly support some of the hypothesised episodes of mammaliaform diversification, albeit not as peaks of absolute diversity.

A key difference in the approach of Luo (2007a) and our corrected diversity estimates is the former's use of morphological diversity as a proxy for mammaliaform diversification through the Mesozoic; Luo's diversification events are based on the appearance of novel morphologies and new clades as well as changes in taxonomic diversity. The emergence of a new clade and the accompanying change in morphological diversity can be heavily influenced by some well-preserved taxa at a given geological time interval, but this may only have a modest impact on the interval's taxonomic diversity (Brusatte et al., 2012). Indeed, measuring diversity solely as simple taxonomic counts can undervalue the true diversity of an interval or the information provided by an individual taxon. Nonetheless, there is concordance in the signals produced by measuring taxonomic diversity, as we have done here, and morphological diversity or disparity (Grossnickle and Polly, 2013; Luo, 2007a). As both approaches offer important and complementary information on deep time diversity dynamics, combining them into comprehensive models provides a more complete understanding of mammaliaform evolution.

Both subsampling and residual approaches suggest relatively low standing species and genus diversity of Mammaliaformes during their inception from the Late Triassic to Early Jurassic. This interval was dominated by morganucodontans and corresponds to Luo's (2007a) episode one. It is followed by elevated diversity in Middle Jurassic units, driven by the radiation of the docodonts, and corresponding to Luo's episode two. However, subsampled and 'residual' diversity of many younger Mesozoic intervals is comparable to that of the Middle Jurassic. Therefore, although younger clades did radiate after the Middle Jurassic (e.g., therians), and certainly affected mammaliaform morphological and ecological diversity, these later episodes did not substantially alter the baseline of Mesozoic mammaliaform taxonomic diversity.

Nevertheless, latest Jurassic and earliest Cretaceous mammaliaform diversity was apparently higher than that of the Middle Jurassic, and statistically well-supported low standing diversity in Cretaceous 2 (Hauterivian–Barremian) marks the transition between faunas dominated by 'Late Jurassic-type' groups (episode 3) and later Cretaceous faunas dominated by tribosphenic mammals in the northern Hemisphere

(episode 4). Although the apparent peak in latest Jurassic–earliest Cretaceous diversity might be an artefact of oversampling of the Morrison Formation (United States) and Lulworth Formation (United Kingdom) (explained above), this pattern of high standing diversity over the Jurassic–Cretaceous boundary is consistent with the occurrence of a distinctly 'Jurassic' fauna observed within Cretaceous 1, with 95% of unique occurrences representing genera either unique to the interval (74%) or of Late Jurassic origin (21%). In contrast, Cretaceous 2 is composed solely of Cretaceous taxa, with only one genus (Loxaulax) carried over from Cretaceous 1, and 40% of occurrences representing taxa also found in later Cretaceous intervals. Furthermore, whereas Cretaceous 1 is dominated by the Lulworth Formation, representing a single environment and geographic region, Cretaceous 2 includes a cosmopolitan set of global faunas with no dominant region or formation. Therefore, underlying diversity is global in Cretaceous 2, but regional in Cretaceous 1, and so should be greater in Cretaceous 2. SQS aims to fairly sample underlying diversity, and should therefore sample more exhaustively in intervals with a larger underlying pool of taxa (Alroy, 2010). For this reason the substantial statistical support for the fall in global subsampled richness between Cretaceous 1 and 2 in our SQS results is even more striking. Such a fall in diversity occurring ~10 myr after the Jurassic–Cretaceous boundary may support the results of previous studies that suggest a delayed, mosaic response to a frequently hypothesised extinction event occurring over the Jurassic–Cretaceous boundary in both terrestrial and marine realms (Jablonski, 2002). However, it is more likely, and more consistent with the evidence, that there simply was no Jurassic–Cretaceous boundary extinction, at least not for mammaliaforms. Instead, our results support recent studies suggesting that there might have been an episode of accelerated turnover during the Early Cretaceous (around the Hauterivian–Barremian; Benson et al., 2013; reviewed by Benson and Druckenmiller, 2014).

