

Phanerozoic marine diversity: rock record modelling provides an independent test of large-scale trends

Andrew B. Smith^{1,*}, Graeme T. Lloyd² and Alistair J. McGowan³

¹*Department of Palaeontology, The Natural History Museum, London SW7 5BD, UK*

²*Department of Earth Sciences, University of Oxford, South Parks Road, Oxford OX1 3AN UK*

³*School of Geographical and Earth Sciences, University of Glasgow, Glasgow G12 8QQ UK*

Sampling bias created by a heterogeneous rock record can seriously distort estimates of marine diversity and makes a direct reading of the fossil record unreliable. Here we compare two independent estimates of Phanerozoic marine diversity that explicitly take account of variation in sampling—a subsampling approach that standardizes for differences in fossil collection intensity, and a rock area modelling approach that takes account of differences in rock availability. Using the fossil records of North America and Western Europe, we demonstrate that a modelling approach applied to the combined data produces results that are significantly correlated with those derived from subsampling. This concordance between independent approaches argues strongly for the reality of the large-scale trends in diversity we identify from both approaches.

Keywords: diversity; evolution; sampling bias; rock record; invertebrates; fossil record

1. INTRODUCTION

Until relatively recently, the geological history of diversity was extracted from the fossil record by tallying the numbers of taxa recorded in the literature from each time interval [1–3]. However, it is now clear that our sampling of the geological record is far from uniform because the amount of rock and range of facies surviving from different time intervals are highly variable [4,5], and there are marked differences in the thoroughness that those outcrops have been sampled and recorded by palaeontologists [6,7]. Since the pioneering work of Raup [8], many studies have confirmed the existence of a strong positive correlation between the geological record available at outcrop and the fossil diversity recovered from that record (see [9] for recent review papers). The reason for this correlation is still debated but involves the interplay between (i) major sea-level cycles driving both biological diversity and the amount of sedimentary record that is deposited (the ‘common cause’ hypothesis of Peters [10]), and (ii) subsequent geological cycles of deposition, tectonism and erosion that affect what survives at outcrop and create a biased rock record [4]. Irrespective of the driving mechanisms, palaeobiologists need to remove any sampling bias signal from raw taxonomic diversity curves to arrive at accurate estimates of biological diversity through time. Two very different approaches have been developed specifically to address the problem of uneven sampling—subsampling of taxonomic lists and modelling with rock sampling proxies.

A subsampling approach to the problem of Phanerozoic diversity became possible with the development of the Paleobiology Database ([6]: see <http://paleodb.org/>),

which lists the taxa recorded from specific localities and time intervals and recently surpassed a million individual occurrences. A variety of subsampling techniques were originally employed to obtain equal-sized counts from each time interval, all of which had some limitations [6,11]. More recently a new subsampling technique called Shareholder Quorum Subsampling (SQS) has been developed [12]. This aims to sample fairly rather than evenly by drawing samples of varying size that represent the same proportional area, or coverage, subtended by the taxon-abundance curve. The SQS approach has the distinct advantage that it is an unbiased estimator and the diversity ratio between two samples can be taken literally. SQS has now been used to explore the pattern of Phanerozoic marine diversity [13,14] and the fossil records of coccolithophorids [15]. Whether subsampled evenly or fairly, the resultant diversity curves obtained in those studies were markedly different from those based on raw taxonomic counts (but see [14]).

The second approach to removing sampling bias starts from an empirical measure of the rock record available for sampling (or a proxy for this), and develops a model that predicts what the fossil record would look like if diversity were uniform over time and driven entirely by the area of rock available for sampling [5,16]. Removing the modelled from the observed diversity generates a curve of the residual change that cannot be predicted from a simple species–rock area relationship, irrespective of whether that relationship arises from sampling bias or common cause mechanism (or a combination of the two). The null hypothesis is that either driving mechanism will create a pattern in which sampled diversity perfectly tracks rock area. Any deviation represents signal where diversity is rising/falling faster than common cause or sampling bias would predict. This approach therefore seeks to identify relative changes and trends in biodiversity that are over and above those

* Author for correspondence (a.smith@nhm.ac.uk).

Electronic supplementary material is available at <http://dx.doi.org/10.1098/rspb.2012.1793> or via <http://rspb.royalsocietypublishing.org>.

predicted from our null hypothesis. It has been used to explore both terrestrial and marine diversity patterns in the fossil record [16–22].

