Sampling bias and the fossil record of planktonic foraminifera on land and in the deep sea

Author(s): Graeme T. Lloyd, Paul N. Pearson, Jeremy R. Young, and Andrew B. Smith
Source: Paleobiology, 38(4):569-584.
Published By: The Paleontological Society
https://doi.org/10.1666/11041.1
URL: http://www.bioone.org/doi/full/10.1666/11041.1

Paleobiology, 38(4), 2012, pp. 569–584

Sampling bias and the fossil record of planktonic foraminifera on land and in the deep sea

Graeme T. Lloyd, Paul N. Pearson, Jeremy R. Young, and Andrew B. Smith

Abstract.—Large-scale trends in planktonic foraminiferal diversity have so far been based on utilization of synoptic biostratigraphic range charts. Although this approach ensures the taxonomic consistency and quality of the data being used, it takes no formal account of any sampling biases that might exist in the fossil record. We demonstrate that the occurrence data of planktonic foraminifera, as recorded in the primary literature, are strongly biased by sampling. We do this by demonstrating that raw diversity curves derived from the land-based and deep-sea records are strikingly different, but that they each correlate with the intensity of sampling in their respective environments, and thus are ultimately controlled by the structure of the geological record in each setting. Because sampling of the Mesozoic record is best in our land record whereas sampling of the Cenozoic is best in our deep-sea record, we combine the two to generate the best-supported estimates of species and genus diversity over time from these data. We correct for sampling bias using shareholder quorum subsampling and a modeling approach. The data are then transformed to generate a range-through plot of species richness that is compared with two earlier estimates of the diversity history where comparable species-in-bin data can be recovered. No robust statistical correlation is found among the three estimates. Although differences in amplitude are to be expected, differences in the actual shape of the curve are surprising. We conclude that these differences stem from the nature of the data themselves, namely the taxonomic scheme adopted and the taxonomic coverage used.

Graeme T. Lloyd. Department of Palaeontology, The Natural History Museum, Cromwell Road, London SW7 5BD, United Kingdom
Paul N. Pearson. School of Earth and Ocean Sciences, University of Cardiff, Park Place, Cardiff CF10 3AT, United Kingdom
Jeremy R. Young. Department of Earth Sciences, University College London, Gower Street, London, WC1E 6BT, United Kingdom
Andrew B. Smith. Department of Palaeontology, The Natural History Museum, Cromwell Road, London SW7 5BD, United Kingdom. E-mail: a.smith@nhm.ac.uk. *Corresponding author*

Accepted: 26 March 2012
Supplemental materials deposited at Dryad: doi 10.5061/dryad.8ts3p

Introduction

Planktonic foraminifera occur in large numbers in the world's oceans and their calcite tests are often a significant and sometimes dominant component of deep-sea sediment. They are used routinely for biostratigraphy—the Cenozoic for example is divided into 49 zones and many subzones based on their occurrence (Wade et al. 2011). This record can be remarkably complete in individual sections and the analysis of foraminifera recovered from cores of near continuous deep-sea sedimentation provides very high resolution signatures of how life in the oceans responded to climatic and oceanographic perturbations (Ezard et al. 2011). Consequently, the planktonic foraminifera are one of the most important groups for studying global change and paleoceanographic evolution (see, for example, the recent review by Aze et al. [2011]). Planktonic foraminifera are passively dispersed by ocean currents, resulting in very low endemism in morphospecies between ocean basins (Bé 1982; Hemleben et al. 1989). There is, however, a strong latitudinal / temperature control on species distributions (Bé 1982) and also considerable variability between oligotrophic environments such as the gyres, where diversity is greatest, and upwelling areas, where a few specialist forms predominate (Hemleben et al. 1989; Arnold and Parker 2003).

This excellent record has been summarized on several occasions in an attempt to generate global diversity curves for the group through time (e.g., Tappan and Loeblich 1973; Loeblich and Tappan 1988; Norris 1991; Hart et al. 2002; Ezard et al. 2011). Such compilations have been made from the selective synthesis of the first and last occurrence datums of each taxon

 0094-8373/12/3804-0005/$1.00

in the biochronologic record rather than from a comprehensive summary of the primary occurrence data, such as are published in the records of the Deep Sea Drilling Project (DSDP; http://www.deepseadrilling.org/) and Ocean Drilling Program (ODP; http://www-odp.tamu.edu/), or available through the Neptune database (Spencer-Cervato 1999; Lazarus 2011). None of these approaches, therefore, have attempted to formally take account of variation in sampling effort over time. Nonetheless, we know that variation in our sampling of the rock record can have a profound effect on the resultant diversity recorded both for macrofossils (Alroy et al. 2008) and microfossils (Lloyd et al. 2012a,b).

On land, sampling intensity is controlled by the interplay of two factors: how much rock survives from each period, and how much interest there is in each period. Planktonic foraminifera live in open water and are most often recovered from offshore fine-grained sediments (marls, chalks, and mudrocks). The preservation of these lithofacies over the continental blocks is primarily controlled by major sea level cycles, with good records of offshore sediments preserved at times of global highstand and with the record of offshore, deep-water marls and chalks diminishing markedly in frequency and abundance toward the Recent (Lloyd et al. 2012b), as sea levels fell globally (Miller et al. 2005).

In the deep sea the situation is very different, as our knowledge of the fossil record there is controlled by the distribution of deep-sea drilling sites. Because of plate tectonics and ocean floor spreading, older rocks are less common and more restricted in their distribution, and progressively more difficult to access than younger rocks. Furthermore, because it is usually not possible to drill older sediments without also drilling through younger sediments, many DSDP/ODP holes recover Neogene sediments even if they are not the primary target. As a consequence the sampling record in the deep sea increases greatly toward the Recent (Lloyd et al. 2011). Thus, the nature of the geological record in each environment directly shapes sampling opportunity. Whereas the strength of the land-based record lies in its Mesozoic record, the strength of the deep-sea record is in the Cenozoic.

Our goal here is to explore the long-term diversity signal that can be recovered from the primary occurrence data of planktonic foraminifera, and to discover to what extent this signal is distorted by sampling bias. We do this by (1) compiling two independent estimates of diversity, one from land-based records, the other from deep-sea records; (2) correcting each for the specific sampling bias; and (3) combining the two data sets to derive our best estimate of planktonic foraminiferal diversity over time. We then compare our results with previous estimates.

Materials and Methods

In this study we limited ourselves to analysis of the fossil record of the North Atlantic region (which we define as 90°N to 20°S and including both the Mediterranean and Caribbean), and its bordering continental regions (eastern North America, the Caribbean islands, western Europe, and north and northwestern Africa) (Fig. 1). This area covers the full latitudinal range, a wide variety of trophic conditions, and the greatest concentration of researcher effort, especially for the Mesozoic part of the record. Species occurrences were entered into a relational database, with taxonomic occurrences assigned to time units that represent a nannofossil or planktonic foraminiferal biozone in either a DSDP/ODP hole or land-based locality (see Lloyd et al. 2011 for details of the database architecture). Deep-sea sites (comprising single or multiple holes) and land-based localities serve as our sampling proxy. This choice was mediated by the requirement of an approximately equivalent proxy between the two (land and deep-sea) records with measurements such as geologic formations or number of maps, both utilized by previous workers (Peters and Foote 2001; Smith 2001; Crampton et al. 2003; Smith and McGowan 2007; Benson et al. 2010; Manion et al. 2011), being inappropriate for the deep sea. Records of planktonic foraminifera from deep-sea sediments were compiled using a mixture of parsing from preexisting databases (NEPTUNE: http://paleodb.org/cgi-bin/bridge.