The last of the four hypothesised episodes (Luo, 2007a) marks the early diversification of the stem taxa of the tribosphenic mammals in the early Cretaceous. SQS and some residual approaches suggest high diversity in Cretaceous 3. This signal might result from the combination of two long intervals in Cretaceous 3 (Aptian: ~12 myr; Albian: ~13 myr), which may result in an artificially high diversity. However, it is possible that the observed increase in Cretaceous 3 also reflects the comparison with low richness in Cretaceous 2. Diversity within Cretaceous 3 is dominated by Albian taxa, with 83% of unique occurrences during Cretaceous 3 found during this stage alone. Therefore diversity during the interval can be confidently interpreted as a robust increase in diversity and more specifically as a primarily Albian episode, with no substantial Aptian influence grossly amplifying diversity. Finally, the high subsampled and residual diversity during Cretaceous 6–Cretaceous 7 corresponds with the apex of the therian radiation (diversification episode four; Luo, 2007a), and appears to reflect a genuine biotic signal, although the period is evidently also affected by regional sampling heterogeneities, in particular by a high representation of North American localities (74 of the 115 Maastrichtian occurrences are North American). This reconstruction of high diversity in the latest Cretaceous contradicts hypotheses of declining terrestrial vertebrate diversity through the end-Cretaceous leading up to the Cretaceous–Paleogene mass extinction (Archibald, 1986; Archibald et al., 2010; Barrett et al., 2009), but agrees with recent analyses of morphological disparity that suggest that mammaliaforms were increasing in disparity

**Fig. 4.** Asian data series through the Mesozoic. (a) SQS subsampled mammaliaform species taxonomic richness. (b) Mammaliaform-bearing formations (MBFs) vs STDE residuals. (c) Mammaliaform-bearing collections (MBCs) vs species STDE residuals. (d) SQS subsampled mammaliaform generic taxonomic richness. (e) MBF vs GTDE residuals. (f) MBC vs GTDE. Vertical lines represent the Triassic/Jurassic boundary, Early/Mid Jurassic boundary, Mid/Late Jurassic boundary, Jurassic/Cretaceous boundary, and Early/Mid Cretaceous boundary respectively. Dashed lines mark standard error 95% confidence intervals. Dash-dot lines mark standard deviation 95% confidence intervals.
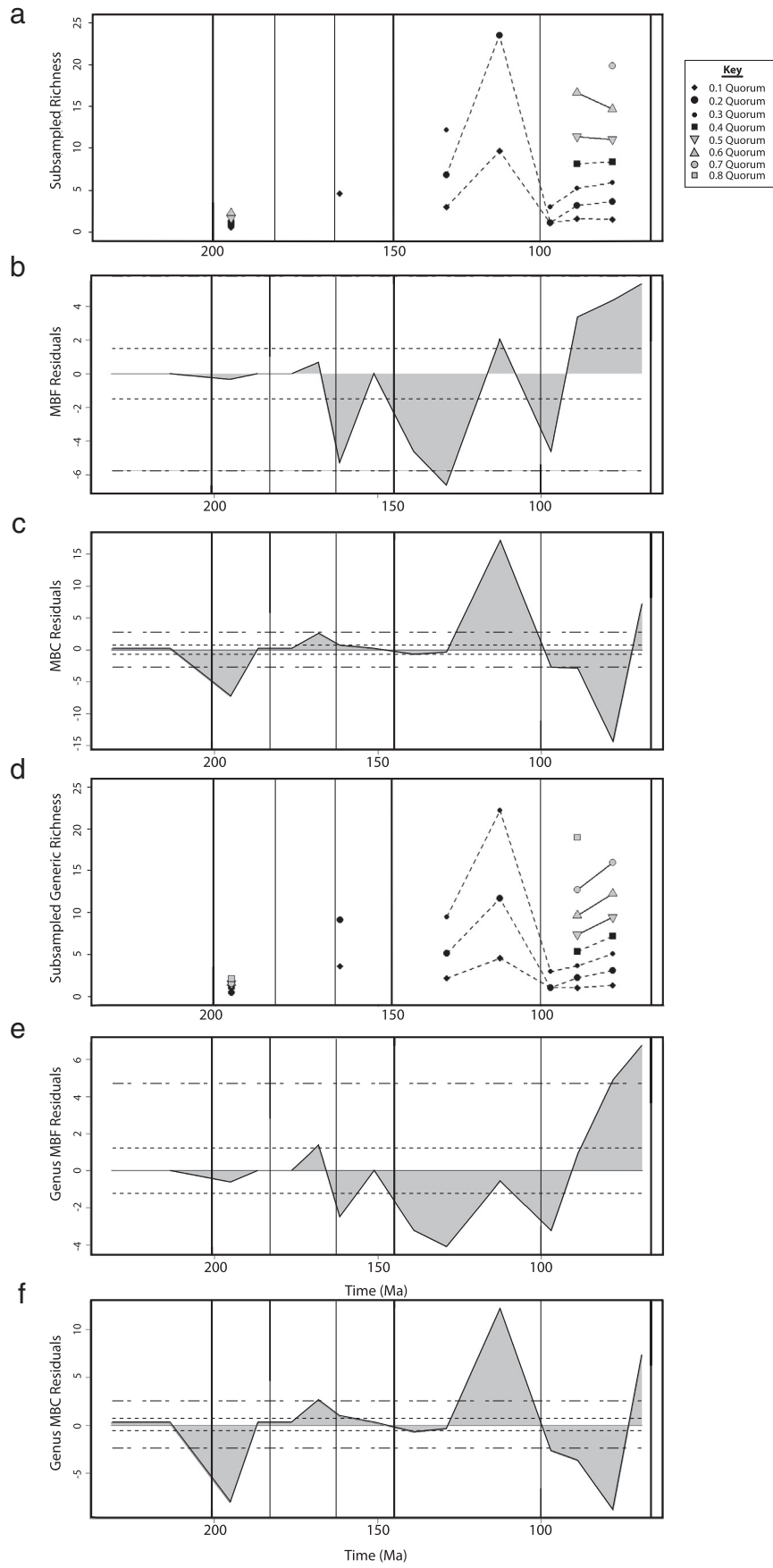
in the latest Cretaceous (Grossnickle and Polly, 2013; Wilson et al., 2012). SQS results also reconstruct a large fall in diversity from Cretaceous 6 to Cretaceous 7, suggesting that mammaliaforms may have been in decline in the last short interval prior to the K/Pg mass extinction. However, it is important to note that the Cretaceous 6 (Campanian) is nearly twice as long as the Cretaceous 7 (Maastrichtian), which might drive this apparent decrease in diversity. Other recent studies have also suggested that the observed fall in diversity from the Campanian to Maastrichtian is due to sampling biases, rather than reflecting a real biotic crisis (Fastovsky and Sheehan, 2005; Fastovsky et al., 2004; Upchurch et al., 2011). Certainly, our analyses do not support a decline in mammaliaform diversity until the Maastrichtian, suggesting that a fall in diversity, if it occurred, did not occur over a substantial period before the K/Pg mass extinction.

