Paleobiology, 49(4), 2023, pp. 733–746
DOI: 10.1017/pab.2023.16

PALEOBIOLOGY
A PUBLICATION OF THE Paleontological SOCIETY

Article

Relative oversampling of carbonate rocks in the North American marine fossil record

Diego Balseiro and Matthew G. Powell*

Abstract.—Paleontologists have long stressed the need to know how sampling the fossil record might influence our knowledge of the evolution of life. Here, we combine fossil occurrences of North American marine invertebrates from the Paleobiology Database with lithologic data from Macrostrat to identify sampling patterns in carbonate and siliciclastic rocks. We aim to quantify temporal trends in sampling effort within and between lithologies, focusing on the proportion of total available volume that has been sampled (sampled fossiliferous proportion, here called κ). Results indicate that the sampled fossiliferous proportion was stable during the Paleozoic, and variable during the post-Paleozoic, but showed no systematic increase through time. Fossiliferous carbonate rocks are proportionally more sampled than siliciclastic rocks, with intervals where the carbonate κ is double the siliciclastic κ. Among possible explanations for the apparent oversampling of fossiliferous carbonate rocks, analyses suggest that barren units, taphonomic dissolution, or data entry errors cannot completely explain sampling patterns. Our results suggest that one of the important drivers might be that paleontologists publish taxonomic descriptions from carbonate rocks more frequently. The higher diversity in carbonate rocks might account for an ease in the description of unknown species and therefore a higher rate of published fossils. Finally, a strong effect in favor of carbonate rocks might distort our perception of diversity through time, even under commonly used standardization methods. Our results also confirm that previous descriptions of an increase in the proportion of sampled fossiliferous rocks over time were driven by the sampling of the nonmarine fossil record.

Diego Balseiro. Facultad de Ciencias Exactas Físicas y Naturales, Universidad Nacional de Córdoba, Avenida Vélez Sarsfield 1611 Ciudad Universitaria, Córdoba X5016GCA Argentina; CONICET, Centro de Investigaciones en Ciencias de la Tierra (CICTERRA), Córdoba X5016GCA, Argentina. E-mail: dbalseiro@unc.edu.ar

Matthew G. Powell. Department of Geology, Juniata College, 1700 Moore Street, Huntingdon, Pennsylvania 16652 USA. E-mail: powell@juniata.edu

Accepted: 20 April 2023
*Corresponding author.

Introduction

Sampling issues have received special attention from paleontologists because of their potential to distort our understanding of the history of life (Raup 1972; Smith and McGowan 2011). Raup (1976) first demonstrated that temporal trends in raw fossil diversity paralleled trends in outcrop area, a relationship that was later found to hold for other proxies of rock quantity (Peters and Foote 2001; Smith and McGowan 2007; Wall et al. 2009), raising the possibility that sampled diversity patterns merely reflect geological factors (Peters and Foote 2001) or that evolutionary processes are not independent of rock volume (i.e., common cause hypothesis; Peters and Foote 2001). Unfortunately, these direct comparisons of sampled diversity and rock availability do not completely capture the actual nature of paleontological sampling. To comprehend sampling of the fossil record, paleontologists must analyze patterns based on sampled rock volumes (together with taphonomic and lithologic information) rather than on the basis of raw estimates of rock volume and diversity, because diversity varies for both biological and sampling-related reasons (Magurran 2004; Hayek and Buzas 2010).

Disentangling the effects of sampling is important, because paleontological sampling may also underlie broadscale environmental and ecological trends through the Phanerozoic, such as the frequency of bottom-level anoxia in the seas (Peters 2007) and/or the filling of habitats as marine life diversified (Smith and McGowan 2008). Peters (2007) proposed that

https://doi.org/10.1017/pab.2023.16 Published online by Cambridge University Press

units may be barren of fossils due to habitat harshness and that the decrease of barren units over time was due to a decline in anoxia. By contrast, Smith and McGowan (2008) proposed that the same trend reflected progressive invasion of habitat as marine life diversified. Both Peters (2007) and Smith and McGowan (2008) interpreted the Phanerozoic decline in the proportion of barren units as a biological reality. However, it the degree to which sampling effects play a role is as yet unclear, because these former studies did not differentiate between the sampled and unsampled rock volume.

Here, we focus on two specific, underexplored issues related to paleontological sampling. First, the type and degree of sampling effects may vary over time as the relative proportions of carbonate and siliciclastic marine sedimentary rocks change (Peters 2006). Paleontological sampling differs between these lithologies due to variation in the degree of lithification, which affects the ability to collect fossils (Hendy 2009; Hawkins et al. 2018; but also see Daley and Bush 2020) and causes the loss of small fossils (Sessa et al. 2009; but see Nawrot 2012). Carbonates may also be prone to early diagenetic dissolution, which may reduce the proportion of fossil-bearing carbonate rocks (Kidwell et al. 2005; Best 2008), although significant disintegration also occurs in siliciclastic sediments (Aller 1982; Tomašových et al. 2019). Previous analyses using the Paleobiology Database (PBDB) have shown that paleontological sampling has high geographic coverage (Alroy 2010b), but lithologic coverage has received less attention. Consequently, it is not yet clear, in a global sense, how temporal variation in the geographic extent of lithology and/or sampling of those lithologies affects secular diversity patterns. A case study from the late Paleozoic showed that marine diversity was controlled by the changing proportion of carbonate and siliciclastic rocks (Balseiro and Powell 2020); this work extends some of those implications to the entire Phanerozoic.

Second, by focusing our analysis on the marine record, we can differentiate sampling effects that are specific to that realm, augmenting studies that described trends for the rock record as a whole (Peters and Heim 2010). The marine and nonmarine fossil records differ in their stratigraphic architectures (Catuneanu 2006), taphonomic pathways (Behrensmeyer et al. 2000), and the stratigraphic and geographic nature of sampling between vertebrate (mostly continental) and invertebrate (mostly marine) faunas (Holland and Loughney 2021; Holland 2022). Hence, to understand sampling patterns, it is necessary to analyze how paleontological sampling varies within a major environment (marine or nonmarine).

Here, we follow the method developed by Peters and Heim (2010) by joining a stratigraphic database (Macrostrat) with a paleontological database (PBDB) to better understand how the changing proportion of carbonate and siliciclastic marine sedimentary rocks affects the North American marine invertebrate fossil record of four diverse and well-preserved major fossil taxa. We quantify temporal trends in sampled fossiliferous volumes within and between lithologies, focusing on the proportion of total available volume sampled (i.e., completeness; Peters and Heim 2010).

Data and Methods

We estimated the sampled, fossiliferous proportion of carbonate and siliciclastic rocks as:

$$\kappa = \text{(volume of sampled, fossiliferous rock)}/\text{(total available volume)} \tag{1}$$

that is, the proportion of the available rock record that has yielded identifiable fossils that have been entered into the PBDB. This relationship has been termed "geological completeness" by Peters and Heim (2010). Although the calculation is a simple proportion, we refer to it here with precision to avoid confusion with "percent fossil-bearing," which would also include unsampled, fossiliferous rock. (Because of potential uncertainty in estimating rock volume, we also calculated sampled fossiliferous proportion using sediment coverage area; see Supplementary Material.)

κ (kappa) was calculated for siliciclastic and carbonate volumes separately as well as for

https://doi.org/10.1017/pab.2023.16 Published online by Cambridge University Press

total rock volumes. We expressed the relative sampled fossiliferous proportion of carbonate (κ_c) and siliciclastic (κ_s) sedimentary rocks over time using a ratio, which is the base 2 logarithm of carbonate sampled fossiliferous proportion divided by siliciclastic sampled fossiliferous proportion:

$$\kappa\text{-ratio} = \log_2(\kappa_c/\kappa_s) \quad (2)$$

Positive values indicate time periods when a greater proportion of the carbonate record has been sampled and entered into the PBDB than the siliciclastic record (carbonate oversampling), whereas negative values indicate the opposite. A ratio of 1 indicates that κ of the best-sampled lithology is double that of the other lithology, while a value of 2 indicates that it is four times higher.