To date, to our knowledge, only one study has explored the congruence between these different approaches to the problem of sampling bias. Lloyd *et al.* [15] applied both methods to the 150 Ma fossil record of coccolithophorids and found that they generated broadly similar results, very different from the raw diversity plots. Here we develop a model-based estimate of Phanerozoic marine diversity in North America and Europe, and compare this with recently published diversity estimates based on SQS subsampling. Any agreement in the diversity curves obtained from such very different approaches will lend strong support to the validity of the results, while areas of disagreement will highlight parts of the curve that are less secure and require further investigation.

2. METHOD AND MATERIALS

All time-series data were plotted using the same 48 time bins that divide the Phanerozoic into intervals of approximately 11 Myr duration (as defined and previously used by Alroy *et al.* [7]). Marine diversity was estimated at genus level. Sampled diversity in each time interval for North America (USA+Canada) and for Western Europe (UK, Spain, Portugal, France, Belgium, The Netherlands, Denmark, Germany, Austria and Switzerland) was independently extracted from the Paleobiology Database (PaleoDB: <http://palaedb.org/>, accessed March 2012). These data were cleaned by removing obvious spelling errors of taxon names and terrestrial organisms preserved in marine sediments (e.g. spores and pollen, terrestrial plants and dinosaurs). Taxonomic lists from the two regions were then pooled and multiple occurrences removed to arrive at a sampled genus diversity for North America plus Western Europe. For our estimate of global sampled genus diversity, we use the Sepkoski compilation (available at <http://strata.geology.wisc.edu/jack/>), binned into our 48 time intervals. Unlike the PaleoDB estimates, which give sampled in bin diversity, the Sepkoski curve uses range-through data. We use two estimates of global diversity corrected for sampling variation by subsampling; the Phanerozoic diversity curve of Alroy *et al.* [7] based on even subsampling of PaleoDB collections, and that of Alroy [13] based on his SQS method. For comparative plotting purposes, time series are rescaled by subtracting the mean value from each datum in a series.

Bedrock area of marine sediments in Western Europe was calculated as the number of 1:50 000 maps with recorded outcrop of each time interval for the UK, Spain and France using the same raw data as Smith & McGowan [5] but binned into 48 rather than 72 time intervals. For our estimate of the comparative rock areas across North America, we use numbers of sediment packages ([23], with updates supplied by Shanan Peters in May 2012). Both bedrock map area and numbers of sedimentary packages provide measures of the relative amount of rock available to palaeontologists from each time bin. In total, there are about twice as many rock proxy records from North America as from Europe (14 532 versus 6878) but slightly fewer genus sampled-in-bin records (14 072 versus 19 013), while the surface area of North America represents about three times that of Europe ($19.6 \times 10^6 \text{ km}^2$ versus $6.0 \times 10^6 \text{ km}^2$: <http://wolfram-alpha.com>). As a first approximation, we simply added the

two proxies together to derive our estimate of rock availability for the two regions combined.

The various time series were first logged to base *e* to remove the effects of outliers. Next, to model what sampled diversity would look like were it to perfectly reflect rock record availability we followed the method of Lloyd [16]. First the rock counts (numbers of maps with outcrop of a given age or stratigraphic columns with sediment packages of a given age) and diversity counts (drawn from the Paleobiology Database) were independently ordered from smallest to largest. The equation of the best-fitting model (using the Akaike information criterion with correction) to these data can then be applied to the rock record in its original time series. Removing the model from the empirical count of sampled genus diversity provides a minimum estimate of the amount of diversity that would be encountered if rock record availability were uniform in each time bin of the Phanerozoic. Confidence intervals (95%) were calculated based on 1.96 s.d. of the model.

Correlation of the raw data was carried out to identify long-term trends using Spearman Ranks. Time series were also detrended by taking generalized differences [24] and the correlations compared with test the short-term correlation in bin-to-bin shifts. Medium-term (multi-time bin) trends are recovered by using the Multivariate Adaptive Regression Splines (MARS) approach of Friedman [25], as implemented by Lloyd [16] and using the ‘Earth’ package for R [26]. This is a statistically robust method for identifying hinge points in a time series that automatically minimizes the residual sum of squares.

All data are available in the electronic supplementary material.