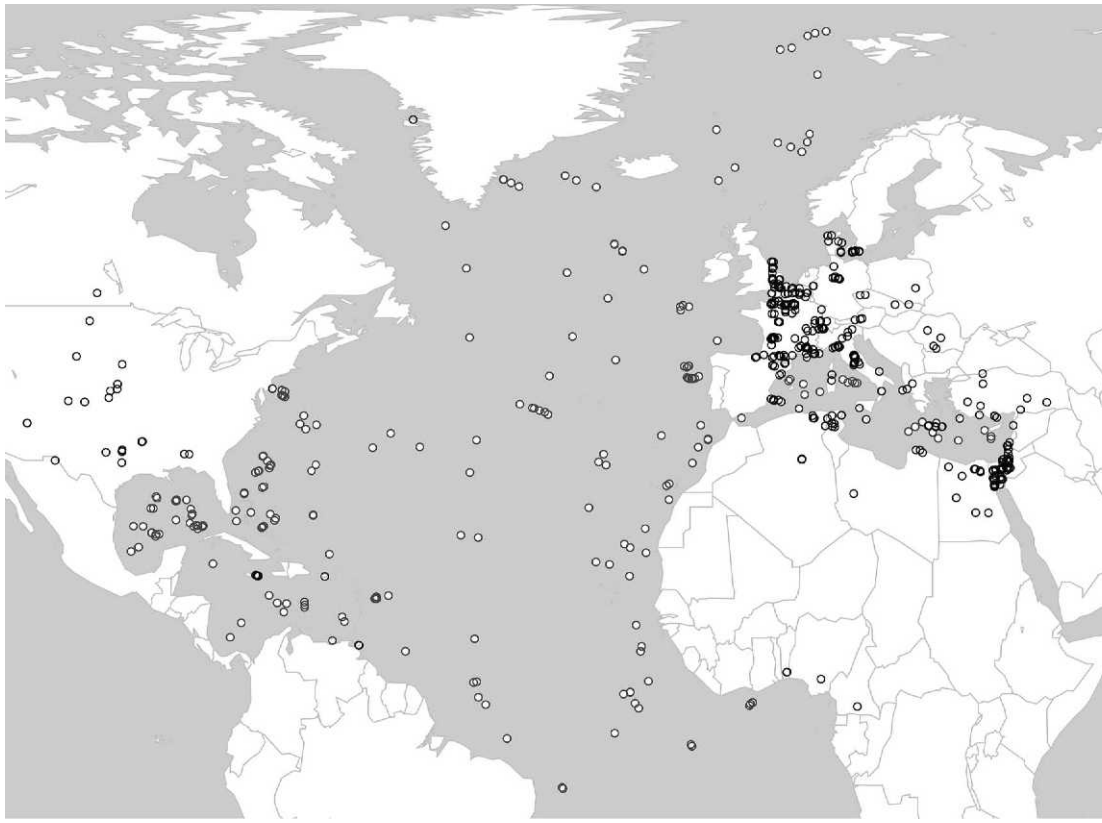

FIGURE 1. Map showing the distribution of sites from which data was collected for this study.

pl?a=displayDownloadNeptuneForm, CHRONOS: http://www.chronos.org and ODSN: http://www.odsn.de) and manual entry, whereas land-based data entry was entirely manual from a survey of published papers since 1970. Land-based localities were included only if they provided a measured section with numbers of samples taken and a full taxonomic list of the species recovered: papers reporting only selected biostratigraphically useful species were not included. The data were filtered for previous data entry errors by using a range-through list (Andrew Fraass unpublished data) and expert observation (P.N.P.) to identify occurrences that lay far outside of the expected range. The majority of these outliers were found to be erroneous, bearing no resemblance to the published distribution chart. Errors were overwhelmingly concentrated in the NEPTUNE data and likely reflect the lengthy and complex history of this database (Lazarus 2011); these were corrected whenever identified. However, we have not attempted to verify the validity of all the taxonomic records, a task that would require study of the original specimens. In total 28,589 occurrences (19,078 deep-sea and 9511 land-based) of planktonic foraminifera were entered, comprising 1708 species names in total. The database was then taxonomically standardized to remove known synonyms using Cenozoic "atlases" (Kennett and Srinivasan 1983; Hemleben et al. 1989; Olsson et al. 1999; Pearson et al. 2006) and expert knowledge (P.N.P.), which left a total of 849 "valid" species (i.e,. species names we have no reason to discount). In cases where continuing disagreement exists among taxonomists over whether a taxon should be split or lumped, proposed species were generally counted as valid. Invalid, indeterminate, and questionable occurrences were removed prior to analysis and junior synonyms were replaced by their senior counterpart. We considered only fossil occurrences of extant taxa, to avoid the "Pull of the Recent" artifact (Raup 1979;

Jablonski et al. 2003): our last time bin did not include the present-day protist biota except where they are preserved as fossils. However, almost all extant species were represented in the Pliocene record.

Biozones were combined to create time bins of approximately equal duration, ensuring that major geologic boundaries (the Jurassic/Cretaceous, Cretaceous/Paleogene, Eocene/Oligocene, and Paleogene/Neogene) coincided with time-bin boundaries. These intervals are associated with major turnover events and any bin spanning such a boundary would have had artificially inflated diversity because of an extinction and recovery flora being time-averaged together. Creation of the time bins followed Lloyd et al. (2011) and proceeded in the following manner. First, a requested bin length is submitted, here 6 Myr. Then, starting from the present (0 Ma) the nearest planktonic foraminiferal and/or coccolithophore zone boundary (actually sampled within the database) to each increment (here 6 Ma, 12 Ma, etc.) is recorded. An additional step inserts major boundaries (Paleogene/Neogene, Eocene/Oligocene, Cretaceous/Paleogene, Jurassic/Cretaceous, and Triassic/Jurassic), replacing the closest value previously recorded if not already present. Table 1 provides details of the binning scheme used here. Actual mean bin length is 5.54 Myr with standard deviation 2.79 Myr. Lists of the species that we accept as valid and which have definitively been recorded as present in each 6-Myr time bin are provided as electronic supplementary data (Supplementary Tables 1–3).

Even after clumping zones together, some units lacked the precise dating required to assign them to a single time bin. Such units were assigned to time bins using the randomization procedure of Lloyd et al. (2011, 2012b) with the process repeated 1000 times to obtain mean and 95% confidence values. To assess whether the time bin duration made a significant difference to our results, we repeated the procedure using both 3-Myr and 6-Myr time bins. Raw diversity curves were produced using the sampled-in-bin approach (Miller and Foote 1996).

Sampling-corrected curves were produced using single-publication occurrence corrected shareholder quorum subsampling (SQS: Alroy 2010a,b,c). Alroy's updated R function (version 2.0, code dated 14 February 2011) was modified to include the single-publication occurrence correction omitted in his R code but included in his Perl code. The function is available directly as sqs() from http://www.graemetlloyd.com/pubdata/functions_2.r. SQS is a novel subsampling approach that takes the area under a frequency curve of species occurrences to estimate subsampled richness. We chose SQS over traditional rarefaction by occurrences because it offers two major benefits. Firstly, traditional rarefaction tends to be highly conservative and often gives flat diversity profiles (e.g., Lloyd et al. 2008) that likely represent a false negative rather than any hidden underlying diversity signal. Very large sample sizes are required to overcome this problem, and with the patchy nature of most fossil records this is not possible for all or even most time bins. Second, a major problem with rarefaction curves is their frequent tendency to cross each other. In practice this means that, for any two bins with crossing rarefaction curves, which one is more diverse depends on the sampling level at which the comparison is made. This occurs far less with SQS curves, meaning relative bin-to-bin comparisons tend to be consistent regardless of the value of q used. (For an example, see text-Fig. 3 of Alroy 2010b.)

We also applied the modeling approach of Smith and McGowan (2007) as an additional estimate for sampling-corrected species richness and developed confidence intervals for these models following Lloyd (2012). This technique builds on the general principle of Raup (1972), which assumes that true diversity is constant and observed diversity is purely a function of sampling. First we developed a model that predicts what sampled diversity would look like if true diversity were uniform over time and proportional to sampling intensity. In practice this means aligning the most sampled datum with the highest richness datum, the second highest with the second highest, and so on. Some model (a linear one in the case of Smith and McGowan 2007 and various in the case of Lloyd 2012) is then fitted, and can be expressed as an equation. Inserting

TABLE 1. Time bins used in this study and the zonal boundaries that were used to define their start. CRZ = concurrent-range Zone; LOZ = lowest-occurrence Zone; TRZ = taxon-range Zone.