As noted above, regional biases may significantly influence hypotheses of clade origination, diversification, and extinction. Some intervals in our dataset exhibited high regional dominance (e.g., Jurassic 6 is dominated by North American occurrences [63% of all occurrences], Cretaceous 1 by British occurrences [67%] and Cretaceous 7 by North American occurrences [64%]), which may explain many of the seemingly dead-end lineages that comprise a large proportion of the Mesozoic mammaliaform record (Luo, 2007a). For example, the diversification of extinct theriiform lineages in the Late Jurassic–Early Cretaceous was shown to be driven by heavily sampled formations, with the majority of Jurassic 6 taxa known only from the North American Morrison Formation, and Cretaceous 1 taxa from the English Lulworth Formation. Other high diversity intervals that correlate with Luo's diversification episodes are followed by relatively long periods of poor fossil representation, highlighted by weak support in both the residual and SQS analyses through Jurassic 5 (following high Jurassic 4 diversity and Luo's episode two), and Cretaceous 4 (following high Cretaceous 3 diversity and Luo's episode four). Thus, the low diversity observed through these weakly represented periods does not necessarily represent a global extinction or a preponderance of short-lived taxa, but rather a lack or loss of their fossil record. This effect can also be observed within the fossil record as long ghost lineages between fossil representatives in genera such as *Morganucodon* (with long gaps between the majority of Late Triassic material and a Mid Jurassic occurrence [Freeman, 1979] and even a possible Early Cretaceous occurrence in England [Butler et al., 2012]). Another example is provided by the substantial missing record between the first eutherian, *Juramaia scansoria*, occurring in the Callovian (~165 mya) (Luo et al., 2011) and the eutherian radiation during the mid Cretaceous (Luo, 2007a). Thus, consideration of biases in the fossil record may have considerable implications for our understanding of the tempo and mode of mammaliaform evolution and extinction.

