Eliason *et al.* Supporting Information

Exceptional preservation and the fossil record of tetrapod integument

Chad M. Eliason[1,2,3,*], Leah Hudson[1], Taylor Watts[1], Hector Garza[1], Julia A. Clarke[1,2,*]

Article DOI: 10.1098/rspb

The following Supporting Information is available for this article:

Supplementary methods: Accounting for uncertainty in Lagerstätten age

Figure S1. Integrating number of Lagerstätten occurrences and proxy data over time.

Figure S2. Number of Lagerstätten is positively correlated with present-day country size.

Figure S3. Number of Lagerstätten is not explained by present-day continental area.

Figure S4. Weak temporal autocorrelation in residuals for the statistical model of Lagerstätten frequency through time.

Figure S5. Number of Lagerstätten varies with $\delta^{18}O$ (proxy for temperature) (a) and global sea levels (b).

Figure S6. Analysis using atmospheric CO_2 data from [1].

Figure S7. Source of sea level data and time bin width has negligible effect on the statistical relationship between number of Lagerstätten and environmental variables.

Figure S8. Observed and predicted number of Lagerstätten found in different depositional environments.

Figure S9. Observed and predicted number of Lagerstätten for different integument types.

Figure S10. Abiotic factors predict the number of Lagerstätten in different depositional environments.

Figure S11. Abiotic and biotic variables significantly predict integument preservation.

Figure S12. Comparative time series analysis of Lagerstätten preservation through time.

Table S1. Results of quasi-Poisson regressions fitting number of Lagerstätten found in different depositional environments as a function of sea level and formation age.

Table S2. Results of quasi-Poisson regressions fitting number of Lagerstätten found in different depositional environments as a function different abiotic and biotic predictors.

Table S3. Results of quasi-Poisson regressions fitting number of Lagerstätten by integument as a function of sea level and formation age.

Table S4. Results of stepwise multiple regressions fitting number of Lagerstätten for different integument types as a function different abiotic and biotic predictors.

Table S5. Co-distribution of integument types through time and in different Lagerstätten localities.

Table S6. Testing the effects of bin size on the relationship between the number of Lagerstätten through time and abiotic or biotic factors.

Table S7. Gap analysis results using minimum and maximum estimate of fossil age ranges.

Supplementary methods

Accounting for uncertainty in Lagerstätten age

The average age uncertainty for a Lagerstätte in our dataset was 7.0 Ma. To account for uncertainty in Lagerstätte age, we used an integration approach. Briefly, we defined time bins (5, 10, 15, 20, 25, and 30 Ma) and used the ranges of uncertainty in fossil age to determine the probability a fossil belongs to a given bin (e.g., for a fossil with a well-bracketed age estimate the probability it belongs to a given time bin would by 1, while a fossil with an age estimate spanning 4 time bins would have a 0.25 probability of belonging to any given time bin; see supplementary figure S1 for details). After determining the probability that each Lagerstätte was found within a given time bin, the proportions for individual bins were summed. R code used for this analysis is available at Dryad (doi:10.5061/dryad.1c5k9).