Time bin	Age at base (Ma)	Nannoplankton zone defining start	Planktonic foraminiferal zone defining start
1	5.12	base NN13	
2	11.63		base N14/M11
3	16.97		base N8/M5
4	21.12		base N5/M2
5	22.96		base N4/M1
6	29.45		base N2/P21/O5/Globigerina angulisuturalis-Chiloguembelina cubensis CRZ
7	34.00		base P18
8	36.95	base CP15	
9	39.80		base P13/E12/Orbulinoides beckmanni TRZ
10	48.60		base P10/E8/Guembelitrioides nuttalli LOZ
11	55.73		base E1/Acarinina sibaiyaensis LOZ
12	58.32	base NP7/CP6	
14	65.50	base NP1/CP1	base P0
15	71.12	base CC23/UC16	
16	83.53	base CC18/NC18	base Globotruncanita elevata Zone
17	85.85	base CC15	
18	89.63	base CC13/NC15	
19	93.41	base UC6	base Helvetoglobotruncana helvetica Zone
20	99.60		base Rotalipora globotruncanoides Zone
21	112.02	base NC8	
22	125.11	base NC6/CC7	
23	129.64		base Hedbergella similis Zone
24	134.38	base CC4	
25	140.24	base CC3/NC3	
26	145.46		base Conoglobigerina gulekhensis Zone
27	150.63	base NJ16b	
28	156.30	base NJ15b	
29	161.61	base NJ14	
30	164.25	base NJ12b	
31	168.24	base NJ11	
32	171.15	base NJ9	
33	174.18	base NJ8b	
34	182.63	base NJ6	
35	189.07	base NJ4a	
36	196.69	base NJ2a	
37	199.60	base NJ1	

Table 2. Correlations between deep sea and land-based measures of sampling and taxonomic richness. Values in bold indicate statistically significant correlations.

		Raw data		Generalized differences	
	Comparison:	ρ	*p*	ρ	*p*
Sampling	Land-based vs. deep-sea	0.13	0.54	0.03	0.88
Species	Land-based richness vs. deep-sea richness (Total)	0.38	0.06	0.19	0.37
	Land-based richness vs. deep-sea richness (Mesozoic)	0.85	**<0.01**	–0.15	0.62
	Land-based richness vs. deep-sea richness (Cenozoic)	–0.60	0.06	0.30	0.41
	Land-based richness vs. land sampling	0.90	**<0.01**	0.73	**<0.01**
	Land-based richness vs. deep-sea sampling	0.12	0.56	–0.10	0.64
	Deep-sea richness vs. deep-sea sampling	0.83	**<0.01**	0.46	**0.02**
	Deep-sea richness vs. land sampling	0.29	0.17	0.24	0.26
Genera	Land-based richness vs. deep-sea richness (Total)	0.57	**<0.01**	0.41	0.05
	Land-based richness vs. deep-sea richness (Mesozoic)	0.90	**<0.01**	–0.07	0.82
	Land-based richness vs. deep-sea richness (Cenozoic)	–0.15	0.67	0.49	0.15
	Land-based richness vs. land sampling	0.84	**<0.01**	0.71	**<0.01**
	Land-based richness vs. deep-sea sampling	0.29	0.17	–0.18	0.39
	Deep-sea richness vs. deep-sea sampling	0.78	**<0.01**	0.22	0.29
	Deep-sea richness vs. land-based sampling	0.32	0.12	0.30	0.15

TABLE 3. Correlations between sampling corrected measures of taxonomic richness. SQS = shareholder quorum subsampling; MR = Residuals from model comparison. Values in bold indicate statistically significant correlations.

		Raw data		Generalized differences	
Richness comparison:		ρ	*p*	ρ	*p*
Species	SQS land-based vs. MR land-based	0.79	**<0.01**	0.82	**<0.01**
	SQS deep-sea vs. MR deep-sea	0.79	**<0.01**	0.75	**<0.01**
	SQS combined vs. MR combined	0.70	**<0.01**	0.60	**<0.01**
	SQS land-based vs. SQS deep-sea	0.33	0.17	0.39	0.12
	MR land-based vs. MR deep-sea	0.30	0.15	0.42	**0.04**
Genera	SQS land-based vs. MR land-based	0.55	**<0.01**	0.33	0.13
	SQS deep-sea vs. MR deep-sea	0.35	0.12	0.77	**<0.01**
	SQS combined vs. MR combined	<0.01	0.98	0.27	0.23
	SQS land-based vs. SQS deep-sea	0.67	**<0.01**	0.38	0.10
	MR land-based vs. MR deep-sea	0.07	0.75	0.45	**0.03**

the sampling proxy into this equation thus returns predicted diversity values that may be subtracted from observed diversity to give residuals that reflect higher or lower than expected taxonomic richness. Subtracting the model predictions from the observed data thus reveals how much of the observed diversity remains unexplained by sampling and thus potentially biological in origin.

For our land rock-record proxies we used the data on outcrop area for western Europe taken from the supplementary data in Smith and McGowan (2007) and the total number of sedimentary rock sections for North America as detailed in Peters (2005).

Correlations are Spearman rank and differencing is generalized differencing (McKinney 1990). All analyses were performed in the freely available statistical programming language R (R Development Core Team 2010) and the code used is available from G.T.L. on request.

Results

Trends in Raw Diversity

Species and genus diversity, as recorded from deep-sea sediments (Fig. 2A,B), rise through the Cretaceous and Cenozoic, reaching a maximum peak in the Neogene, but with significant dips after both the uppermost Cretaceous and upper Eocene. Note that the number of species recorded in the last time bin is significantly greater than the current estimated diversity of planktonic foraminifera primarily because of binning and the high species turnover that took place during this final 6-Myr interval. By contrast, species and genus diversity as recorded by our land-based records (Fig. 2D,E) begins in the Middle Jurassic, rises steeply to an overall peak value in the last bin of the Cretaceous, then declines through the Cenozoic, with marked drops after the uppermost Cretaceous and upper Eocene. At face value, therefore, the two records point to very different diversification histories. Despite the apparent disparity between these two records the raw genus data are significantly positively correlated (Table 2). However, this is driven by the Mesozoic portion of the data; the Cenozoic portion is borderline significantly negatively correlated, and all correlations are nonsignificant when the data are differenced (Table 2).

Sampling Proxy Results

The variation in our sampling proxy—the number of sites or localities that have been sampled for planktonic foraminifera—is also shown in Figure 2. The deep-sea record (Fig. 2C) has the overall shape of an exponential curve, although this is largely due to the Neogene's steep rise far above Mesozoic and Paleogene values. The land-based record (Fig. 2F) reaches an overall peak value in the Maastrichtian and declines rapidly after the Cretaceous. Correlation of these two time series is not significant (Table 2).

Correlations between Sampling and Richness

To test for a sampling bias we correlated our richness data with our sampling proxy data (Table 2). For the raw species data all combinations are positive and significant, although richness correlates most strongly with its

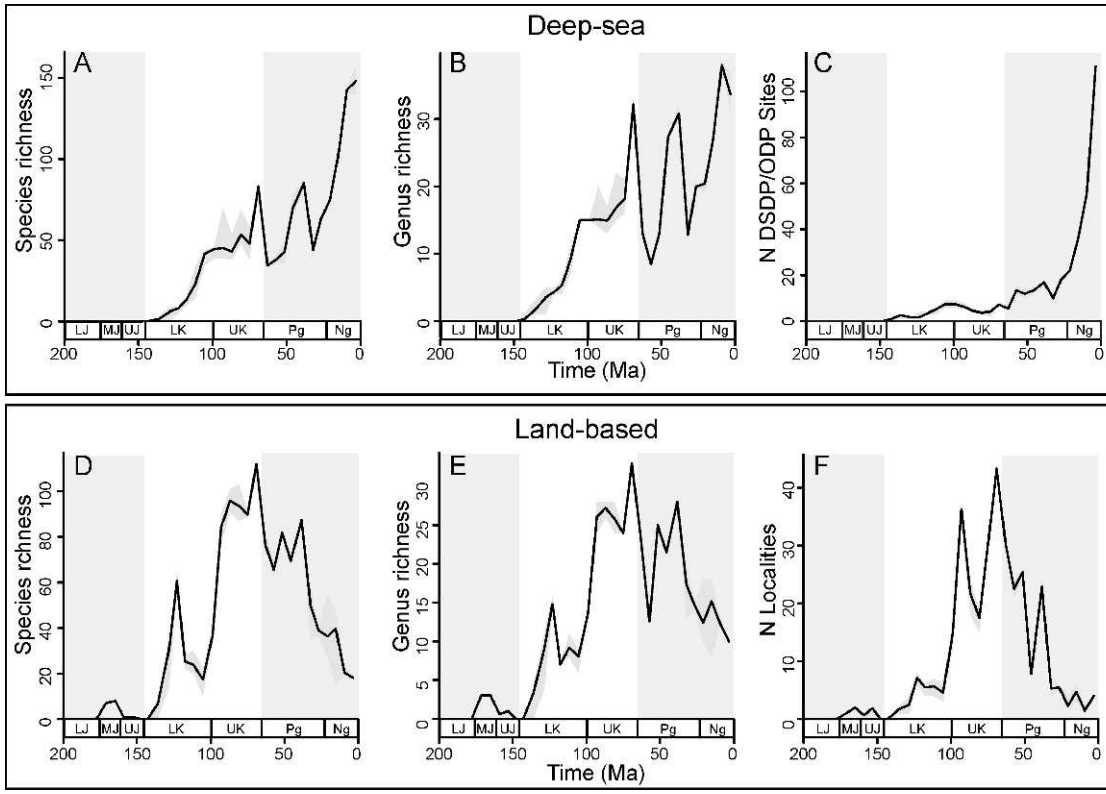

FIGURE 2. Comparisons of species richness (A, D), genus richness (B, E), and sampling (C, F). for deep-sea and land-based records plotted in 6-Myr time bins over the last 200 Myr. Sampling data are the numbers of DSDP/ODP sites or land-based localities from which the planktonic foraminiferal lists in our database have been collated. Geological epochs are as follows: LJ = Lower Jurassic, MJ = Middle Jurassic, UJ = Upper Jurassic, LK = Lower Cretaceous, UK = Upper Cretaceous, Pg = Paleogene and, Ng = Neogene. Jurassic and Cenozoic intervals shaded.