The substantial missing records we have highlighted can affect current views of mammaliaform evolution. For example, common references to a poor fossil record may be used to support molecular divergence time estimates that are substantially older than the first appearance of a clade in the fossil record, such as placental mammals (Eizirik et al., 2001; Kumar and Hedges, 1998; Springer et al., 2003), although this gap has been reduced substantially in recent molecular studies (dos Reis et al., 2012; Meredith et al., 2011). This apparent missing lineage, combined with a paucity of Cretaceous fossils in some geographic areas, especially Gondwana, has led some workers to suggest a hidden radiation of placental mammals within the southern continents (Hedges et al., 2006). However, improved sampling and analysis of the Mesozoic mammaliaform record in Gondwanan landmasses has failed to yield any evidence for this hypothesis (Goswami, 2012; Goswami et al., 2011; Hunter and Janis, 2006). Cretaceous 6–7 also presents the largest proportion of Gondwanan occurrences throughout the entire Cretaceous (13%), with some endemic eutherian occurrences (*Deccanolestes*, *Kharmerungulatum*), but no well-supported occurrences of placentals.

The suggestion of a long but un-sampled Cretaceous mammalian fossil record has also been met with criticism on the grounds of preservation rate, with Foote et al. (1999) showing that a missing mammalian fossil record, assuming a monotonic increase in diversity between a molecular time of placental origination and the first placental fossil representation, would require a Cretaceous preservation rate of at least an order of magnitude worse than actually observed. A caveat of the methodology employed by Foote et al. (1999) is that it treats preservation as a time-homogenous Poisson process; we have here shown preservation to be markedly heterogeneous, consistent with other recent taxic diversity studies of Cretaceous terrestrial organisms (Brocklehurst et al., 2012; Butler et al., 2009; Upchurch et al., 2011). This evident plasticity in preservation potential is highlighted by weak residual strength in much of the 'mid' to early Late Cretaceous, and might relate to the ~20 myr gap still existing between recent molecular divergence estimates for the origin of placentals and their fossil record (dos Reis et al., 2012; Goswami, 2012) (suggesting that nearly 25% of placental mammal history remains unrepresented in the fossil record). However, Cretaceous 6 provides one of the most statistically robust subsampled diversity estimates of the entire study, and so a lack of any crown placental representation through the interval might well be genuine, giving weight to the palaeontological argument of latest Cretaceous or even earliest Cenozoic origin for placental mammals.

## Acknowledgements

## References

Alroy, J., 2000. New methods for quantifying macroevolutionary patterns and processes. Paleobiology 26/4, 707–733.

Alroy, J., 2008. Dynamics of origination and extinction in the marine fossil record. Proc. Natl. Acad. Sci. U. S. A. 105, 11536–11542.

Alroy, J., 2009. Speciation and extinction in the fossil record of North American mammals. In: Butlin, R., Bridle, J., Schluter, D. (Eds.), Speciation and Patterns of Diversity. Cambridge University Press, Cambridge, pp. 301–323.

Alroy, J., 2010. Geographical, environmental and intrinsic biotic controls on Phanerozoic marine diversification. Palaeontology 53, 1211–1235.

Alroy, J., Marshall, C.R., Miller, A.I., 1998. The palaeobiology database. Available at http://www.paleobiodb.org. Accessed on the February second, 2013.

Alroy, J., Marshall, C.R., Bambach, R.K., Bezusko, K., Foote, M., Fursich, F.T., Hansen, T.A., Holland, S.M., Ivany, L.C., Jablonski, D., Jacobs, D.K., Jones, D.C., Kosnik, M.A., Lidgard, S., Low, S., Miller, A.I., Novack-Gottshall, P.M., Olszewski, T.D., Patzkowsky, M.E., Raup, D.M., Roy, K., Sepkoski, J.J., Sommers, M.G., Wagner, P.J., Webber, A., 2001. Effects of sampling standardization on estimates of Phanerozoic marine diversification. Proc. Natl. Acad. Sci. U. S. A. 9811, 6261–6266.