We estimated sampled and total rock volumes for 12,083 strictly marine units in 863 columns from the Macrostrat database (https://macrostrat.org) limited to North America, which we operationally defined as continental United States (i.e., excluding Hawai'i and U.S. territories) and Canada. Macrostrat columns are geographic regions defined using a Delaunay tessellation (also known as Voronoi diagrams of Thiessen polygons) (Peters and Heim 2010). Columns are divided into units, which are genetically, lithologically, or chronologically distinct bodies of rock within the column (Peters et al. 2018); Macrostrat units usually correspond to a geographic subset of a formal stratigraphic unit, such as a formation (Fig. 1). Macrostrat records both outcropping and subsurface units, but given that sampling of macrofossils is essentially limited to outcrops, we eliminated all subsurface units from the analyses. We calculated rock volumes as the stratigraphic thickness of a unit within a time interval multiplied by the area of the Macrostrat column in which it occurred (Balseiro and Powell 2020). Units that crossed an interval boundary were proportionally allocated to each interval based on its duration within each interval. Durations of units are calculated using Macrostrat's continuous-age model (Peters et al. 2018). The volume of carbonate or siliciclastic rocks within each time unit was estimated by multiplying the total volume of a unit by the proportion of each lithology that is recorded in Macrostrat. This methods allows estimation of individual carbonate and siliciclastic volumes from mixed carbonate/siliciclastic units.

Fossil occurrences of trilobites, brachiopods, mollusks, and cnidarians were obtained from the PBDB (https://paleobiodb.org). We restricted our analysis to these four clades because they are the most abundant taxa in the PBDB, have similar preservation potential, and do not depend on exceptional preservation to be taxonomically identifiable at the genus level, as is the case for some echinoderms. We then joined the paleontological data with the stratigraphic data from Macrostrat based on the collection identification number (Peters and Heim 2010).

The lithology of each collection was characterized as carbonate or siliciclastic. Carbonates were defined as lithologies containing the words "carbonate", "limestone", "reef rocks", "bafflestone", "bindstone", "dolomite", "framestone", "grainstone", "lime mudstone", "packstone", "rudstone", and "wackestone". Siliciclastics were defined as lithologies containing the words "shale", "siliciclastic", "breccia", "claystone", "conglomerate", "gravel", "mudstone", "quartzite", "sandstone", "siltstone", and "slate". We coded for carbonate or siliciclastic lithology when the two primary lithology fields in the PBDB did not record different lithologies or specify the lithology from which the fossils came (Balseiro and Powell 2020). We then discarded all collections that did not come from fully marine Macrostrat units. The final dataset consists of 25,455 collections from 2557 Macrostrat units from the Fortunian (lowermost Cambrian) to Piacenzian (Pliocene).

We divided the Phanerozoic into similar time bins of 10 Myr duration (Supplementary Material). Mean duration of bins is 9.83 Myr, with maximum duration of 19.2 Myr (Visean) and minimum duration of 5 Myr (Ladinian); 80% of the bins are within ±2 Myr of the mean.

Using this dataset, we estimated sampled fossiliferous volumes as the sum of volumes of units that recorded at least one paleontological collection (Fig. 1). Carbonate and siliciclastic volumes were similarly calculated as the volumes of units that recorded

https://doi.org/10.1017/pab.2023.16 Published online by Cambridge University Press

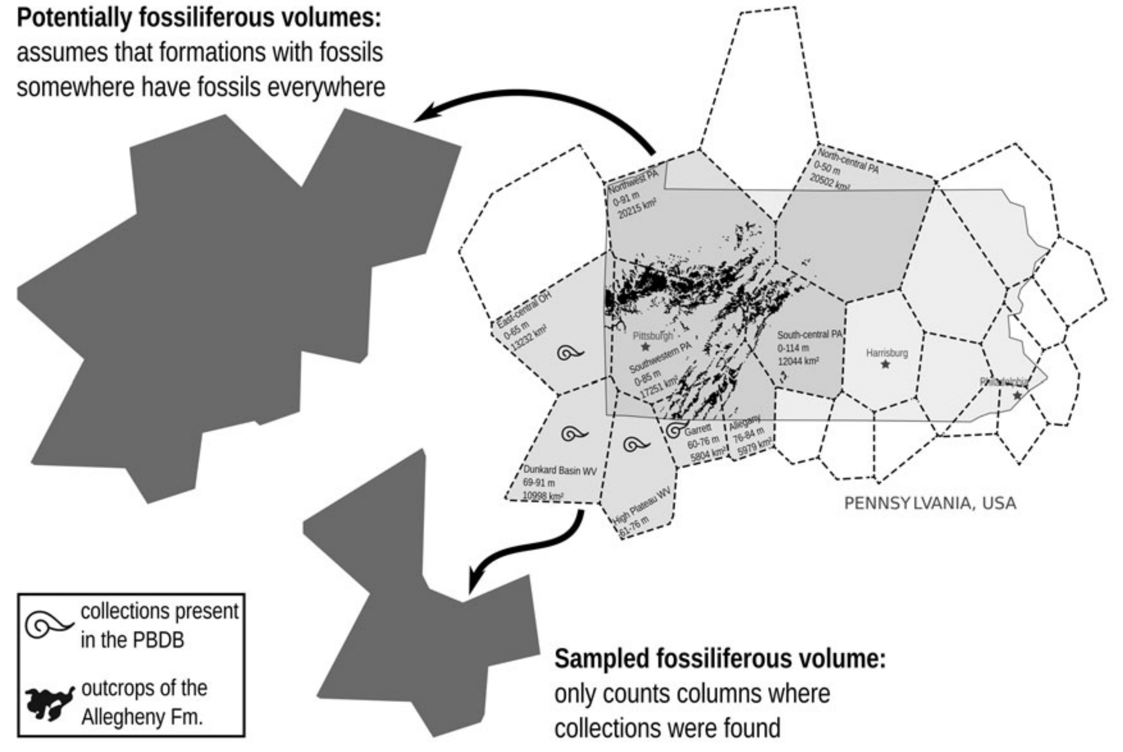

FIGURE 1. Illustrated example of the method used to estimate sampled and potentially fossiliferous volumes using the Allegheny Formation in Pennsylvania. Created based on data from USGS, Macrostrat, and Pennsylvania Bureau of Topographic and Geologic Survey.

paleontological collections from each lithology. In the case of mixed carbonate/siliciclastic units, we computed only the volume of the lithology corresponding to the collections present in the unit. In other words, if a mixed unit contained only siliciclastic collections, then only the siliciclastic volume of the unit was computed as sampled. Sampled volumes calculated this way are approximations of the minimum rock volume that has been actually sampled, because many localities with sampled fossils are either unpublished or have not been registered in the PBDB (Marshall et al. 2018), and because we eliminated all collections that lacked lithologic information or that could not be undoubtedly identified as either siliciclastic or carbonate. To overcome these limitations, we also calculated potentially fossiliferous volumes, which assumed that lithostratigraphic formations with at least one fossiliferous unit recorded in the PBDB were fossiliferous in every unit of that formation. Potentially fossiliferous volume was calculated by summing the volumes of all Macrostrat units that belong to the same formal lithostratigraphic unit, where at least one of those Macrostrat units recorded at least one PBDB collection (Fig. 1).