respective sampling record. Differencing the data further emphasizes this, as correlations between richness and sampling *between* the two records becomes nonsignificant, and even negative, whereas correlation *within* records remains significant and positive. The raw genus level data are also all significantly correlated, but this time only the *within* land-based correlation remains significant after differencing. Overall correlation between richness and sampling is strongest within the land-based record (Table 2), suggesting a stronger bias here than in the deep sea.

Sampling-Corrected Richness

Shareholder Quorum Subsampling.—Species richness following SQS is shown in Figure 3A, D, G, and J. (Throughout we present results using q, the sampling quorum, of 0.4; although very similar results are obtained for other values of q, larger values often resulted in missing data for some time bins.) The deep-sea records of species and genus richness are very similar (Fig. 3A,G), with richness rising with slight fluctuations to an overall Maastrichtian peak, before the largest bin-to-bin fall across the Cretaceous/Cenozoic boundary. After this drop richness again rises, although more markedly in the genus data than at the species level. The land-based record (Fig. 3D,J) is broadly similar but with maximum richness not reached until the Cenozoic, and with a less pronounced drop after the end-Cretaceous. Whereas the deep-sea and land-based species richness records are uncorrelated for both the raw and differenced data (Table 3), the two genus-level estimates appear more congruent, and are significantly correlated (Table 3), although only for raw data.

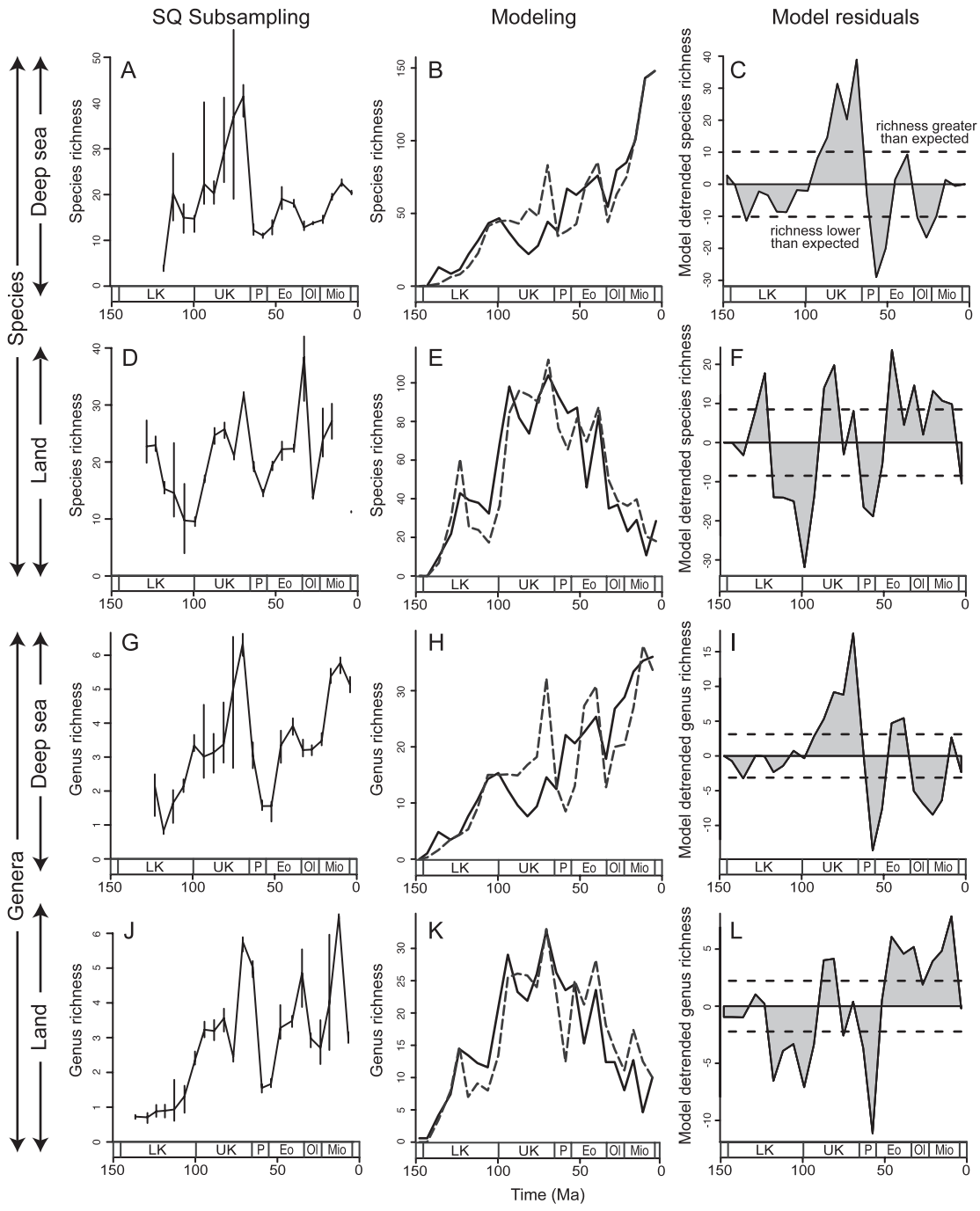

FIGURE 3. Estimates of planktonic foraminiferal species and genus richness derived from deep-sea and land-based sedimentary records. A, D, G, J, Shareholder quorum subsampling (Alroy 2010a) with single publication occurrence correction ($q = 0.4$ in all cases). Vertical bars indicate 95% CI based on 1000 randomizations where uncertainly dated units were assigned to a single time bin. B, E, H, K, Richness modeled assuming true diversity is constant and proportional to rock record sampled (Lloyd 2011). Dashed line, empirical data; solid line, model prediction. C, F, I, L, Residuals obtained by removing modeled from empirical species richness. Dashed lines mark 95% confidence estimates of model. Geological epochs are as follows: LJ = Lower Jurassic, MJ = Middle Jurassic, UJ = Upper Jurassic, LK = Lower Cretaceous, UK = Upper Cretaceous, P = Paleocene, Eo = Eocene, Ol = Oligocene, Mi = Miocene.

Modeling.—By assuming that true richness is constant and the observed richness is driven purely by sampling, we constructed model estimates for apparent species and genus richness over time that closely match what is observed in the empirical data (Fig. 3B,E,H,K). However, the fit is not perfect and several divergences between modeled and empirical data are evident. Subtraction of the model predictions from the observed data indicates where species richness is greater or less than predicted from sampling and allows an independent estimate for sampling-corrected richness. For deep-sea species there are four time intervals of note where our sampling proxy alone cannot explain observed richness (Fig. 3C,I). These are a richness low in the Valanginian, a prolonged high in the Upper Cretaceous (including an overall Maastrichtian peak), a richness low in the upper Paleocene–lower Eocene, and finally a low in the Oligocene. The land-based record again offers a different picture (Fig. 3F,L) with a prolonged low in the mid-Cretaceous, a high in the Santonian–Campanian, a Paleocene low, and a prolonged period of elevated richness in the Eocene–Miocene. Land and deep-sea curves do follow similar trajectories and, although raw comparison of the two time series indicates no significant correlation, both species and genus level curves are significantly correlated after differencing (Table 3).

Pooled Data.—Our land and deep-sea data sets are strongest at different times, with the Mesozoic record more extensively sampled in the former and the Cenozoic record better sampled in the latter. We therefore pooled our land-based and deep-sea data together and repeated the same sampling-correcting methods as above, to derive a combined estimate for planktonic foraminifera richness. Raw species and genus richness (Fig. 4E,F) both rise through the Cretaceous to a late Cretaceous peak before dropping at the Cretaceous/Paleogene boundary. There is a smaller mid-Eocene peak and an upper Eocene through Oligocene trough before sampled diversity rises again. Highest diversity is reached in the Plio-Pleistocene (~150 species) but is only slightly higher than the late Cretaceous peak (~135 species).