3. RESULTS

(a) Comparison of subsampling approaches

The two subsampling methods (Alroy *et al.*'s [7] uniform sampling and Alroy's [13] fair sampling) generate closely similar time series with high levels of statistical correlation at both long and short time-scales (see the electronic supplementary material, table S1). SQS provides the less dampened curve, as expected, and is used for all subsequent comparisons. The fit of the two subsampling curves to Sepkoski's range-through data is poorer but still significant (see the electronic supplementary material, table S1), largely owing to the post-Triassic rise in diversity present in all three. Major differences exist between the two curves, as highlighted previously [6,7]. Compared with the Sepkoski curve, the SQS curve shows a more extended rise to an Early Devonian peak, a marked dip in diversity in the Late Devonian and Carboniferous, a less pronounced end Permian drop, and a more pronounced Early Jurassic trough (figure 1).

(b) Comparison of rock modelling approaches

The rock area records of Western Europe and North America differ markedly (figure 2*a–c*). Key differences include a proportionally greater Early Palaeozoic record in North America, greater Late Triassic–Jurassic representation in Western Europe, and the very poor representation of Permo-Triassic marine sequences in Western Europe. Only from the Cretaceous onwards do the two curves track one another, as both records respond to the opening of the Atlantic and falling global sea levels.

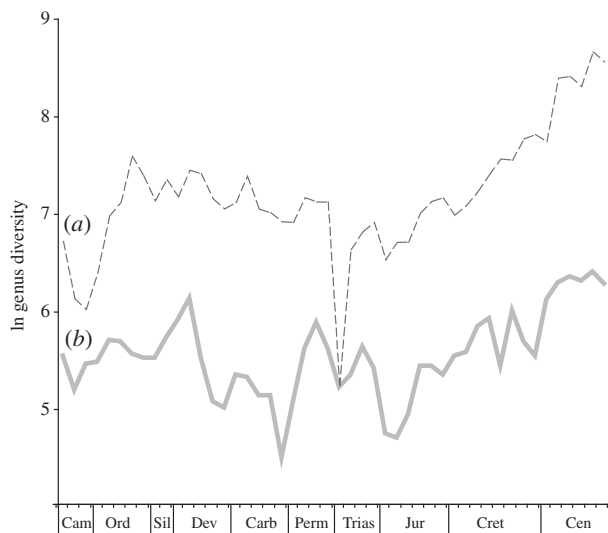


Figure 1. Comparative global marine biodiversity through time based on (a) raw counts of recorded taxa with range interpolation between first and last occurrences [27], and (b) sampled taxa recorded in the Paleobiology Database and corrected for variation in sampling using SQS (from [13]).

Sampled genus diversity is also different in the two regions (figure 2*a,b,d*), with North America registering proportionally higher sampled diversity in the early Palaeozoic and Permian, and lower diversity from the Late Triassic onwards, mirroring differences in their respective rock records. A strong positive correlation between sampled diversity and rock area proxy is found for both regions (see the electronic supplementary material, table S2). As previously reported by Peters & Heim [23], the correlation between our rock proxies and estimates of diversity strengthens as the datasets converge in their geographical scope. Non-significant levels of correlation are found between model and empirical data when models are constructed from rock proxy data and diversity data that differ in spatial extent, but correlation becomes strongly significant when both rock proxy and diversity data are drawn from identical regions (see the electronic supplementary material, table S2).

The match between modelled and empirical diversities is good although far from perfect, and the residuals plot reveals the diversity change over and above that expected from species–rock area considerations (figure 3 and table 1). Not surprisingly, a different residual curve is generated for each region when the model is subtracted from the observed diversity, the two showing little similarity until the Cretaceous. The lack of statistically significant correlation between the two models reinforces the finding that diversity curves and rock record curves show strong regional signatures that are asynchronous at the global scale [28].

Combining rock proxies for North America and Western Europe and applying the modelling approach to our pooled sampled diversity for the same regions provides an estimate of diversity across a large part of the Northern Hemisphere (figure 3*e*). Again the rock-based model provides a good but far from perfect fit to the observed sampled diversity (table 1), with four intervals when diversity is greater than predicted and four when it is less (figure 3*f*). However, the only medium-term trend identified by MARS analysis

is the rise in diversity since the Jurassic. Diversity highs occur in the Mid-Devonian, Mid-Permian, Early Cretaceous and Cenozoic; and lows in the Late Cambrian, Late Devonian, Early Jurassic and Mid-Cretaceous; all of which are statistically significant.