Alroy, J., Aberhan, M., Bottjer, D.J., Foote, M., Fursich, F.T., Harries, P.J., Hendy, A.J.W., Holland, S.M., Ivany, L.C., Kiessling, W., Kosnik, M.A., Marshall, C.R., McGowan, A.J., Miller, A.I., Olszewski, T.D., Patzkowsky, M.E., Peters, S.E., Villier, L., Wagner, P.J., Bonuso, N., Borkow, P.S., Brenneis, B., Clapham, M.E., Fall, L.M., Ferguson, C.A., Hanson, V.L., Krug, A.Z., Layou, K.M., Leckey, E.H., Nurnberg, S., Powers, C.M., Sessa, J.A., Simpson, C., Tomasovych, A., Visaggi, C., 2008. Phanerozoic trends in the global diversity of marine invertebrates. Science 3215885, 97–100.

Archibald, J.D., 1986. Sedimentology, stratigraphy, and extinctions during the Cretaceous–Paleogene transition at Bug-Creek, Montana — comment. Geology 1410, 892–893.

Archibald, J.D., Clemens, W.A., Padian, K., Rowe, T., Macleod, N., Barrett, P.M., Gale, A., Holroyd, P., Sues, H.D., Arens, N.C., Horner, J.R., Wilson, G.P., Goodwin, M.B., Brochu, C.A., Lofgren, D.L., Hurlbert, S.H., Hartman, J.H., Eberth, D.A., Wignall, P.B., Currie, P.J., Weil, A., Prasad, G.V.R., Dingus, L., Courtillot, V., Milner, A., Milner, A., Bajpai, S., Ward, D.J., Sahni, A., 2010. Cretaceous extinctions: multiple causes. Science 3285981, 973.

Barrett, P.M., McGowan, A.J., Page, V., 2009. Dinosaur diversity and the rock record. Proc. R. Soc. B Biol. Sci. 2761667, 2667–2674.

Behrensmeyer, A.K., Fursich, F.T., Gastaldo, R.A., Kidwell, S.M., Kosnik, M.A., Kowalewski, M., Plotnick, R.E., Rogers, R.R., Alroy, J., 2005. Are the most durable shelly taxa also the most common in the marine fossil record? Paleobiology 314, 607–623.

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate — a practical and powerful approach to multiple testing. J. R. Stat. Soc. Ser. B Methodol. 571, 289–300.

Benson, R.B.J., Butler, R.J., 2011. Uncovering the diversification history of marine tetrapods: ecology influences the effect of geological sampling biases. Geol. Soc. Lond., Spec. Publ. 358, 191–208.

Benson, R.B.J., Druckenmiller, P.S., 2014. Faunal turnover of marine tetrapods during the Jurassic–Cretaceous transition. Biol. Rev. 891, 1–23.

Benson, R.B.J., Upchurch, P., 2013. Diversity trends in the establishment of terrestrial vertebrate ecosystems: interactions between spatial and temporal sampling biases. Geology 41, 43–46.

Benson, R.B.J., Mannion, P.D., Butler, R.J., Upchurch, P., Goswami, A., Evans, S.E., 2013. Cretaceous tetrapod fossil record sampling and faunal turnover: implications for biogeography and the rise of modern clades. Palaeogeogr. Palaeoecol. 372, 88–107.

Benton, M.J., Dunhill, A.M., Lloyd, G.T., Marx, F.G., 2011. Assessing the quality of the fossil record: insights from vertebrates. Geol. Soc. Lond., Spec. Publ. 358, 63–94.

Benton, M.J., Ruta, M., Dunhill, A.M., Sakamoto, M., 2013. The first half of tetrapod evolution, sampling proxies, and fossil record quality. Palaeogeogr. Palaeoclimatol. Palaeoecol. 372, 18–41.

Brocklehurst, N., Fröbisch, J., 2014. Current and historical perspectives on the completeness of the fossil record of pelycosaurian-grade synapsids. Palaeogeogr. Palaeoclimatol. Palaeoecol. 399, 114–126.

Brocklehurst, N., Upchurch, P., Mannion, P.D., O'Connor, J., 2012. The completeness of the fossil record of Mesozoic birds: implications for early avian evolution. Plos One 76.

Brocklehurst, N., Kammerer, C.F., Fröbisch, J., 2013. The early evolution of synapsids, and the influence on their fossil record. Paleobiology 39, 470–490.