The SQS curves are shown in Figure 4A and B. The species curve begins with a steep rise in SQS richness to a mid-Lower Cretaceous high before a substantial fall within the Aptian. There is then a steady climb through the rest of the Cretaceous to an overall peak in the Upper Cretaceous (Maastrichtian). SQS richness falls across the Cretaceous/Paleogene boundary and remains rather depressed thereafter. The genus-level results are similar, but with a complete absence of the Barremian–Aptian high and an overall steeper rise in the Cenozoic, with the richness peak in the upper Miocene of similar magnitude to the Maastrichtian one.

Modeling estimates of richness are shown alongside the empirical curves in Figure 4E and F. Here there are several notable discrepancies that are best revealed by subtraction of the modeling data from the empirical data (Fig. 4C,D). Species and genus data show a very similar distribution of residuals with a high in the mid-Cretaceous, a low through the Albian–Cenomanian, a high and overall peak in the Santonian–Campanian, a maximal low in the Paleocene, and a high in the middle Eocene. A Plio-Pleistocene low is also apparent in the genus-level data.

Correlations between the resulting time series of the two sampling-correcting methods show much greater congruence at the species than the genus level, with the former being significant for both raw and differenced data and the latter being nonsignificant in both cases (Table 3).

Discussion

Comparison with Previous Diversity Estimates

Because the raw diversity curves for land and deep-sea planktonic foraminifera diverge so markedly after the Oligocene, it is clear that simple counts of the numbers of taxa being recorded in any one time interval in our database are heavily dependent upon sampling, itself constrained by the structure of the geological records preserved in those respective environments. These differences in sampling intensity over time need to be removed to derive a less biased estimate of diversity.

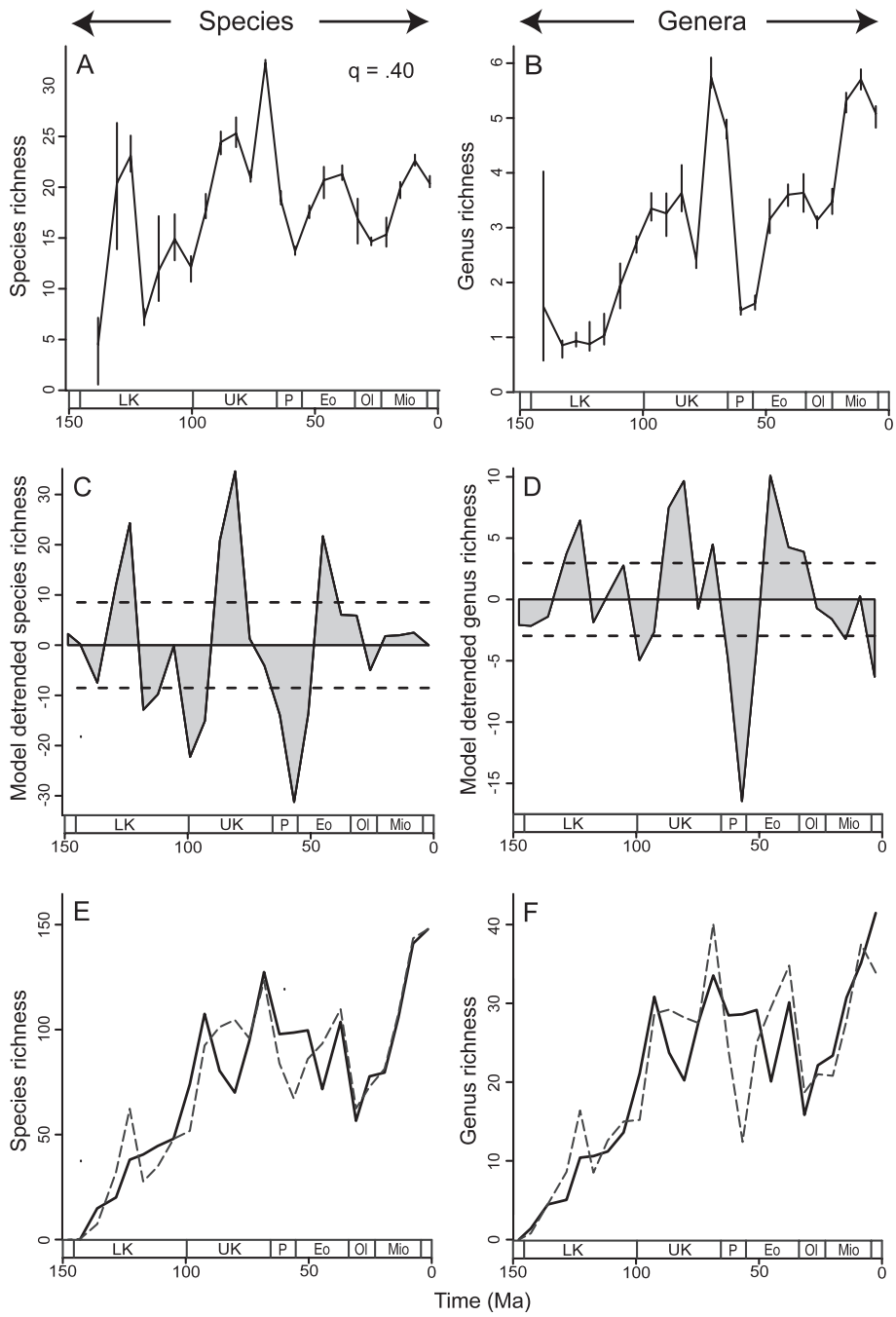

FIGURE 4. Estimates of planktonic foraminiferal taxonomic richness derived from combined deep-sea and land-based occurrence lists. A, B, Species and genus richness from shareholder quorum subsampling (Alroy 2010a) with single publication occurrence correction ($q = 0.4$ in all cases). Vertical bars as in Figure 2. C, D, Residuals obtained by removing modeled from empirical species and genus richness. Dashed lines mark 95% confidence estimates. E, F, Species and genus richness modeled assuming true diversity is constant and proportional to rock record sampled (Lloyd 2011). Dashed line, empirical data; solid line, model prediction. Geological epochs as in Figure 3.

Consequently, our SQS analysis of the combined land and deep-sea records (Fig. 4A) is taken as our best estimate of long-term trends in planktonic foraminifera species diversity from these data. Independent support for this comes from our modeling approach (Fig. 4C), which points to a statistically very similar result.

The curves presented here are not the first estimates of planktonic foraminifera richness. Tappan and Loeblich (1988: Fig. 22) presented one of the earliest diversity curves for the group. Broadly similar features can be noted here, such as a Mesozoic rise toward a Maastrichtian peak, followed by a major dip and a second rise to a mid-Eocene high, then on to an Oligocene low before a final peak in the Miocene. However, any differences could be partially accounted for purely by the differing time-binning used (they applied geologic stages). Consequently, we sought data sets containing first and last occurrence data that could be assigned to our 3- or 6-Myr bins for quantitative comparison. (No sampled-in-bin curves have been produced for planktonic foraminifera as far as we are aware.) We found two candidate data sets: the online Plankrange database (Stewart and Pearson 2000) and the recent compilation of Aze et al. (2011). Both are global in their scope. The Plankrange database is based on an unpublished Ph.D. thesis by D. R. M. Stewart in which species ranges were compiled from 33 published sources listing preexisting unified taxonomies drawn up by specialists between 1957 and 1997. It contains first and last occurrence data on some 600 species drawn from both deep-sea cores and land-based sections. The Aze et al. database is limited to the Cenozoic and includes only macroperforate planktonic foraminifera. However, because macroperforate taxa dominate the curve, we consider the data to still be broadly comparable. An important difference from our database is that the Aze et al. compilation includes only occurrences that one of us (P.N.P., based on personal experience) was confident represented valid species determinations and correct ranges. Here, by contrast, we have included any species formally erected that was not obviously invalid or a known junior synonym and all firm occurrences unless specifically identified as reworked. The Aze et al. database thus contains information on just 297 Cenozoic species (79% of those included here). For comparative purposes our pooled data were used to produce range-through curves and the results are plotted in Figure 5 as simple range-through curves using the same 6-Myr time bins.