(c) Comparison of rock model and subsampling approaches

When subsampling and modelling estimates of diversity are compared (figure 3*g*) the long-term trends are reassuringly similar. Features common to both plots include: a rise to Devonian 3 followed by a subsequent sharp fall in Devonian 4, a second peak in the Permian and fall in the Early Triassic, an extended period of low diversity in the early Jurassic and a steady rise in diversity to a Cenozoic high. The raw data are strongly correlated ($\rho = 0.628$; figure 2*f*), but after differencing the strength of correlation drops and becomes non-significant ($\rho = 0.234$; $p = 0.11$). Three time intervals exhibit striking differences between rock model and subsampling approaches. In the Mid-Ordovician, rock modelling indicates diversity is higher than expected whereas subsampling suggests diversity is much lower than expected, while in the Late Carboniferous and Early Permian subsampling implies a diversity that is much lower than that predicted from rock record modelling. Finally, in the Mid-Cretaceous our rock model predicts a much lower and more extended drop in diversity than subsampling.

Plotting sampling intensity in the Paleobiology Database against our rock record proxies (see the electronic supplementary material, figure S1*b*) reveals unexpectedly large oversampling in Ordovician 4, Devonian 3 and through the interval Carboniferous 5 to Permian 4.

4. DISCUSSION

The history of biodiversity through geological time has to be estimated from the rock record that survives today and is accessible to palaeontologists at outcrop. All subsampling methods are dependent on the fossil collections available to us. If the collections cannot be matched for facies/habitat then the geological record will introduce an inherent bias of varying severity before a single fossil is collected. This constraint imposes an intrinsic heterogeneity to the fossil record even before the very first fossil is collected, and which must be taken into account somehow if we are to derive an accurate estimate of relative biodiversity change through time. An ecological census that failed to account for the effects of habitat would rightly be regarded as flawed and palaeobiologists should heed Erwin's call to model the causes of biodiversity fluctuations, which must incorporate an understanding of biases [29]. Two approaches currently try to remove the effects of this sampling heterogeneity: (i) correcting for variation in numbers of fossil collections through subsampling [6,7,12–14]; and (ii) modelling to find what proportion of the sampled diversity curve remains unexplained by variation in the area of rock record surviving [5,16–22]. Both try to isolate the biological signal that lies hidden in the fossil record but apply different assumptions and use different data (lists of taxa in collections and measures of rock availability). The former corrects for variation in the relative numbers of collections that have been made, while