Brusatte, S.L., Butler, R.J., Prieto-Marquez, A., Norell, M.A., 2012. Dinosaur morphological diversity and the end-Cretaceous extinction. Nat. Commun. 3.

Butler, R.J., Barrett, P.M., Kenrick, P., Penn, M.G., 2009. Diversity patterns amongst herbivorous dinosaurs and plants during the Cretaceous: implications for hypotheses of dinosaur/angiosperm co-evolution. J. Evol. Biol. 223, 446–459.

Butler, R., Barrett, P., Benson, R., Brusatte, S., Andres, B., 2011a. The taxonomic diversity and morphological disparity of pterosaurs: untangling sampling biases, the impact of Lagerstätten, and diversification trajectories. J. Vertebr. Paleontol. 31, 78–87.

Butler, R., Benson, R.B.J., Carrano, M.T., Mannion, P.D., Upchurch, P., 2011b. Sea level, dinosaur diversity and sampling biases: investigating the 'common cause' hypothesis in the terrestrial realm. Proc. R. Soc. B Biol. Sci. 2781709, 1165–1170.

Butler, P.M., Sigogneau-Russell, D., Ensom, P.C., 2012. Possible persistence of the morganucodontans in the Lower Cretaceous Purbeck Limestone Group Dorset, England. Cretac. Res. 331, 135–145.

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 3015631, 358–360.

dos Reis, M., Inoue, J., Hasegawa, M., Asher, R.J., Donoghue, P.C.J., Yang, Z.H., 2012. Phylogenomic datasets provide both precision and accuracy in estimating the timescale of placental mammal phylogeny. Proc. R. Soc. B Biol. Sci. 2791742, 3491–3500.

Eizirik, E., Murphy, W.J., O'Brien, S.J., 2001. Molecular dating and biogeography of the early placental mammal radiation. J. Hered. 922, 212–219.

Fastovsky, D.E., Sheehan, P.M., 2005. The extinction of the dinosaurs in North America. GSA Today 15, 4–10.

Fastovsky, D.E., Huang, Y., Hsu, J., Martin-McNaughton, J., Sheehan, P.M., Weishampel, D.B., 2004. Shape of Mesozoic dinosaur richness. Geology 3210, 877–880.

Foote, M., Hunter, J.P., Janis, C.M., Sepkoski, J.J., 1999. Evolutionary and preservational constraints on origins of biologic groups: divergence times of eutherian mammals. Science 2835406, 1310–1314.

Freeman, E.F., 1979. a Middle Jurassic mammal bed from Oxfordshire. Palaeontology 221, 135–166.

Fröbisch, J., 2008. Global taxonomic diversity of anomodonts (Tetrapoda, Therapsida) and the terrestrial rock record across the Permian–Triassic boundary. Plos One 311.

Fröbisch, J., 2013. Vertebrate diversity across the end-Permian mass extinction—separating biological and geological signals. Palaeogeogr. Palaeoclimatol. Palaeoecol. 372, 50–61.

Good, I.J., 1953. The population frequencies of species and the estimation of population parameters. Biometrika 403-4, 237–264.

Goswami, A., 2012. A dating success story: genomes and fossils converge on placental mammal origins. Evodevo 3.

Goswami, A., Prasad, G.V.R., Upchurch, P., Boyer, D.M., Seiffert, E.R., Verma, O., Gheerbrant, E., Flynn, J.J., 2011. A radiation of arboreal basal eutherian mammals beginning in the Late Cretaceous of India. Proc. Natl. Acad. Sci. U. S. A. 10839, 16333–16338.

Grossnickle, D.M., Polly, P.D., 2013. Mammal disparity decreases during the Cretaceous angiosperm radiation. Proc. R. Soc. B 280, 20132110.

Hammer, Ø., Harper, D.A.T., Ryan, P.D., 2001. www.uv.es/~pardomv/pe/2001_1/past/pastprog/past.pdf,acessadoem.

Hannisdal, B., Peters, S.E., 2011. Phanerozoic Earth system evolution and marine biodiversity. Science 3346059, 1121–1124.

Hedges, S.B., Kumar, S., Van Tuinen, M., 2006. Constraining fossil calibrations for molecular clocks. Bioessays 287, 770–771.