Supplementary Figures

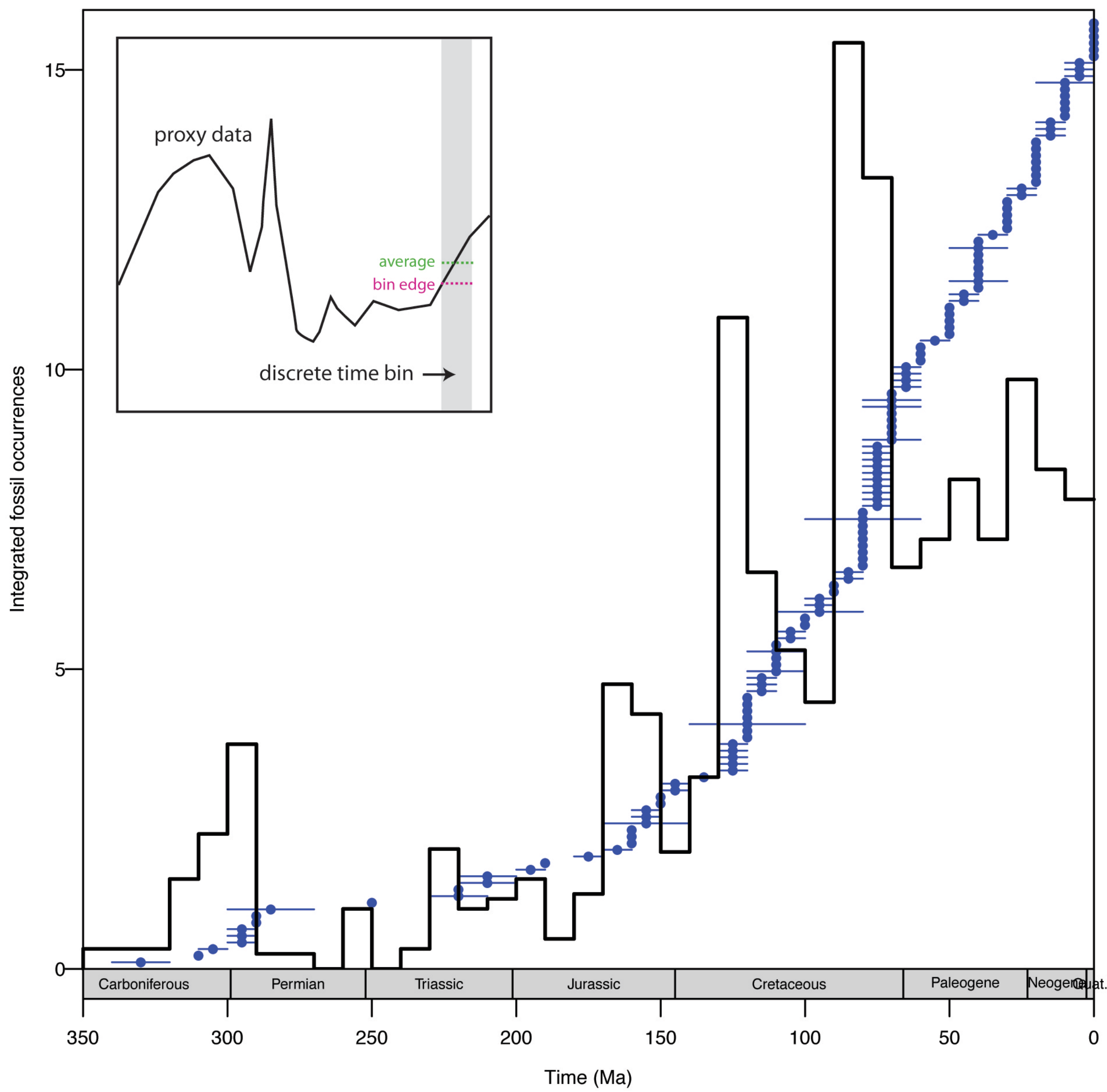

Figure S1. Integrating number of Lagerstätten occurrences and proxy data over time. Graph shows individual Lagerstätte median ages (blue points) with associated error (horizontal lines). Black line is the number of Lagerstätten (kernel density estimate) integrated over each 10-My time bin to account for uncertainty in age. Inset shows two approaches for integrating sample proxy data over discrete time intervals. Method 1 interpolates data along specific time intervals (light green) and method 2 averages the proxy data over the entire intervals (dark green points).