It is immediately clear that the curves show marked differences (Fig. 5), and this is borne out by correlation (Supplementary Table 4). Our data generally gives higher richness estimates than the older Plankrange data set, presumably due to two factors: (1) the extension of ranges of shared species and (2) the addition of some 250 valid species. The same is true when comparing Aze et al. (2011) with our data, suggesting that our less stringent criteria for including species may account for some differences. In two places (during the lower and upper Eocene) the Aze et al. data records higher diversity, possibly because it is a global database whereas our data are drawn exclusively from the Atlantic region. However, although differences in species richness are to be expected from the three databases, a significant difference in the shape of these three time series is striking and unexpected. Significant levels of correlation are found only between the raw (and probably autocorrelated) range-through data and the SQS estimate from those same data. Other workers comparing successive richness estimates (Sepkoski 1993; Alroy 2000) have largely found an increase in numbers only, and not the significant differences in shape shown here.

Detailed examination of the three curves (Fig. 5) reveals a number of discrepancies. Perhaps most marked of these is the progressive rise from the Oligocene to the present in our data that is absent in both Aze et al. (2011) and Plankrange. Although this is partly ameliorated once sampling is accounted for through SQS (solid black line, Fig. 5), there is still little congruence after about 45 Ma, with a small rise in the upper Miocene in our data set against a steep drop in both the Aze et al. (2011) and Plankrange curves. Similarly, the shareholder quorum estimate predicts that Cenozoic richness reached its peak in the upper Miocene, whereas Aze et al. (2011) peaks in the Pliocene (like our raw combined

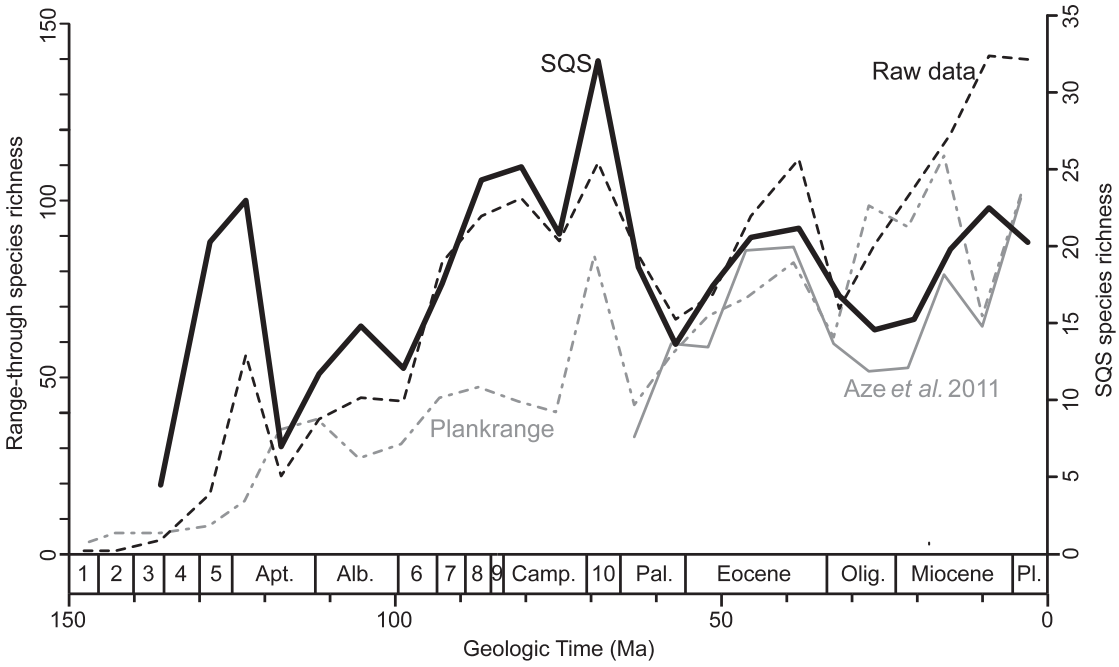

FIGURE 5. Comparison of the species data presented here with two synoptic (range-through) species curves. The dashed black line is a range-through curve created from the pooled data presented here. The solid line is the shareholder quorum subsampling (Alroy 2010a) species richness estimate from Figure 3A. The solid gray line is from a recent list of Cenozoic-only macroperforate planktonic foraminifera (Aze et al. 2011). The dashed gray line is from a more comprehensive, but older compilation of all planktonic foraminfera (Stewart and Pearson 2000). Numbering and abbreviations for stages and epochs as follows: 1, Tithonian; 2, Berriasian; 3, Valanginian; 4, Hauterivian; 5, Barremian; 6, Cenomanian; 7, Turonian; 8, Coniacian; 9, Santonian; 10, Maastrichtian; Apt, Aptian; Alb, Albian; Camp, Campanian; Pal, Paleocene; Olig, Oligocene; Pl, Plio-Pleistocene.

data) and Plankrange in the middle Miocene. Conversely both the raw and sampling-corrected Cenozoic diversity lows are in the upper Paleocene, whereas in the Aze et al. (2011) and Plankrange curves the low occurs in the lower–middle Paleocene.

What then causes these differences? A sampling bias, at least insofar as it is measured here, does not appear to explain the difference, as in all cases using the shareholder quorum estimate actually weakens the correlations (Supplementary Table 4), showing a significant fit only to the raw data presented here. Nor is this primarily a problem of binning, as correlation remains poor irrespective of whether 6-Myr, 3-Myr, 2-Myr, or 1-Myr binning strategies are used (data not shown). Consequently, it seems most likely that the difference is attributable to the nature of the data themselves and to the taxonomic rigor that has gone into their compilation. Recent taxonomic revisions (such as that of Pearson et al. 2006) can dramatically affect such data, and this may largely explain the difference between the Aze et al. (2011) data and our data on the one hand and the Plankrange data on the other. For example, the major difference between Plankrange and Aze et al. plots seen in the Oligocene and lower Miocene may arise from a monographic effect. Plankrange uses the more finely split taxonomy of Spezzaferri (1994), whereas Aze et al. (2011) and our analysis, based on more recent work, used a much more conservative taxonomy and consequently recorded lower diversity.

The difference between our results and those of Aze et al. (2011) arises for a different reason related to the taxonomic inclusivity of the two analyses. Our database has 375 Cenozoic species whereas the Aze et al. database has 297. Of the 78 additional taxa in our data set 54 (69%) are microperforate foraminifera, a group excluded from the Aze et al. analysis. Microperforate species show a

different diversity trajectory through the Cenozoic, and when we remove them from our analyses, correlation between the two data sets improves markedly (raw data: Spearman rho = 0.61, $p = 0.05$; generalized differences: Spearman rho = 0.53, $p = 0.12$). When we further reduce our data to include only those taxa in common with the Aze et al. (2011) data set, the resultant correlation is higher still (raw data: 0.87, $p < 0.01$; generalized differences: 0.79, $p < 0.01$). Although not perfect it suggests that the difference between the two diversity estimates comes predominantly from differences in the taxa included, and not from differences in their recorded ranges.

Our data set is more taxonomically inclusive than either Plankrange or the Aze et al. compilation and consequently encompasses a greater diversity of taxonomic opinion. Diversity of taxonomic opinion in our data almost certainly increases as the record gets better through the Cenozoic and the number of individual papers documenting the fossil record increases (Lloyd et al. 2012a). As well as increasing levels of species richness in our database, this may also result in species range extensions (as some taxonomists apply names more loosely than others), in a greater likelihood that some reworked individuals will have been mis-recorded as valid species occurrences, and in more cases of taxonomic misidentification that can be corrected only by revisiting the original samples (e.g., Lazarus 2011). All of these effects are expected to dampen and delay drops in diversity in our raw data, as is indeed seen at the end-Cretaceous and end-Eocene.

Comparison with the Coccolithophore Record

Because the approach used here is essentially identical to that applied by Lloyd et al. (2012b) to coccolithophores, it is worth noting what similarities and differences emerge when comparing the two groups.