Hu, Y., Meng, J., Wang, Y., Li, C., 2005. Large Mesozoic mammals fed on yound dinosaurs. Nature 433, 149–152.

Hunter, J.P., Janis, C.M., 2006. "Garden of Eden" or "Fool's Paradise"? Phylogeny, dispersal, and the southern continent hypothesis of placental mammal origins. Paleobiology 323, 339–344.

Irmis, R., Whiteside, J., 2011. Non-biotic controls of Permo-Triassic tetrapod diversity: implications for understanding the End-Permian extinction on land. J. Vertebr. Paleontol. 31, 213–214.

Jablonski, D., 1999. The future of the fossil record. Science 2845423, 2114–2116.

Jablonski, D., 2002. Survival without recovery after mass extinctions. Proc. Natl. Acad. Sci. U. S. A. 9912, 8139–8144.

Ji, Q., Luo, Z.X., Yuan, C.X., Tabrum, A.R., 2006. A swimming mammaliaform from the Middle Jurassic and ecomorphological diversification of early mammals. Science 3115764, 1123–1127.

Johnson, J.B., Omland, K.S., 2004. Model selection in ecology and evolution. Trends Ecol. Evol. 192, 101–108.

Kielan-Jaworowska, Z., Cifelli, R., Luo, Z.X., 2004. Mammals from the Age of Dinosaurs: Origins, Evolution, and StructureColumbia University Press, New York; Chichester.

Kowalewski, M., 2002. The fossil record of predation: an overview of analytical methods. Palaeontol. Soc. Pap. 8, 3–42.

Kumar, S., Hedges, S.B., 1998. A molecular timescale for vertebrate evolution. Nature 3926679, 917–920.

Lloyd, G.T., 2012. A refined modelling approach to assess the influence of sampling on palaeobiodiversity curves: new support for declining Cretaceous dinosaur richness. Biol. Lett. 81, 123–126.

Lloyd, G.T., Davis, K.E., Pisani, D., Tarver, J.E., Ruta, M., Sakamoto, M., Hone, D.W.E., Jennings, R., Benton, M.J., 2008. Dinosaurs and the Cretaceous Terrestrial Revolution. Proc. R. Soc. B Biol. Sci. 2751650, 2483–2490.

Luo, Z.X., 2007a. Transformation and diversification in early mammal evolution. Nature 4507172, 1011–1019.

Luo, Z.X., 2007b. Successive diversifications in early mammalian evolution. In: Anderson, J.S., Sues, H.-D. (Eds.), Major Transitions in Vertebrate Evolution. Indiana University Press, Bloomington and Indianapolis, Indiana, pp. 337–391.

Luo, Z.X., Wible, J.R., 2005. A Late Jurassic digging mammal and early mammalian diversification. Science 3085718, 103–107.

Luo, Z.X., Crompton, A.W., Sun, A.L., 2001. A new mammaliaform from the Early Jurassic and evolution of mammalian characteristics. Science 2925521, 1535–1540.

Luo, Z.X., Kielan-Jaworowska, Z., Cifelli, R., 2002. In quest for a phylogeny of Mesozoic mammals. Acta Palaeontol. Pol. 471, 1–78.

Luo, Z.X., Ruf, I., Schultz, J.A., Martin, T., 2011. Fossil evidence on evolution of inner ear cochlea in Jurassic mammals. Proc. R. Soc. B Biol. Sci. 2781702, 28–34.

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. Biol. Rev. 861, 157–181.

Mannion, P.D., Benson, R.B.J., Upchurch, P., Butler, R.J., Carrano, M.T., Barrett, P.M., 2012. A temperate palaeodiversity peak in Mesozoic dinosaurs and evidence for Late Cretaceous geographical partitioning. Glob. Ecol. Biogeogr. 219, 898–908.

McKinney, M.L., 1990. In: McNamara, K.J. (Ed.), Classifying and Analysing Evolutionary Trends Evolutionary Trends. University of Arizona Press, Tuscon, pp. 28–58.

Meng, J., Hu, Y.M., Wang, Y.Q., Wang, X.L., Li, C.K., 2006. A mesozoic gliding mammal from northeastern china. Nature 4447121, 889–893.