We tested whether time series of κ_c, κ_s, and κ-ratio were best explained by a random process, a directional trend, a stable dynamic, or a complex combination of these (Hunt et al. 2015) adopting a full maximum-likelihood approach using the paleoTS package for R (Hunt 2021). A fully random dynamic would be modeled as an unbiased random walk (URW), a directional trend as a generalized random walk (GRW), and a stable dynamic as stasis (Hunt 2008), whereas a combination of these possibilities implies a shift in the underlying dynamic. We estimated the variance needed for the analysis by bootstrapping columns 1000 times. Akaike weights based on Akaike information criterion corrected for small

https://doi.org/10.1017/pab.2023.16 Published online by Cambridge University Press

sample sizes (AICc) were used to select the most-supported models given the data. Akaike weights can be interpreted as the probability that a given model is the best one among a set of models analyzed, given the data (Burnham and Anderson 2002).

All analyses were carried out in R (R Core Team 2018). Data and R scripts used in the analysis are available as Supplementary Material.

Results

Carbonate rocks were sampled more intensively than siliciclastic rocks throughout most of the Phanerozoic (Fig. 2A). The average κ-ratio for all intervals was 1.14, indicating that κ_c was 2.2 times higher than κ_s rocks, overall. Only two brief intervals exhibited more than one consecutive epoch of greater κ_s: the first centered near the Triassic/Jurassic boundary (Triassic 5–Jurassic 1), in which the average κ-ratio was −1 (i.e., κ_s was one time higher than κ_c), and the second centered on the Late Cretaceous (Cretaceous 6–Cretaceous 7), in which the average κ_ratio was −0.42 (i.e., κ_s was 1.3 times higher than κ_c). A third longer interval, spanning the Late Jurassic–Early Cretaceous, is also relevant, as the κ-ratio should be highly negative but cannot be computed because κ_c is 0. When these intervals are excluded from the calculation, the average κ-ratio for the remainder of the Phanerozoic was 1.3, indicating that κ_c was 2.5 times higher than κ_s. Geologically brief intervals of higher κ_s occurred in Silurian 1 and Devonian 5. The time series analysis indicates that the mean higher κ_c during most of the Phanerozoic is not caused by chance alone, but responds to a stable pattern that fluctuates around 1 (i.e., κ_c is 2 times higher than κ_s; Table 1, but see Supplementary Material for further analysis.). The early Paleozoic (Cambrian 1–Ordovician 4) shows an even higher κ-ratio, with a trend around 2.58.

Carbonate and siliciclastic rocks exhibited somewhat dissimilar patterns of κ (the correlation of κ_c and κ_s was just 0.39). Carbonate rocks exhibited high κ_c during Permian 1–Jurassic 4, during which average κ_c rose to 37%, compared with the Phanerozoic average of 25%, and even exceeded 70% during Triassic 1–Triassic 2 (Fig. 2B,C). Carbonates also exhibited higher κ_c during Paleogene 3–Neogene 3, when the average value rose to 47%. By contrast, there were few sustained intervals of unusually high κ_s, which on average was 13%, exhibited isolated spikes of higher values in Triassic 1–Triassic 2, Jurassic 1, Jurassic 4–Jurassic 5, Cretaceous 6–Cretaceous 8, and Paleogene 4–Neogene 1. Average κ_s during these time intervals was 33% and reached a Phanerozoic maximum of 42% during Triassic 2. This conclusion is supported by time series analyses, which show that both trends are best explained by complex models including two or more shifts between stable sampling trends (stasis; Table 1). The shifts, however, are not coincident between lithologies, suggesting the absence of a single explanation.

Inspection of sampled fossiliferous volume and total available volume shows that the incongruent patterns are caused by opposite underlying dynamics: κ_c is driven primarily by changes in total volume, whereas κ_s is driven primarily by changes in sampled volume (Fig. 3). The average total carbonate volume during Permian 1–Jurassic 5 (85,880 km^3) was 39% of the Phanerozoic average (230,324 km^3), whereas during that same time interval the average sampled carbonate volume (31,295 km^3) was 78% of the Phanerozoic sampled average (40,282 km^3). In other words, carbonate sampled volume was relatively unchanged even as its total volume decreased. Siliciclastic rocks experienced the opposite pattern. During intervals of unusually high κ_s, total volumes were only slightly above the Phanerozoic average, by 17% (584,380 km^3 compared with the average of 499,252 km^3), while sampled volumes increased by 178% (to 188,884 km^3 compared with the average of 67,683 km^3).

In a general sense, variation in sampled volumes can be explained by variation in total rock volumes, as shown by the correlation between first differences of these metrics ($r = 0.63$, $p = 2 \times 10^{-7}$), indicating that rock availability controls paleontological sampling at a large scale, as originally suggested by Raup (1972, 1976; see also Peters and Heim 2010). However, this general relationship obscures the fact that sampling of carbonate and siliciclastic lithologies is significantly different. Notably, carbonate rocks are consistently more fossil-bearing