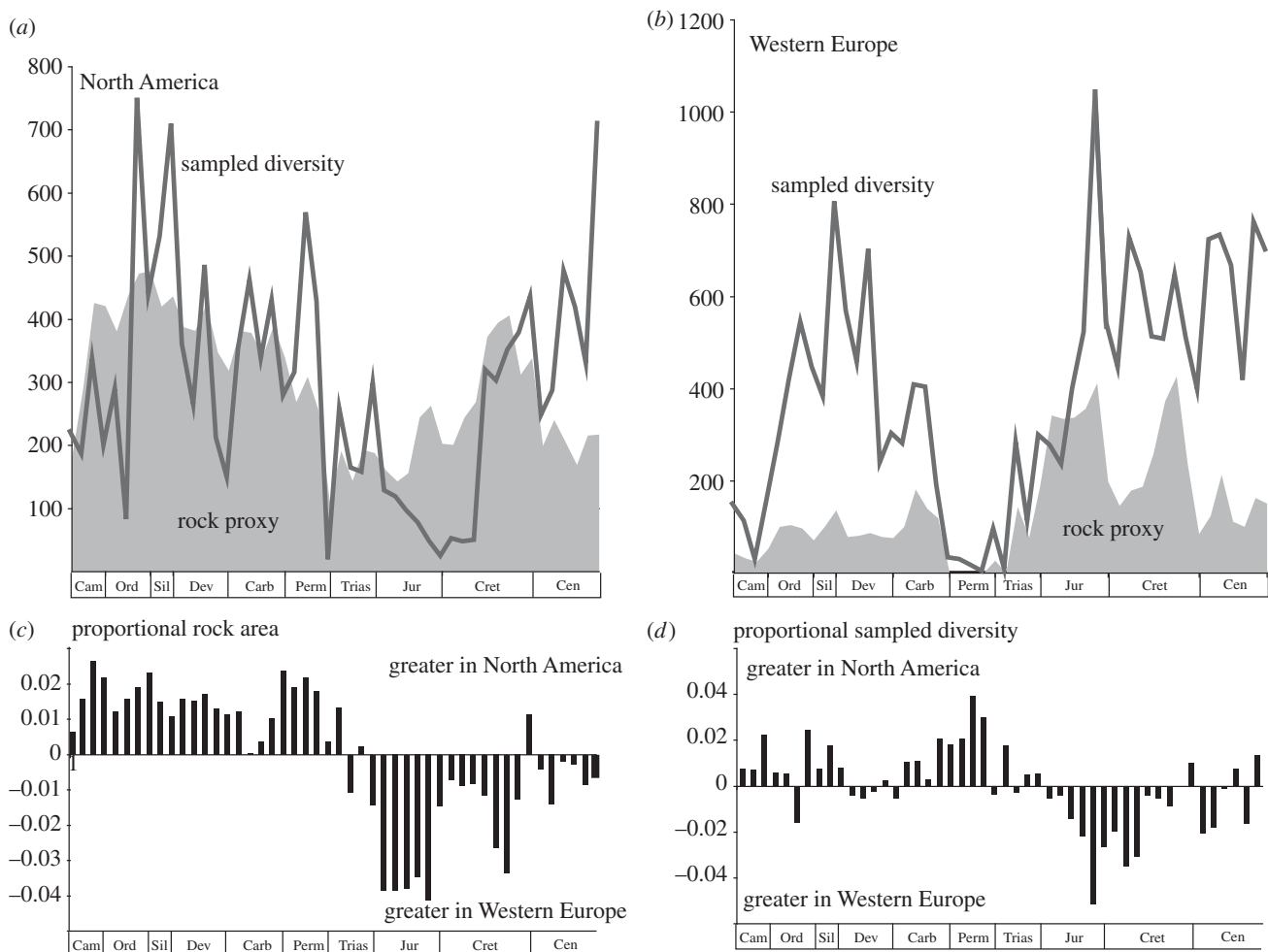


Figure 2. (a) Plots of sampled genus richness versus rock area proxy for North America. (b) Same for Western Europe. (c) Residuals after subtracting North American rock record from Western Europe rock record showing comparative strengths of the two records. (d) Same for sampled diversity.

the latter corrects for variation introduced by the different amount of outcrop that has survived from each time interval. Of course both diversity and outcrop area might be controlled by the spatial extent of epic seas in the geological past, in which case correcting for rock area will underestimate true diversity. In this study, we can only identify changes in diversity that are greater or smaller than expected from the null model. However, once sampling irregularities have been removed from our database there is no correlation ($r^2 = 0.001$) between rock proxy and SQS diversity, suggesting that the significant correlation ($r^2 = 0.43$) between rock proxy and raw diversity counts arises primarily from rock bias.

A second problem is that the datasets being compared are not exactly comparable in their geographical extent, as the rock record modelling is based on just North America plus Western Europe while the subsampling is based on global data in the Paleobiology Database. However, the Paleobiology Database is heavily biased towards records from North America and Europe with, for example, 80 per cent of Silurian Paleobiology Database records coming from these two regions. Consequently, diversity estimates derived from the Paleobiology Database will predominantly reflect patterns specific to those two regions. We find it very reassuring, therefore, that despite these caveats our two independent approaches, which

begin from very different proxies of sampling intensity, converge on the same large-scale Phanerozoic pattern.

Congruence among independent metrics provides strong evidence that the patterns being identified are well founded, and our results confirm the validity of many aspects of Alroy's [13] diversity curve over the earlier genus-level curve of Sepkoski [3]. Patterns for which we have corroboratory evidence for include: (i) the long-term trend of increasing diversity from Cambrian 2 to Devonian 3, (ii) an extended phase of low diversity in the Triassic and Early Jurassic, and (iii) a long-term trend of increasing diversity from the Jurassic to Recent, punctuated by a small fall in the Mid-Cretaceous. Mismatch is most striking at three time intervals: Ordovician 3, Carboniferous 5 and Cretaceous 5. In the first of these, rock area modelling identifies Ordovician 3 as a period of unexpectedly low diversity. This reflects the extremely low reported sampling of this time interval in the Paleobiology Database rather than any marked shift in rock area availability (see the electronic supplementary material, data figure S1). By contrast, Carboniferous 5 is identified as having a very low diversity fauna in the subsampling analysis but not the rock area modelling, which shows a more or less steadily increasing diversity from the Late Devonian through to the Mid-Permian. This is a period when there is virtually no marine record in