Meredith, R.W., Janecka, J.E., Gatesy, J., Ryder, O.A., Fisher, C.A., Teeling, E.C., Goodbla, A., Eizirik, E., Simao, T.L.L., Stadler, T., Rabosky, D.L., Honeycutt, R.L., Flynn, J.J., Ingram, C.M., Steiner, C., Williams, T.L., Robinson, T.J., Burk-Herrick, A., Westerman, M., Ayoub, N.A., Springer, M.S., Murphy, W.J., 2011. Impacts of the Cretaceous terrestrial revolution and KPg extinction on mammal diversification. Science 3346055, 521–524.

Peters, S.E., 2005. Geologic constraints on the macroevolutionary history of marine animals. Proc. Natl. Acad. Sci. U. S. A. 10235, 12326–12331.

Peters, S.E., Foote, M., 2001. Biodiversity in the Phanerozoic: a reinterpretation. Paleobiology 274, 583–601.

Raup, D.M., 1972. Taxonomic diversity during Phanerozoic. Science 1774054, 1065–1071.

Raup, D.M., Gould, S.J., Schopf, T.J.M., Simberloff, D.S., 1975. Stochastic models of phylogeny and evolution of diversity — reply. J. Geol. 831, 126–127.

Rose, K.D., 2006. The Beginning of the Age of MammalsJohn Hopkins University Press, Maryland.

Rowe, T., 1988. Definition, diagnosis, and origin of Mammalia. J. Vertebr. Paleontol. 83, 241–264.

Sames, B., Cifelli, R.L., Schudack, M.E., 2010. The nonmarine Lower Cretaceous of the North American Western Interior foreland basin: new biostratigraphic results from ostracod correlations and early mammals, and their implications for paleontology and geology of the basin — an overview. Earth Sci. Rev. 1013-4, 207–224.

Smith, A.B., 2001. Large-scale heterogeneity of the fossil record: implications for Phanerozoic biodiversity studies. Philos. Trans. R. Soc. B Biol. Sci. 3561407, 351–367.

Smith, A.B., 2007. Marine diversity through the Phanerozoic: problems and prospects. J. Geol. Soc. 164, 731–745.

Smith, A.B., Benson, R.B.J., 2013. Marine diversity in the geological record and its relationship to surviving bedrock area, lithofacies diversity, and original marine shelf area. Geology 412, 171–174.

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.

Smith, A.B., Lloyd, G.T., McGowan, A.J., 2012. Phanerozoic marine diversity: rock record modelling provides an independent test of large-scale trends. Proc. R. Soc. B Biol. Sci. 2791746, 4489–4495.

Springer, M.S., Murphy, W.J., Eizirik, E., O'Brien, S.J., 2003. Placental mammal diversification and the Cretaceous–Tertiary boundary. Proc. Natl. Acad. Sci. U. S. A. 1003, 1056–1061.

Upchurch, P., Barrett, P.M., 2005. Sauropodomorph diversity through time. In: Rogers, K.C., Wilson, J.A. (Eds.), The Sauropods: Evolution and Paleobiology. University of California Press, California, pp. 125–156.

Upchurch, P., Mannion, P.D., Butler, R.J., Benson, R.B.J., Carrano, M.T., 2011. Geological and anthropogenic controls on the sampling of the terrestrial fossil record: a case study from the Dinosauria. Geol. Soc. Lond., Spec. Publ. 358, 209–240.

Valentine, J.W., 1985. Phanerozoic Diversity Patterns: Profiles in MacroevolutionAmerican Association for the Advancement of Science, San Francisco, California, Pacific Division.

Wilson, G.P., 2005. Mammalian faunal dynamics during the last 1.8 million years of the Cretaceous in Garfield County, Montana. J. Mamm. Evol. 121-2, 53–76.

Wilson, G.P., 2013. Mammals across the K/Pg boundary in northeastern Montana, U.S.A.: dental morphology and body-size patterns reveal extinction selectivity and immigrant-fueled ecospace filling. Paleobiology 39, 429–469.

Wilson, G.P., Evans, A.R., Corfe, I.J., Smits, P.D., Fortelius, M., Jernvall, J., 2012. Adaptive radiation of multituberculate mammals before the extinction of dinosaurs. Nature 4837390, 457–460.

Zhou, C.F., Wu, S.Y., Martin, T., Luo, Z.X., 2013. A Jurassic mammaliaform and the earliest mammalian evolutionary adaptations. Nature 5007461, 163–167.