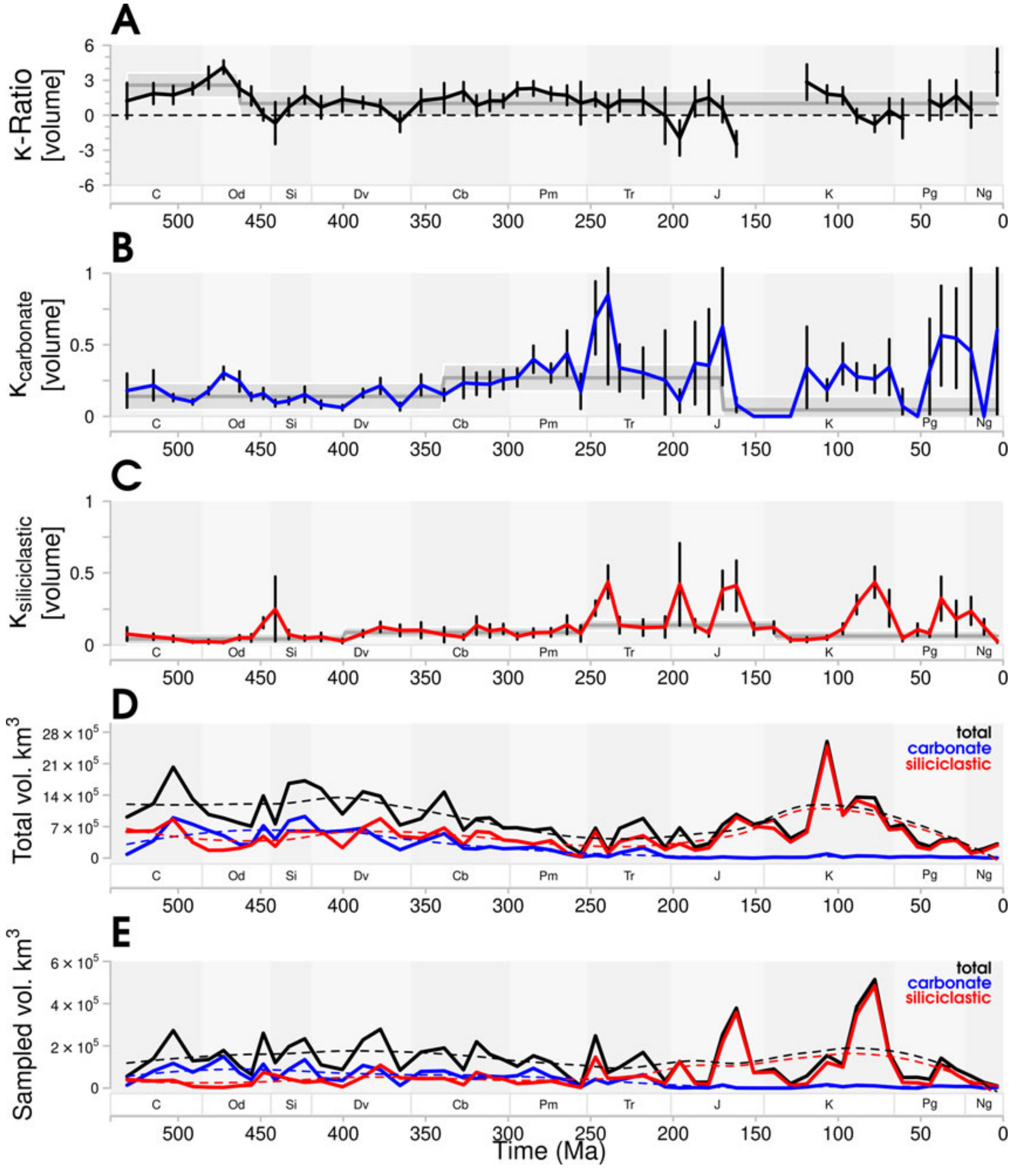

FIGURE 2. A, κ-ratio trend. B, Sampled proportion (κ) of carbonate rocks. C, Sampled proportion (κ) of siliciclastic rocks. Whiskers indicate 1 SD. Best-supported time series models are shown as dark gray lines with 95% probability envelopes in light gray. D, Available total marine sedimentary rock volume. E, Sampled marine sedimentary rock volume. Dashed lines are loess regressions, a local polynomial regression, with a smoothing parameter (alpha) of 0.4. Data plotted at interval midpoints.

(in that they have a higher κ throughout the Phanerozoic), and the sampled fossiliferous proportion (κ) of each lithology appears to have different underlying drivers.

Discussion

We consider four hypotheses to explain the oversampling of carbonate rocks relative to siliciclastic rocks (Table 2). Some were

TABLE 1. Results of the maximum-likelihood estimates for time series modeling; boldface indicates the most-supported model. AICc, Akaike information criterion corrected for small sample sizes; κ, sampled fossiliferous proportion; GRW, general random walk (i.e., directional trend); URW, unbiased random walk (i.e., random trend); GRW-stasis, combination of an initial directional and a later stable trends; URW-stasis, an initial random dynamic followed by a stable trend; Stasis-GRW, an initial stable trend followed by a directional trend; Stasis-URW, an initial stable trend followed by a random trend.

Data	Model	logL	K	AICc	Akaike weight
κ-ratio	GRW	−80.1	3	166.6	0.003
	URW	−80.1	2	164.3	0.01
	Strict stasis	−91.5	1	185.1	0
	Stasis	−80.9	2	166.1	0.004
	Two stasis intervals	**−74.2**	**4**	**157.3**	**0.345**
	Three stasis intervals	−72.1	6	158.1	0.24
	Four stasis intervals	−69.3	8	158.2	0.22
	GRW-stasis	−72.9	6	159.7	0.1
	URW-stasis	−74.8	5	160.9	0.058
	Stasis-GRW	−77.8	5	167	0.003
	Stasis-URW	−78	4	164.8	0.008comp
κ_{cb}	GRW	33.6	3	−60.8	0.047
	URW	33.6	2	−60.9	0.133
	Strict stasis	−154	1	310	0
	Stasis	28	2	−51.8	0.001
	Two stasis intervals	32.6	4	−56.4	0.005
	Three stasis intervals	**39.5**	**6**	**−65.3**	**0.437**
	Four stasis intervals	41.3	8	−63.5	0.182
	GRW-stasis	35.7	6	−57.6	0.009
	URW-stasis	35.4	5	−59.6	0.026
	Stasis-GRW	35.5	5	−59.8	0.028
	Stasis-URW	35.8	4	−68.9	0.131
κ_{sl}	GRW	65.3	3	−124.1	0.008
	URW	65.3	2	−126.3	0.023
	Strict stasis	36.1	1	−70.1	0
	Stasis	63.4	2	−122.6	0.004
	Two stasis intervals	69.2	4	−129.8	0.128
	Three stasis intervals	71.8	6	−129.8	0.128
	Four stasis intervals	**75.6**	**8**	**−132**	**0.390**
	GRW-stasis	72.6	6	−131.5	0.305
	URW-stasis	68.3	5	−125.4	0.014
	Stasis-GRW	64.4	5	−117.6	0
	Stasis-URW	64.4	4	−120	0.001
κ_{total}	GRW	59.8	3	−113.2	0.005
	URW	59.8	2	−115.4	0.016
	Strict stasis	18.1	1	−34.2	0
	Stasis	61.7	2	−119.2	0.1
	Two stasis intervals	63.7	4	−118.6	0.079
	Three stasis intervals	**68.4**	**6**	**−1228**	**0.64**
	Four stasis intervals	69.4	8	−119.6	0.134
	GRW-stasis	62.5	6	−111.3	0.002
	URW-stasis	62	5	−112.8	0.002
	Stasis-GRW	61.6	5	−112.1	0.003
	Stasis-URW	61.6	4	−114.4	0.01
Potentially fossiliferous κ-ratio	GRW	−51.5	3	109.4	0.06
	URW	−51.6	2	107.3	0.173
	Strict stasis	**−51.6**	**1**	**105.1**	**0.507**
	Stasis	−51.6	2	107.3	0.173
	Two stasis intervals	−51.8	4	110.4	0.037
	Three stasis intervals	−50.2	6	114.1	0.006
	Four stasis intervals	−50	8	119	0.001
	GRW-stasis	−49.8	6	113.3	0.009
	URW-stasis	−50.8	5	112.3	0.011
	Stasis-GRW	−51.3	5	113.8	0.007
	Stasis-URW	−51.6	4	111.9	0.018

https://doi.org/10.1017/pab.2023.16 Published online by Cambridge University Press

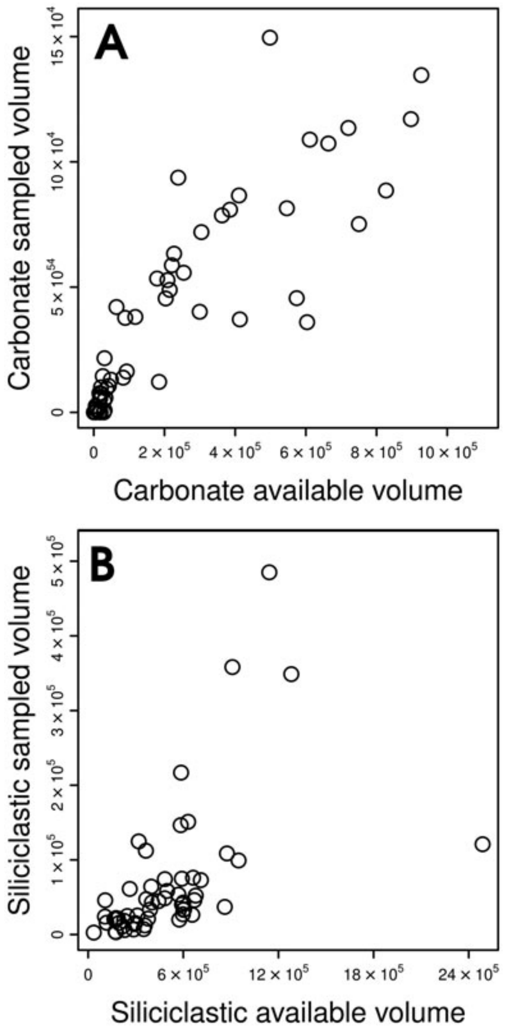

FIGURE 3. Scatter plot of available and sampled volumes. A, Carbonate rocks. B, Siliciclastic rocks.

previously discussed by Peters and Heim (2010) but deserve additional scrutiny, as the factors accounting for unequal κ by lithology (κ_c and κ_s) might differ greatly from those accounting for a trend of increasing κ through the Phanerozoic. Moreover, the difference in κ between lithologies is much larger than the previously described difference between mean Paleozoic κ and Cretaceous–Cenozoic κ (Peters and Heim 2010).