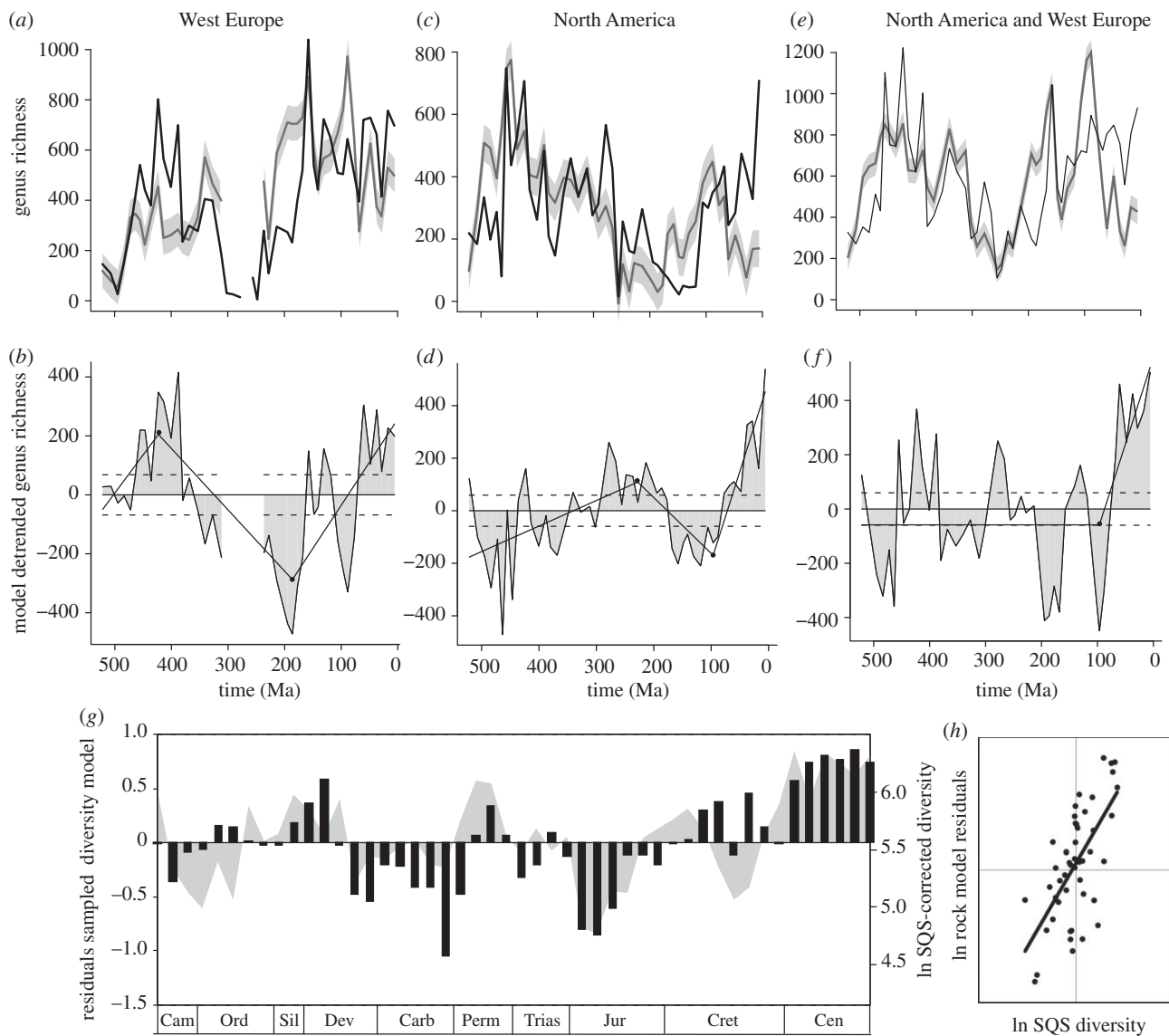


Figure 3. (a) Plot of sampled genus richness for North America in the Paleobiology Database (black line) versus a model of what diversity would look like assuming true diversity is invariant and sampled diversity were driven solely by the rock area available in North America (grey line with 95% confidence intervals). (b) Residuals from subtracting the modelled diversity from the observed sampled diversity shown in (a). (c,d) Same for Western Europe. (e,f) Same for combined European and North American data. (g) Residuals from subtracting modelled diversity from the observed sampled diversity for Europe and North America combined (shaded area) plotted against SQS (from [13]), normalized around the mean (bars). (h) Correlation of modelled versus subsampled estimates of marine diversity.