Overall, the two have very similar results. These include significantly different taxonomic diversity estimates between the deep-sea and the land, and extremely similar sampling curves (see Fig. 2C,F). Both records show the same disparity in post-Eocene trajectories, with a huge rise seen in both deep-sea richness and sampling, but a more level or even declining trend in both curves on the land. The only major difference is that, with coccolithophores, correcting for sampling in both records does show a notable convergence between the two records, something that is less marked or even divergent here (Table 3). However, it is still clear that sampling bias is a major factor for both groups, because this is the only reasonable explanation for major differences between two diversity curves produced for the same, essentially cosmopolitan, pelagic groups. Consequently both stud-

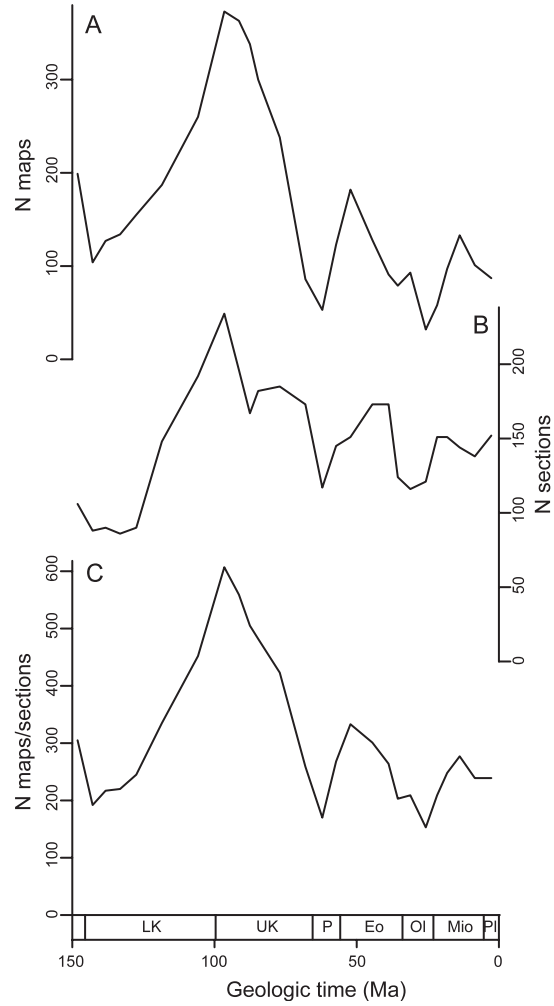

FIGURE 6. Estimates of land-based marine sedimentary rock abundance through time. A, Western Europe, based on numbers of geological map areas with outcrop (from Smith and McGowan 2007). B, North America, based on number of sedimentary packages listed in COSUNA charts (from Peters 2005). C, Combined western European and North American rock proxy data.

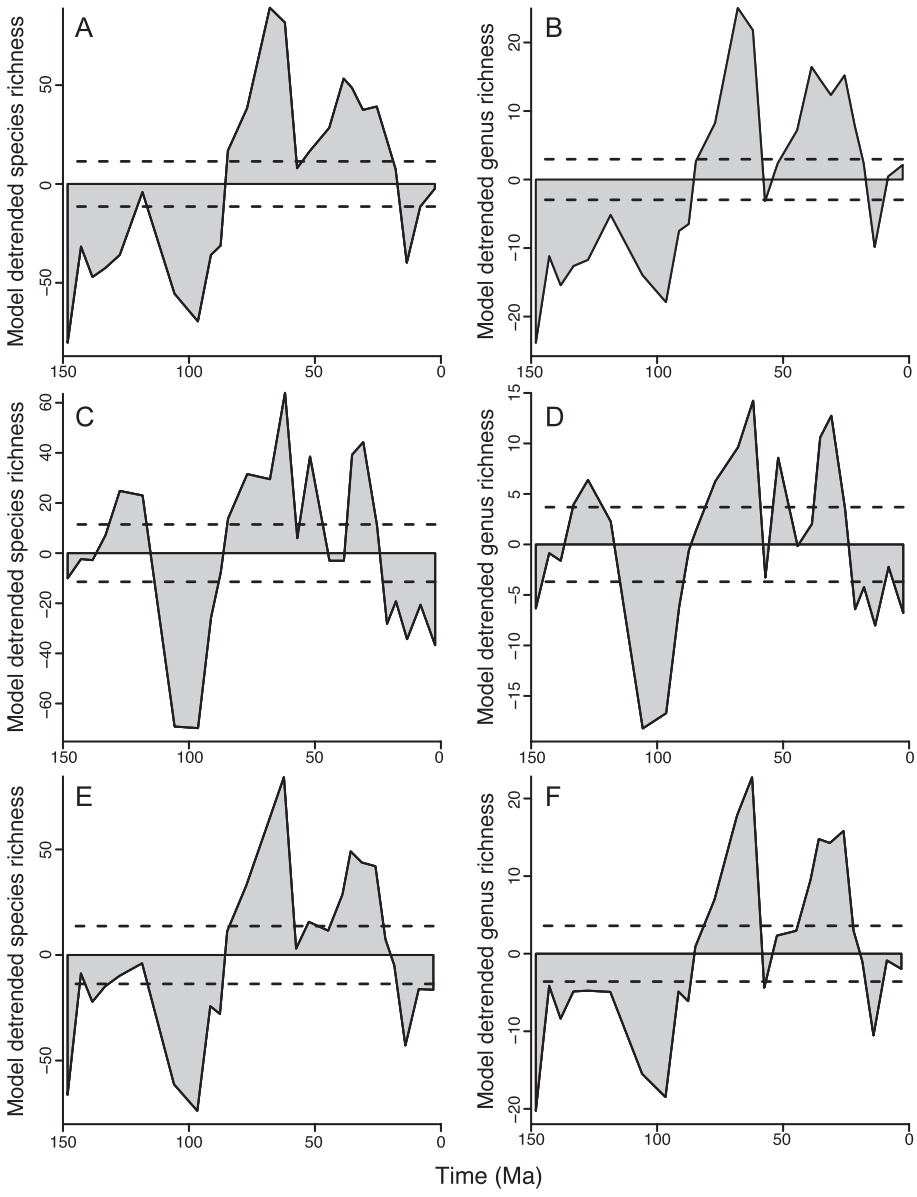

FIGURE 7. Residuals obtained by removing modeled from empirical species and genus richness using the land-based data only and a geologic proxy (number of maps). A, B, Species and genus residuals using western European maps (Smith and McGowan 2007). C, D, Species and genus richness using North American COSUNA charts (from Peters 2005). E, F, Species and genus residuals using the combined western European and North American map data as a proxy. In all cases dashed lines mark the 95% confidence estimates of the model.

ies support the continued use of sampling-corrected diversity estimates, whether through subsampling (Alroy 2010a) or modeling (as here).

Comparison with Rock Record Proxies

The shape of the species curve after correcting with SQS (Fig. 4A) and the outcrop area of marine sedimentary rock on land (Fig. 6) do not track each other closely. When only the land-based data and a geological proxy are used, residuals obtained by removing modeled from empirical species and genus richness show large and significant excursions indicating major mismatch (Fig. 7). Furthermore, compar-

ing our best estimate of diversity, using the combined land and deep-sea diversity data after correcting for sampling, with the combined North American and European rock record proxies showed no significant correlation (Spearman rho = 0.09, $p > 0.5$). Consequently, it is clear that diversity of planktonic foraminifera is not responding directly to the major transgression-regression cycles over the European and North American cratons bounding the Atlantic. It would appear that large-scale transgressive-regressive cycles over the cratonic blocks played little obvious role in shaping the diversity of planktonic foraminifera over time.

Conclusions

Traditionally, large-scale trends in micropaleontological diversity have been compiled using synoptic biostratigraphic range charts. Although such summaries are one step removed from the primary data and may never encompass the full diversity of recorded species, they have the advantage of focusing on groups with the most thoroughly studied fossil record where there is a high degree of taxonomic uniformity. They are, after all, based on synoptic publications by the leading taxonomists in their field who have expended considerable effort on identifying highest and lowest occurrences of species and calibrating those datums to the geological timescale. On the other hand they could, through their selectivity of taxonomic occurrence data, omit genuine but unexpected records. By contrast, analyzing the primary taxonomic occurrence data, as we have done here, relies on the actual records created by many taxonomic recorders. Such data must encompass greater taxonomic inconsistency and error, as not all recorders will have the same taxonomic experience, but has the major advantages that it is all encompassing and can be corrected for sampling bias. Sampling bias, as we have shown here, is no trivial matter, and a direct reading of the occurrence data in large regional databases of marine microplankton through time currently tells us more about the sampling pattern than about their true underlying diversity. Correcting for sampling generates a very different estimate of diversity than is obtained from the raw species counts of either land or deep-sea records, as expected, but it is also different from that obtained by traditional synoptic approaches. Although it is not always clear how best to remove sampling bias, it is reassuring that our two approaches, with their very different assumptions, point to a similar species-level curve.

Acknowledgments

This research was supported by National Environment Research Council (NERC) research grant NE/F016905/1. We would like to thank Andy Fraass for sharing his unpublished data on foraminiferal species ranges.