Barren Units.—A possible explanation for greater κ_c than κ_s is that the original siliciclastic depositional environments were relatively devoid of macroscopic life. Peters (2007) described this as the "problem with the Paleozoic" and hypothesized that barren units were due to the anoxia of Paleozoic epeiric seas. However, any factor that may account for environmental harshness can explain the pattern described by Peters (2007). It is plausible that the effect of environmental harshness is greater for siliciclastic environments, because nearly all carbonate sediments are biogenic (although not necessarily of invertebrate origin), and because carbonate sediments are deposited in shallow-marine conditions that are more likely to be suitable for life. The observed trend in the κ-ratio (Fig. 2) does support Peters's (2007) proposition, given that carbonates are particularly oversampled during the Paleozoic. However, the hypothesis is difficult to test directly, because the data do not distinguish between barren units and those that are merely unsampled; both are recorded as Macrostrat units without an associated PBDB collection.

We attempted to differentiate barren and unsampled units by calculating "potentially fossiliferous volume," which classifies units as fossiliferous if the unit belongs to a lithostratigraphic formation that has at least one other fossiliferous Macrostrat unit, even if that specific unit had no PBDB collections associated with it. (The approach assumes that formations are equally fossiliferous everywhere they occur.) Carbonate and siliciclastic rocks have similar κ when calculated this way, as the κ-ratio (1) follows a stable trend very close to zero (~0.3) (Fig. 4A, Table 1), (2) has values closer to zero in each interval (paired Wilcoxon signed-rank test $V = 1008$, $p = 1.4 \times 10^{-5}$), and (3) is less variable (F-test: $F = 2.5$, $p = 0.001$) than raw values. This indicates that higher κc is due to undersampling of fossiliferous siliciclastic formations, rather than a greater number of barren siliciclastic formations. This conclusion is further supported by the fact that unpublished gray data in museum collections cover a much larger geographic and stratigraphic extent than published data (Marshall et al. 2018), which suggests that much of the supposedly unfossiliferous stratigraphic record is actually unpublished or unsampled, rather than truly barren. This reinforces the idea that the described lithologic inequality is a consequence of greater carbonate sampling, given

TABLE 2. Factors affecting sampled fossiliferous volume addressed in this study.

Filter	Potential effect	Test
Ecological	Carbonate sedimentary environments were less likely to be barren.	Recalculate fossiliferous volume to include unsampled units of formations that contain fossils elsewhere ("potentially fossiliferous volume") to estimate the effect of collections that were not sampled or entered into the Paleobiology Database
Taphonomic	Fossils from carbonate rocks were more likely to be preserved.	Calculate sampling coverage for aragonitic and calcitic taxa separately to estimate the effect of preferential dissolution
Data entry	Lithologic information from siliciclastic collections is less likely to be entered into the Paleobiology Database.	Redefinition of all non-carbonate lithologies as siliciclastic to estimate the effect of preferential exclusion of siliciclastic collections from the original analysis
Publication	Fossils from carbonate rocks are more likely to be reported in the literature.	Taxonomic diversity remeasured after triple rarefaction to equalize Macrostrat columns, sampled volumes, and collections to estimate whether less sampling effort is needed to find new species in carbonate rocks relative to siliciclastic rocks

that the formation-scale estimation is, potentially, less influenced by unsampled units. Finally, it is unlikely that uninhabited environments were widespread enough geographically or temporally to explain a consistent pattern over much of the Phanerozoic. Therefore, a pervasive effect of environmental harshness is unlikely to be a main driver of lower κs.

An additional potential driver of barrenness is that sedimentation rates can be very high in certain siliciclastic environments (Sadler, 1981), raising the possibility that the observed lower κ_s is due to "dilution" of fossiliferous zones by sediment. Siliciclastic units in our data do have higher mean sedimentation rates (44.3 m/Myr for carbonates vs. 70.7 m/Myr for siliciclastics; $t = -8.2$, $p < 2 \times 10^{-16}$) and marginally lower collection densities (0.18 collections per meter for carbonates vs. 0.14 collections per meter for siliciclastics; $t = 1.4$, $p = 0.16$). However, dilution can only affect our results if more barren siliciclastic units had

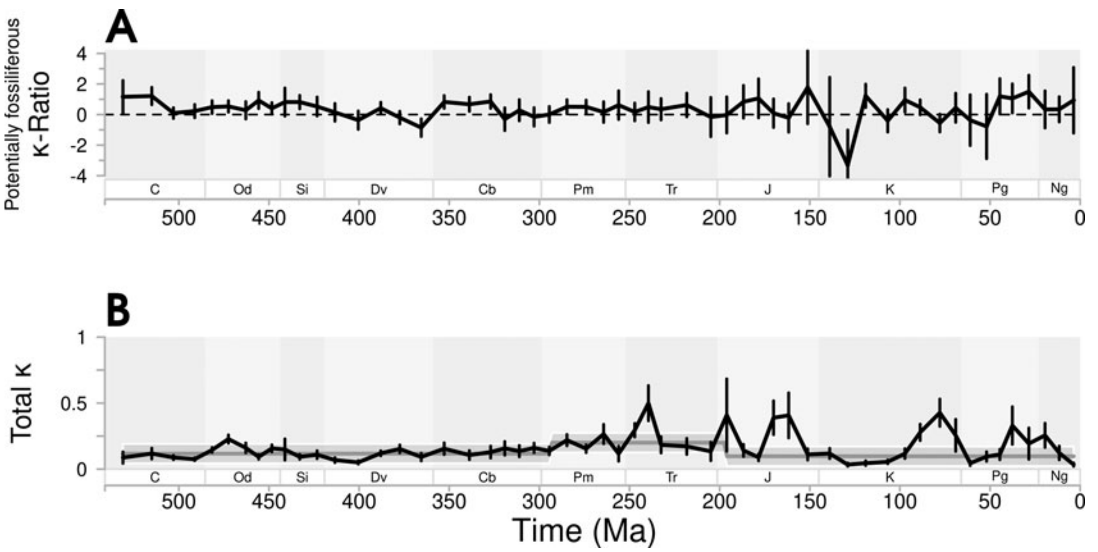

FIGURE 4. A, Trend in κ (sampled proportion) ratio estimated with potentially fossiliferous volumes. B, κ for marine sedimentary rocks. Whiskers indicate 1 SD. Best-supported time series model is shown as a dark gray line with 95% probability envelopes in light gray.

https://doi.org/10.1017/pab.2023.16 Published online by Cambridge University Press

been defined per unit of time, which would inflate total available volume; intra-unit collection density does not factor into our calculation. Siliciclastic units exhibit longer mean time durations than carbonate units (6.0 Myr for carbonates vs. 6.9 Myr for siliciclastics; $t = -4.6$, $p = 5 \times 10^{-6}$), indicating that siliciclastic units are not subdivided more than carbonate units.

Taphonomic Effects.—Early diagenetic dissolution is quite common in Paleozoic sediments (Cherns and Wright 2009, 2011) and has been identified as the cause of barren intervals in siliciclastic sedimentary successions (Schovsbo 2001). If early diagenetic dissolution was more common in siliciclastic than in carbonate sediments (Alexandersson 1976), siliciclastic rocks would appear to be less fossiliferous. This possibility is unlikely, given that carbonate rocks may be more likely to experience dissolution (Kidwell et al. 2005; Best 2008; Foote et al. 2015; but see Tomašových et al. 2019) or can be subjected to other mechanisms contributing to disintegration, such as bioerosion (Best and Kidwell 2000), and that genus-level taxonomic identification may be made from molds. Nevertheless, we tested the hypothesis by analyzing the sampled, fossiliferous proportion of aragonitic and calcitic fossils separately. Because aragonite is less stable than calcite (Cherns and Wright 2000), a pattern of preferential dissolution in siliciclastic sediments should be stronger in aragonitic than calcitic fossils. Aragonitic taxa, however, have higher κ_s than κ_c, (Fig. 5), indicating that, if there is any dissolution effect, it is against carbonate facies (paired Wilcoxon signed-rank tests: total, $V = 745$, $p = 0.056$; Paleozoic, $V = 238$, $p = 0.042$, post-Paleozoic, $V = 139$, $p = 0.7$). This reinforces the described pattern of higher carbonate oversampling.