Table 1. Statistical correlations of the match between Paleobiology Database sampled diversities and model-based diversity residuals for the same geographical regions based on Spearman Rank. (p , probability; PaleoDB, Paleobiology Database. Statistically significant results in bold.)

	raw	generalized differences
Europe: rock proxy model versus PaleoDB diversity (figure 3a)	$\rho = 0.488$	$\rho = 0.478$; $p = 0.002$
North America: rock proxy model versus PaleoDB diversity (figure 3c)	$\rho = 0.478$	$\rho = 0.416$; $p = 0.004$
Europe + North America: rock proxy model versus PaleoDB diversity (figure 3e)	$\rho = 0.320$	$\rho = 0.417$; $p = 0.003$
Europe + North America: rock proxy model residuals versus SQS diversity from [13] (figure 3g)	$\rho = 0.648$	$\rho = 0.234$; $p = 0.110$

Western Europe and it may be that signal from outside the North America–Western Europe region is strongly influencing this part of the SQS curve. Furthermore, this interval is unusually densely sampled in the Paleobiology Database (see the electronic supplementary material,

data figure S1) contributing to their anomalously high sampled diversities. Third, in the Mid-Cretaceous, where rock area modelling suggests a major drop in diversity, subsampling suggests only a minor dip. This was a time of maximum flooding when deep-water chalk

facies flooded across extensive swathes of the continental blocks. Sediments of this age are thus widespread but represent only a very restricted range of marine facies [30]. Here outcrop area is a poor indicator of sampled diversity. As recently noted by Holland [27], the link between diversity, sea-level and spatial distribution of lithofacies can be idiosyncratic.

Large-scale patterns in diversity change have now been shown to match major trace isotope trends and large-scale plate tectonic cycles [14]. These large-scale cycles, which are in the order of *ca* 100–150 Myr ago, shown up as cycles of rising and falling diversity irrespective of our approach to estimating diversity and thus seem securely founded. However, smaller bin-to-bin shifts in diversity are much less consistently recognized, with our two regional patterns, after generalized differencing, showing no significant correlation. Generalized differencing also fails to find any significant correlations between SQS and rock modelling estimates of diversity. A strong regional signature to the rock record has previously been noted [26], and our work confirms this is also paralleled in the biodiversity signal. Thus, support for regarding small-scale shifts as having true biological significance is limited, and the interpretation of bin-to-bin fluctuations should be treated with extreme care.

We would like to thank our three anonymous referees for their constructive criticisms of earlier drafts of this paper, and Shanan Peters for making his most recent compilation of North American sediment packages available to us. This paper represents contribution no. 164 of the Paleobiology Database.