Literature Cited

Alroy, J. 2000. Successive approximations of diversoty curves: ten more years in the library. Geology 28:1023–1026.

———. 2010a. The shifting balance of diversity among major marine animal groups. Science 329:1191–1194.

———. 2010b. Geographical, environmental and intrinsic biotic controls on Phanerozoic marine diversification. Palaeontology 53:1211–1235.

———. 2010c. Fair sampling of taxonomic richness and unbiased estimation of origination and extinction rates. *In* J. Alroy and G. Hunt, eds. Quantitative methods in paleobiology. Paleontological Society Papers 16:55–80.

Alroy, J., M. Aberhan, D. J. Bottjer, M. Foote, F. T. Fürsich, P. J. Harries, A. J. W. Hendy, S. M. Holland, L. C. Ivany, W. Kiessling, M. A. Kosnik, C. R. Marshall, A. J. McGowan, A. I. Miller, T. D. Olszewski, M. E. Patzkowsky, S. E. Peters, L. Villier, P. J. Wagner, N. Bonuso, P. S. Borkow, B. Brenneis, M. E. Clapham, L. M. Fall, C. A. Ferguson, V. L. Hanson, A. Z. Krug, K. M. Layou, E. H. Leckey, S. Nürnberg, C. M. Powers, J. A. Sessa, C. Simpson, A. Tomašových, and C. C. Visaggi. 2008. Phanerozoic trends in the global diversity of marine invertebrates. Science 321:97–100.

Arnold, A., and W. Parker. 2003. Biogeography of planktonic foraminifera. Pp. 103–122 *in* B. K. Sen Gupta, ed. Modern foraminifera. Kluwer Academic, New York.

Aze, T., T. H. G. Ezard, A. Purvis, H. K. Coxall, D. R. M. Stewart, B. S. Wade, and P. N. Pearson. 2011. A phylogeny of macroperforate planktonic foraminifera from fossil data. Biological Reviews of the Cambridge Philosophical Society 86:900–927.

Bé, A. W. H. 1982. Biology of planktonic foraminifera. *In* T. W. Broadhead, ed. Foraminifera: notes for a short course. Studies in Geology 6:51–92. University of Tennessee, Knoxville.

Benson, R. B. J, R. J. Butler, J. Lindgren, and A. S. Smith. 2010. Mesozoic marine tetrapod diversity: mass extinctions and temporal heterogeneity in geological megabiases affecting vertebrates. Proceedings of the Royal Society of London B 277:829–834.

Crampton, J. S., A. G. Beu, R. A. Cooper, C. M. Jones, B. Marshall, and P. A. Maxwell. 2003. Estimating the rock volume bias in paleobiodiversity studies. Science 301:358–360.

Ezard, T. H. G., T. Aze, P. N. Pearson, and A. Purvis. 2011. Interplay between climate and species' ecology drives macroevolutionary dynamics. Science 332:349–351.

Hart, M. B., M. J. Oxford, and W. Hudson. 2002. The early evolution and palaeobiogeography of Mesozoic planktonic

foraminifera. Geological Society of London Special Publications 194:115–125.

Hemleben, C., M. Spindler, and O. R. Anderson. 1989. Modern planktonic foraminifera. Springer, New York.

Jablonski, D., K. Roy, J. W. Valentine, R. M. Price, and P. S. Anderson. 2003. The impact of the Pull of the Recent on the history of marine diversity. Science 16:1133–1135.

Kennett, J. P., and M. S. Srinivasan. 1983. Neogene planktonic foraminifera. Hutchinson Ross, Stroudsburg, Penn.

Lazarus, D. B. 2011.The deep sea microfossil record: potential and current data quality. *In* A. J. McGowan and A. B. Smith, eds. Comparing the geological and fossil records: implications for biodiversity. Geological Society of London Special Publication 358:141–166.

Lloyd, G. T. 2012. A refined modelling approach to assess the influence of sampling on palaeobiodiversity curves: new support for declining Cretaceous dinosaur richness. Biology Letters 8:123–126.

Lloyd, G. T., K. E. Davis, D. Pisani, J. E. Tarver, M. Ruta, M. Sakamoto, D. W. E. Hone, R. Jennings, and M. J. Benton. 2008. Dinosaurs and the Cretaceous terrestrial revolution. Proceedings of the Royal Society of London B 275:2483–2490.

Lloyd, G. T., A. B. Smith, and J. R. Young. 2011. Quantifying the deep sea rock and fossil record bias using coccolithophores. *In* A. J. McGowan and A. B. Smith, eds. Comparing the geological and fossil records: implications for biodiversity. Geological Society of London Special Publication 358:167–178.

Lloyd, G. T., J. Y. Young, and A. B. Smith. 2012a. Taxonomic structure of the fossil record is shaped by sampling bias. Systematic Biology 61:80–89.

———. 2012b. Comparative quality and fidelity of the deep-sea and land-based nannofossil records. Geology 40:155–158.

Loeblich, A. R., Jr., and H. Tappan. 1988. Foraminiferal genera and their classifications. Van Nostrand Reinhold, New York.

Mannion, P. D., P. Upchurch, M. T. Carrano, and P. M. Barrett. 2011. Testing the effect of the rock record on diversity: a multidisciplinary approach to elucidating the generic richness of sauropodomorph dinosaurs through time. Biology Reviews 86:157–181.

McKinney M. L. 1990. Classifying and analysing evolutionary trends. Pp. 28–58 *in* K. J. McNamara, ed. Evolutionary trends. Belhaven, London.

Miller, A. I., and M. Foote. 1996. Calibrating the Ordovician radiation of marine life: implications for Phanerozoic diversity trends. Paleobiology 22:304–309.

Miller, K. G., M. A. Kominz, J. V. Browning, J. D. Wright, G. S. Mountain, M. E. Katz, P. J. Sugarman, B. S. Cramer, N. Christie-Blick, and S. F. Pekar. 2005. The Phanerozoic record of global sea-level change. Science 310:1293–1298.

Norris, R. D. 1991. Biased extinction and evolutionary trends. Paleobiology 17:388–399.

Olsson, R. K., C. Hemleben, W. A. Berggren, and B. T. Huber 1999. Atlas of Paleocene planktonic foraminifera. Smithsonian Contributions to Paleobiology 85:1–252.

Pearson, P. N., R. K. Olson, B. T. Huber, C. Hemleben, and W. A. Berggren, eds. 2006. Atlas of Eocene planktonic foraminifera. Cushman Foundation Special Publication 41. Cushman Foundation, Washington, D.C.

Peters, S. E. 2005. Geologic constraints on the macroevolutionary history of marine animals. Proceedings of the National Academy of Sciences USA 102:12326–12331.

Peters, S. E., and M. Foote, 2001. Biodiversity in the Phanerozoic: a reinterpretation. Paleobiology 27:583–601.

Raup, D. M. 1972. Taxonomic diversity during the Phanerozoic. Science 177:1065–1071.

———. 1979. Biases in the fossil record of species and genera. Bulletin of the Carnegie Museum of Natural History 13:85–91.

R Development Core Team. 2010. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. ISBN 3-900051-07-0. http://www.R-project.org.

Sepkoski, J. J. 1993. Ten years in the library: new data confirm paleontological patterns. Paleobiology 19:43–51.

Smith, A. B. 2001. Large-scale heterogeneity of the fossil record: implications for Phanerozoic biodiversity studies. Philosophical Transactions of the Royal Society of London B 356:351–367.

Smith, A. B., and A. J. McGowan. 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.

Spencer-Cervato, C. 1999. The Cenozoic deep sea microfossil record: explorations of the DSDP/ODP sample set using the Neptune Database. Palaeontologica Electronica 2:4.

Spezzaferri, S. 1994. Planktonic foraminiferal biostratigraphy and taxonomy of the Oligocene and lower Miocene in the oceanic record: an overview. Palaeontographica Italica 81:1–187.

Stewart, D. R. M., and P. N. Pearson. 2000. PLANKRANGE: a database of planktonic foraminiferal ranges. Electronic database with documentation available at http://palaeo.gly.bris.ac.uk/Data/plankrange.html (updated December 2002)

Tappan, H., and A. R. Loeblich Jr. 1973. Evolution of the oceanic plankton. Earth-Science Reviews 9:207–240.

Wade, B. S., P. N. Pearson, W. A. Berggren, and H. Pälike. 2011. Review and revision of Cenozoic tropical planktonic foraminiferal biostratigraphy and calibration to the geomagnetic astronomical time scale. Earth-Science Reviews 104:111–142.