Data Entry Effects.—κ_s could be lower if paleontologists are less likely to report in the literature, or to enter into the PBDB, the lithology of siliciclastic fossiliferous beds. The systematic exclusion of lithologic information from originally siliciclastic collections in the dataset would reduce κ_s, because these collections would not be classified as siliciclastic based on our analytical protocol. We tested this possibility by reanalyzing κ_s after relaxing the definition of siliciclastic collections, including all non-carbonate lithologies as siliciclastic whether or not they were specifically identified as such. This increased the number of occurrences considered to be siliciclastic by 56%, from 54,907 to 85,663. Although this approach reduced the κ-ratio for some Paleozoic intervals, overall it yielded results very similar to the previously described patterns (Fig. 6, $r = 0.93$, $p = 2 \times 10^{-16}$). Differences in Phanerozoic median values are not significant based on a Wilcoxon signed-rank test ($W = 1530$, $p = 0.054$), nor are differences in the variance of κ-ratio values (F-test: $F = 1.04$, $p = 0.88$). Even in this very improbable scenario (there is no a priori reason to suspect such a strong bias on the original description of lithologic information or during compilation of the PBDB), κ_c still exceeds κ_s in many intervals. A systematic bias in data entry, even if present, cannot be strong enough as to explain the observed high differences in κ between lithologies.

Publication Effects.—Because the PBDB is based primarily on published data, the observed inequality between κ_c and κ_s may also result from more publications of fossil data from carbonate rocks. This would be unexpected, given that carbonates are more difficult to sample than siliciclastic rocks (Hendy 2011). However, if the evenness of taxonomic occurrences is greater in carbonate environments, then more new taxa will be recovered at equal sampling effort (Powell and Kowalewski 2002). If new taxa are more likely to be published in the literature than previously discovered taxa, as is known for the paleontological literature (Alroy 2010a; Close et al. 2018), this would lead to a publication bias in favor of carbonate rocks. A simple triple rarefaction, equalizing the amount of Macrostrat columns, sampled volumes, and collections, shows that carbonate rocks usually record higher diversity than siliciclastic rocks at the same sampling effort (Fig. 7). Therefore, carbonate rocks are more likely to contain previously undiscovered taxa, increasing the number of places where this lithology has been reported in published data. This possible explanation is bolstered by the observation that published records cover less area than museum collections (Marshall et al. 2018); that is, many of the unpublished

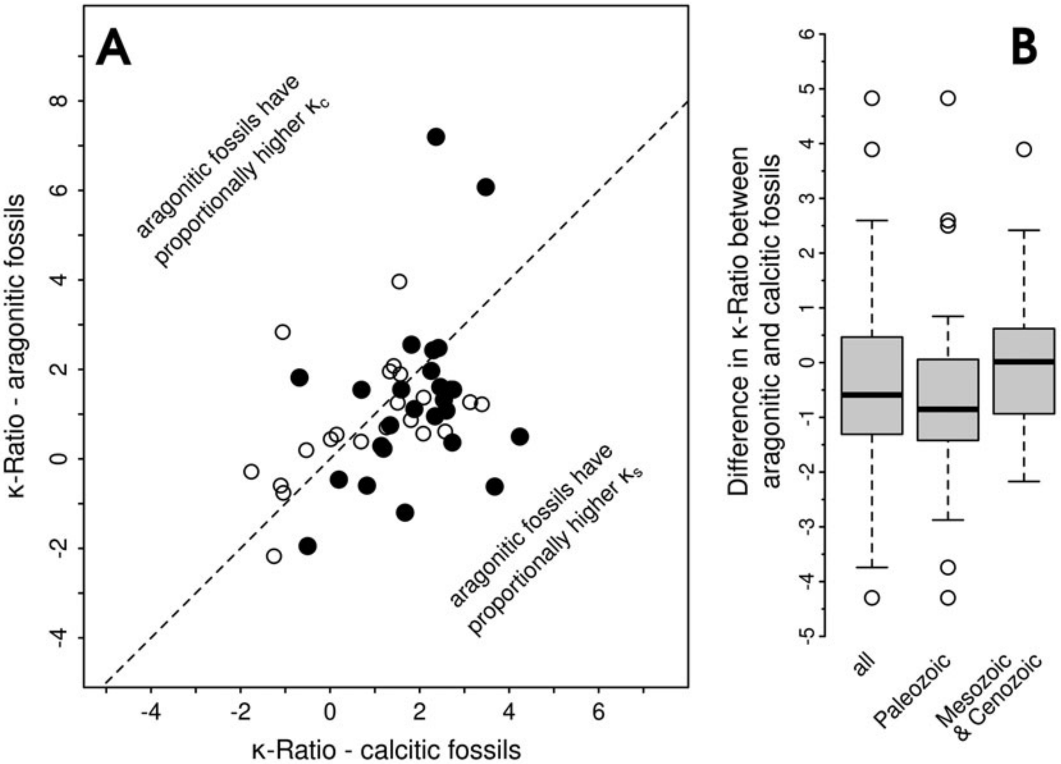

FIGURE 5. A, Scatter plot of sampled proportion ratios for calcitic and aragonitic fossils. Dashed line indicates a perfect 1:1 relationship; in the lower triangle, aragonitic fossils record lower κ-ratio (i.e., proportionally higher siliciclastic sampled proportion relative to carbonate sampled proportion) than calcitic fossils. B, Box plots of paired differences between κ-ratio estimated with aragonitic and calcitic fossils. Negative values indicate that aragonitic fossils record lower κ-ratio.

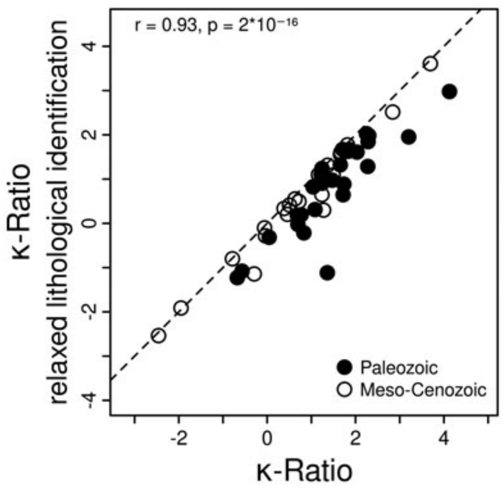

FIGURE 6. Scatter plot of raw κ-ratio and κ-ratio estimated with relaxed lithologic identification (i.e., all non–strictly carbonate fossiliferous beds coded as siliciclastic beds). Dashed line indicates a perfect 1:1 relationship. Black dots are Paleozoic intervals, white dots are post-Paleozoic intervals.

occurrences in museums probably belong to taxa already known elsewhere.

Because of the complexity of factors that affect fossil sampling and the limitations of fossil data, our analysis cannot rule out a role for any specific factor, and probably all of them contribute to some extent. Overall, however, our analysis finds little to support relative oversampling of carbonate rocks being a consequence of a greater proportion of barren siliciclastic units, greater dissolution of fossils from siliciclastic rocks, or biased data entry errors and omissions when reporting siliciclastic lithologies. We do find support that differences between carbonate and siliciclastic environments in the taxonomic distribution of occurrences may favor greater relative publication of fossils coming from carbonate environments.