REFERENCES

- Sepkoski, J. J., Bambach, R. K., Raup, D. M. & Valentine, J. W. 1981 Phanerozoic marine diversity and the fossil record. *Nature* **293**, 435–437. (doi:10.1038/293435a0)
- Benton, M. J. 1995 Diversity and extinction in the history of life. *Science* **268**, 52–58. (doi:10.1126/science.7701342)
- Sepkoski, J. J. 1997 Biodiversity: past, present, and future. *J. Paleontol.* **71**, 533–539.
- Smith, A. B. 2001 Large-scale heterogeneity of the fossil record: implications for Phanerozoic biodiversity studies. *Phil. Trans. R. Soc. Lond. B* **356**, 351–367. (doi:10.1098/rstb.2000.0768)
- Smith, A. B. & McGowan, A. J. 2007 The shape of the Phanerozoic diversity curve. How much can be predicted from the sedimentary rock record of Western Europe? *Palaeontology* **50**, 765–777. (doi:10.1111/j.1475-4983.2007.00693.x)
- Alroy, J. *et al.* 2001 Effects of sampling standardization on estimates of Phanerozoic marine diversification. *Proc. Natl Acad. Sci. USA* **98**, 6261–6266. (doi:10.1073/pnas.111144698)
- Alroy, J. *et al.* 2008 Phanerozoic trends in the global diversity of marine invertebrates. *Science* **321**, 97–100. (doi:10.1126/science.1156963)
- Raup, D. M. 1972 Taxonomic diversity during the Phanerozoic. *Science* **177**, 1065–1071. (doi:10.1126/science.177.4054.1065)
- McGowan, A. J. & Smith, A. B. (eds) 2011 *Comparing the geological and fossil records: implications for biodiversity studies*, vol. 358, pp. 247. London, UK: Geological Society of London Special Publications.
- Peters, S. E. 2005 Geological constraints on the macro-evolutionary history of marine animals. *Proc. Natl Acad. Sci. USA* **102**, 12 326–12 331. (doi:10.1073/pnas.0502616102)
- Bush, A. M., Markey, M. J. & Marshall, C. R. 2004 Removing bias from diversity curves: the effects of spatially organized biodiversity on sampling-standardization. *Paleobiology* **30**, 666–686. (doi:10.1666/0094-8373(2004)030<0666:RBFDCCT>2.0.CO;2)
- Alroy, J. 2010 Fair sampling of taxonomic richness and unbiased estimation of origination and extinction rates. Quantitative methods in paleobiology. *Paleontol. Soc. Pap.* **16**, 55–80.
- Alroy, J. 2010 The shifting balance of diversity among major marine animal groups. *Science* **329**, 1191–1194. (doi:10.1126/science.1189910)
- Hannisdal, B. & Peters, S. E. 2011 Phanerozoic Earth system evolution and marine biodiversity. *Science* **334**, 1121–1124. (doi:10.1126/science.1210695)
- Lloyd, G. T., Young, J. R. & Smith, A. B. 2012 Comparative quality and fidelity of the deep-sea and land-based fossil records. *Geology* **40**, 155–158. (doi:10.1130/G32561.1)
- 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.* **8**, 123–126. (doi:10.1098/rsbl.2011.0210)
- Barrett, P. M., McGowan, A. J. & Page, V. 2009 Dinosaur diversity and the rock record. *Proc. R. Soc. B* **276**, 2667–2674. (doi:10.1098/rspb.2009.0352)
- Benson, R. J. B., Butler, R. J., Lindgren, J. & Smith, A. S. 2010 Mesozoic marine tetrapod diversity: mass extinctions and temporal heterogeneity in geological megabiases affecting vertebrates. *Proc. R. Soc. B* **227**, 829–834. (doi:10.1098/rspb.2009.1845)
- Benson, R. J. B. & Butler, R. J. 2011 Uncovering the diversification history of marine tetrapods: ecology influences the effect of geological sampling biases. In *Comparing the geological and fossil records: implications for biodiversity studies*, vol. 358 (eds A. J. McGowan & A. B. Smith), pp. 191–208. London, UK: Geological Society of London Special Publications.
- 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. *Proc. R. Soc. B* **278**, 1165–1170. (doi:10.1098/rspb.2010.1754)
- Mannion, P. D., Upchurch, P., Carrano, M. T. & Barrett, P. M. 2010 Testing the effect of the rock record on diversity: a multidisciplinary approach to elucidating the generic richness of sauropodomorph dinosaurs through time. *Biol. Rev.* **86**, 157–181. (doi:10.1111/j.1469-185X.2010.00139.x)
- Upchurch, P., Mannion, P. D., Benson, R. B. J., Butler, R. J. & Carrano, M. T. 2011 Geological and anthropogenic controls on the sampling of the terrestrial fossil record: a case study from the Dinosauria. In *Comparing the geological and fossil records: implications for biodiversity studies*, vol. 358 (eds A. J. McGowan & A. B. Smith), pp. 209–240. London, UK: Geological Society of London Special Publications.
- Peters, S. E. & Heim, N. A. 2010 The geological completeness of paleontological sampling in North America. *Paleobiology* **36**, 61–79. (doi:10.1666/0094-8373-36.1.61)
- McKinney, M. L. 1990 Classifying and analysing evolutionary trends. In *Evolutionary trends* (ed. K. J. McNamara), pp. 28–58. London, UK: Belhaven.
- Friedman, J. H. 1991 Multivariate adaptive regression splines. *Ann. Stat.* **19**, 1–67. (doi:10.1214/aos/1176347963)

- 26 Milborrow, S. 2011 Earth: multivariate adaptive regression spline models. R package v. 2.6–2. (derived from mda:mars by Trevor Hastie and Rob Tibshirani). See <http://CRAN.R-project.org/package=earth>
- 27 Holland, S. M. 2011 Sea level change and the area of shallow-marine habitat: implications for marine biodiversity. *Paleobiology* **38**, 187–204.
- 28 McGowan, A. J. & Smith, A. B. 2008 Are global Phanerozoic marine diversity curves truly global? A study of the relationship between regional rock records and global Phanerozoic marine diversity. *Paleobiology* **34**, 80–103. (doi:10.1666/07019.1)
- 29 Erwin, D. H. 2009 A call to the custodians of deep time. *Nature* **462**, 282–283. (doi:10.1038/462282a)
- 30 Smith, A. B., Gale, A. S. & Monks, N. E. A. 2001 Sea-level change and rock record bias in the Cretaceous: a problem for extinction and biodiversity studies. *Paleobiology* **27**, 241–253. (doi:10.1666/0094-8373(2001)027<0241:SLCARR>2.0.CO;2)