Variable sampled fossiliferous proportion over time and carbonate relative oversampling could affect estimates of regional and global

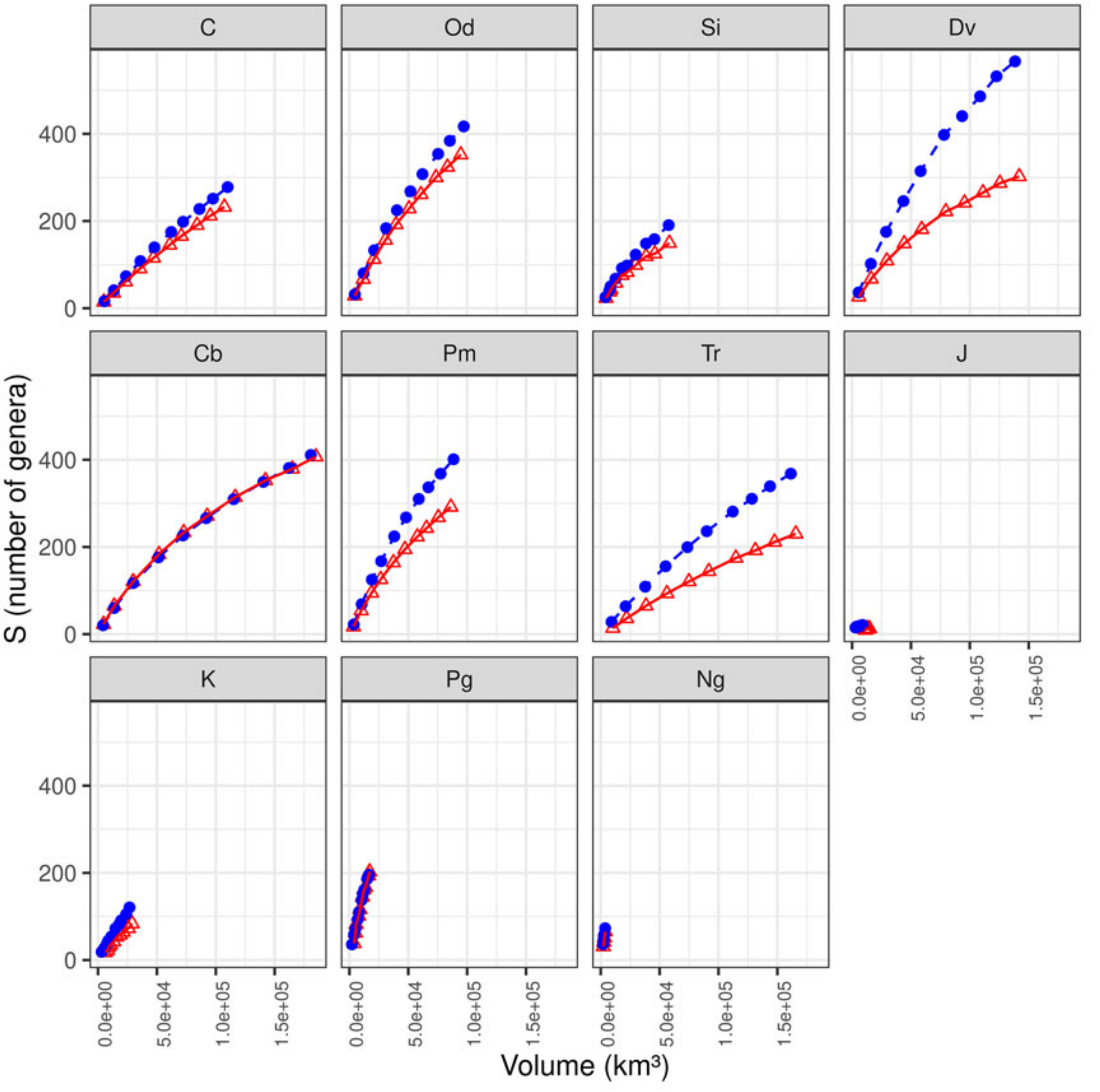

FIGURE 7. Triple-rarefaction curves for period-length bins. Dots and dashed lines are carbonate rocks, and triangles and continuous lines are siliciclastic rocks. Rarefactions account for an equal amount of Macrostrat columns, sediment volumes, and collections.

diversity. The available methods to overcome sampling intensity biases when analyzing diversity (Alroy 2020) do not solve for variable sampling coverage, which is a first-order control on diversity even after standardizing for sampling intensity (Wall et al. 2009). Therefore, if we want to comprehend diversity changes related to the fluctuating inhabited areas (e.g., Balseiro and Powell 2020), we should evaluate diversity from areas proportional to their original extent. The sampled fossiliferous proportion (κ) seems sufficiently stable (Fig. 5B) at the temporal and geographic scales of the current analysis for us to believe that there is no significant bias in the original data. However, within more specific time intervals, it may be advantageous to consider geological completeness when comparing diversity between intensively sampled intervals (e.g., Late Cretaceous) relative to poorly sampled intervals (e.g., Early Cretaceous). Moreover, differences in sampling across environments/

https://doi.org/10.1017/pab.2023.16 Published online by Cambridge University Press

regions can further skew diversity estimation (Wall et al. 2009). Because many taxa are substrate specialists (Foote 2006; Hopkins et al. 2014), considerable differences in the carbonate–siliciclastic κ-ratio can bias both composition and diversity. Therefore, it would also be good to evaluate the effect of variable κ_c and κ_s when studying biotic trends.

The Phanerozoic trend in total sampled fossiliferous proportion also bears on the issue of a marine or terrestrial origin for the documented increase sampling coverage of the fossil record as a whole, first noted by Peters and Heim (2010). Our results localize this increase to the terrestrial record, because κ of the marine fossil record shows no evident increase through the Phanerozoic, except for a rise limited to the Permian–Jurassic (Fig. 5B, Table 1). Such a trend contrasts with the results of Peters and Heim (2010), who described a rise in κ during the Late Cretaceous and a steady high level of sampling during the Cenozoic. The difference between results could be caused by new collections that have been added to the PBDB in the decade between our studies. However, κ for the marine geological record estimated using occurrences entered before 2010 is highly correlated with our current results (first differences $r = 0.87$, $p = 2 \times 10^{-16}$), indicating that new occurrences are unlikely to account for the discrepancy. Instead it appears that increasing κ of all rocks is likely limited to the continental fossil record included in Peters and Heim's (2010) analysis. Indeed, Peters and Heim (2010) already raised the possibility that the continental record was responsible for their observed increase in κ, but dismissed it due to the stable trend in the number of fossiliferous continental stratigraphic units during the post-Paleozoic. Our analysis, however, confirms that the rise is likely to be limited to the nonmarine stratigraphic record, as the marine fossil record shows a stable pattern (Fig. 4B).

Acknowledgments

This contribution was made possible thanks to the ANPCyT (MINCyT, Argentina) financial support and Juniata College's lodging facilities and assistance in Huntingdon, PA. D.B. thanks the Powell-Miller family for their generosity and warm reception. We thank associate editor A. Tomašových (Slovak Academy of Sciences), S. Holland (University of Georgia), and two anonymous reviewers for their thoughtful suggestions, which significantly improved this contribution. This is Paleobiology Database official publication no. 453.

Declaration of Competing Interests

The authors declare no competing interests.

Data Availability Statement

Additional analyses, data and R scripts used in this article are available as online Supplementary Material, which is available in the Digital repository of the Universidad Nacional de Córdoba at: http://hdl.handle.net/11086/546892.

Literature Cited

Alexandersson, E. T. 1976. Actual and anticipated petrographic effects of carbonate undersaturation in shallow seawater. Nature 262:653–657.

Aller, R. C. 1982. Carbonate dissolution in nearshore terrigenous muds: the role of physical and biological reworking. Journal of Geology 90:79–95.

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

Alroy, J. 2020. On four measures of taxonomic richness. Paleobiology 46:158–175.

Balseiro, D., and M. G. Powell. 2020. Carbonate collapse and the late Paleozoic ice age marine biodiversity crisis. Geology 48:118–122.

Behrensmeyer, A. K., S. M. Kidwell, and R. A. Gastaldo. 2000. Taphonomy and paleobiology. Paleobiology 26:103–137.

Best, M. M. R. 2008. Contrast in preservation of bivalve death assemblages in siliciclastic and carbonate tropical shelf settings. Palaios 23:796–809.

Best, M. M. R., and S. M. Kidwell. 2000. Bivalve taphonomy in tropical mixed siliciclastic-carbonate settings. I. Environmental variation in shell condition. Paleobiology 26:80–102.

Burnham, K. P., and D. R. Anderson. 2002. Model selection and multimodel inference: a practical information-theoretic approach, 2nd ed. Springer, New York.

Catuneanu, O. 2006. Principles of sequence stratigraphy. Elsevier, Amsterdam.

Cherns, L., and V. P. Wright. 2000. Missing molluscs as evidence of large-scale, early skeletal aragonite dissolution in a Silurian sea. Geology 28:791–794.

Cherns, L., and V. P. Wright. 2009. Quantifying the impacts of early diagenetic aragonite dissolution on the fossil record. Palaios 24:756–771.

Cherns, L., and V. P. Wright. 2011. Skeletal mineralogy and biodiversity of marine invertebrates: size matters more than seawater chemistry. *In* A. J. McGowan and A. B. Smith, eds. Comparing the geological and fossil records: implications for biodiversity studies. Geological Society of London Special Publication 358:9–17.

Close, R. A., S. W. Evers, J. Alroy, and R. J. Butler. 2018. How should we estimate diversity in the fossil record? Testing richness estimators using sampling-standardised discovery curves. Methods in Ecology and Evolution 9:1386–1400.

Daley, G. M., and Bush, A. M. 2020. The effects of lithification on fossil assemblage biodiversity and composition: an experimental test. Palaeontologia Electronica 23(3):a53.

Foote, M. 2006. Substrate affinity and diversity dynamics of Paleozoic marine animals. Paleobiology 32:345–366.

Foote, M., J. S. Crampton, A. G. Beu, and C. S. Nelson. 2015. Aragonite bias, and lack of bias, in the fossil record: lithological, environmental, and ecological controls. Paleobiology 41:245–265.

Hawkins, A. D., M. Kowalewski, and S. Xiao. 2018. Breaking down the lithification bias: the effect of preferential sampling of larger specimens on the estimate of species richness, evenness, and average specimen size. Paleobiology 44:326–345.

Hayek, L. and Buzas, M. 2010. Surveying natural populations: quantitative tools for assessing biodiversity, 2nd ed. Columbia University Press, New York.

Hendy, A. J. W. 2009. The influence of lithification on Cenozoic marine biodiversity trends. Paleobiology 35:51–62.

Hendy, A. J. W. 2011. Taphonomic overprints on Phanerozoic trends in biodiversity: lithification and other secular megabiases. Pp. 19–77 *in* P. A. Allison and D. J. Bottjer, eds. Taphonomy: process and bias through time. Springer, Dordrecht, Netherlands.

Holland, S. 2022. The structure of the nonmarine fossil record: predictions from a coupled stratigraphic–paleoecological model of a coastal basin. Paleobiology 48:372–396.

Holland, S., and K. M. Loughney. 2021. The stratigraphic paleobiology of nonmarine systems. Elements of Paleontology. Cambridge University Press, Cambridge.

Hopkins, M. J., C. Simpson, and W. Kiessling. 2014. Differential niche dynamics among major marine invertebrate clades. Ecology Letters 17:314–323.

Hunt, G. 2008. Gradual or pulsed evolution: when should punctuational explanations be preferred? Paleobiology 34:360–377.

Hunt, G. 2021. paleoTS: analyze paleontological time-series, R package version 0.5.3. https://cran.r-project.org/web/packages/paleoTS/index.html.

Hunt, G., M. J. Hopkins, and S. Lidgard. 2015. Simple versus complex models of trait evolution and stasis as a response to environmental change. Proceedings of the National Academy of Sciences USA 112:4885–4890.

Kidwell, S. M., M. M. R. Best, and D. S. Kaufman. 2005. Taphonomic trade-offs in tropical marine death assemblages: differential time averaging, shell loss, and probable bias in siliciclastic vs. carbonate facies. Geology 33:729–729.

Magurran, A. E. 2004. Measuring ecological diversity. Blackwell Science, Oxford, U.K.

Marshall, C. R., S. Finnegan, E. C. Clites, P. A. Holroyd, N. Bonuso, C. Cortez, E. Davis, G. P. Dietl, P. S. Druckenmiller, R. C. Engo, C. Garcia, K. Estes-Smargiassi, A. Hendy, K. A. Hollis, H. Little, E. A. Nesbitt, P. Roopnarine, L. Skibinski, J. Vendetti, and L. D. White. 2018. Quantifying the dark data in museum fossil collections as palaeontology undergoes a second digital revolution. Biology Letters 14:2–5.

Nawrot, R. 2012. Decomposing lithification bias: preservation of local diversity structure in recently cemented storm-beach carbonate sands, San Salvador Island, Bahamas. Palaios 27:190–205.

Peters, S. E. 2006. Macrostratigraphy of North America. Journal of Geology 114:391–412.

Peters, S. E. 2007. The problem with the Paleozoic. Paleobiology 33:165–181.

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

Peters, S. E., and N. A. Heim. 2010. The geological completeness of paleontological sampling in North America. Paleobiology 36:61–79.

Peters, S. E., J. M. Husson, and J. Czaplewski. 2018. Macrostrat: a platform for geological data integration and deep-time earth crust research. Geochemistry, Geophysics, Geosystems 19:1393–1409.

Powell, M. G., and M. Kowalewski. 2002. Increase in evenness and sampled alpha diversity through the Phanerozoic: comparison of early Paleozoic and Cenozoic marine fossil assemblages. Geology 30:331–334.

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

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

R Core Team. 2018. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org.

Sadler, P. M. 1981. Sediment accumulation rates and the completeness of stratigraphic sections. Journal of Geology 89:569–584.

Schovsbo, N. H. 2001. Why barren intervals? A taphonomic case study of the Scandinavian Alum Shale and its faunas. Lethaia 34:271–285.

Sessa, J. A., M. E. Patzkowsky, and T. J. Bralower. 2009. The impact of lithification on the diversity, size distribution, and recovery dynamics of marine invertebrate assemblages. Geology 37:115–118.

Smith, A. B., and A. J. McGowan. 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., and A. J. McGowan. 2008. Temporal patterns of barren intervals in the Phanerozoic. Paleobiology 34:155–161.

Smith, A. B., and A. J. McGowan. 2011. The ties linking rock and fossil records and why they are important for palaeobiodiversity studies. *In* A. J. McGowan and A. B. Smith, eds. Comparing the geological and fossil records: implications for biodiversity studies. Geological Society, London, Special Publications 358:1–7.

Tomašových, A., S. M. Kidwell, C. R. Alexander, and D. S. Kaufman. 2019. Millennial-scale age offsets within fossil assemblages: result of bioturbation below the taphonomic active zone and out-of-phase production. Paleoceanography and Paleoclimatology 34:954–977.

Wall, P. D., L. C. Ivany, and B. H. Wilkinson. 2009. Revisiting Raup: exploring the influence of outcrop area on diversity in light of modern sample-standardization techniques. Paleobiology 35:146–167.