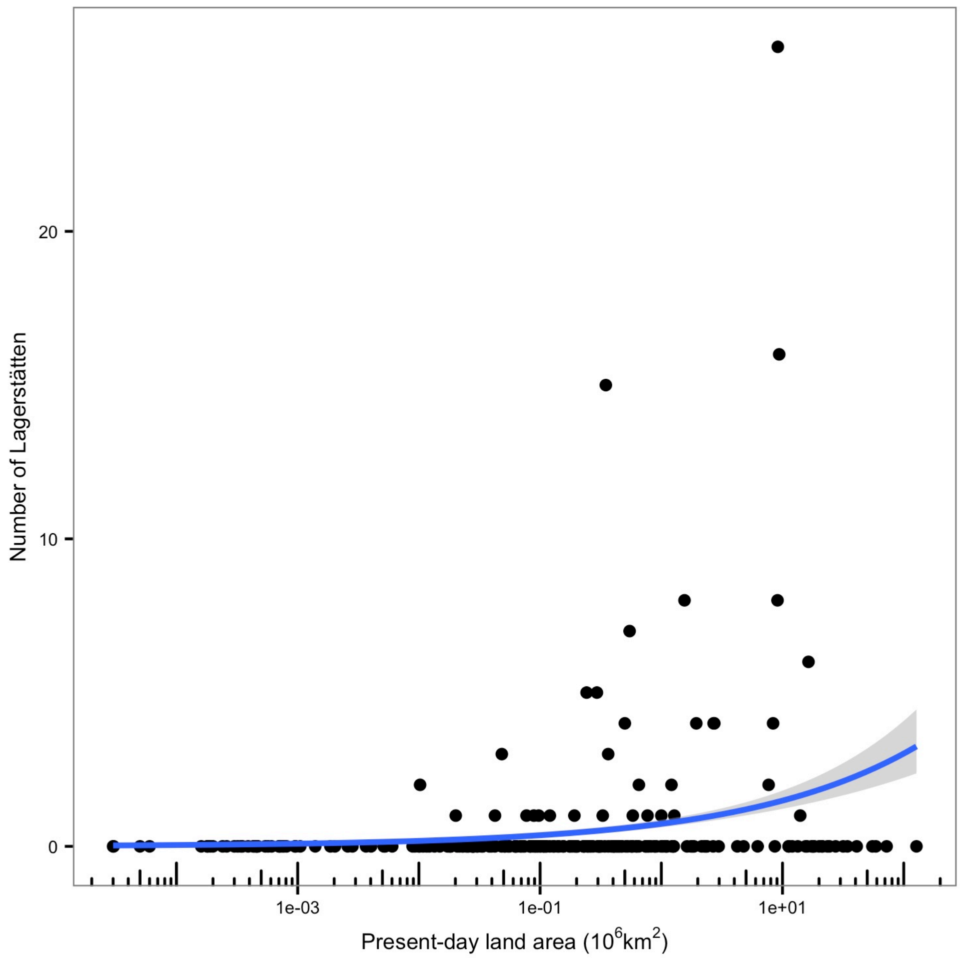

Figure S2. Number of Lagerstätten is positively correlated with present-day country size. Points are total number of Lagerstätten found in each country. Dashed line is Poisson regression fit. The relationship was statistically significant ($\beta = 0.31 \pm 0.03$ SE, $z = 9.62$, $p < 0.001$, $n = 241$).

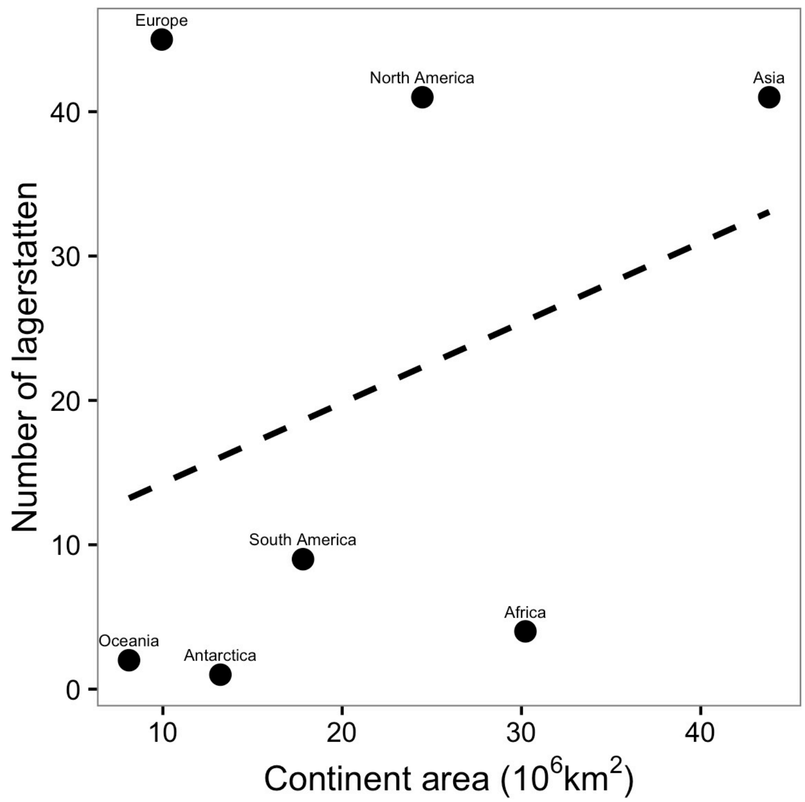

Figure S3. Number of Lagerstätten is not explained by present-day continental area. Points are total number of Lagerstätten found in each continent. Dashed line is linear regression fit. The relationship was not statistically significant (Pearson's $r = 0.34$, $p = 0.45$; Spearman's $rho = 0.25$, $p = 0.59$).

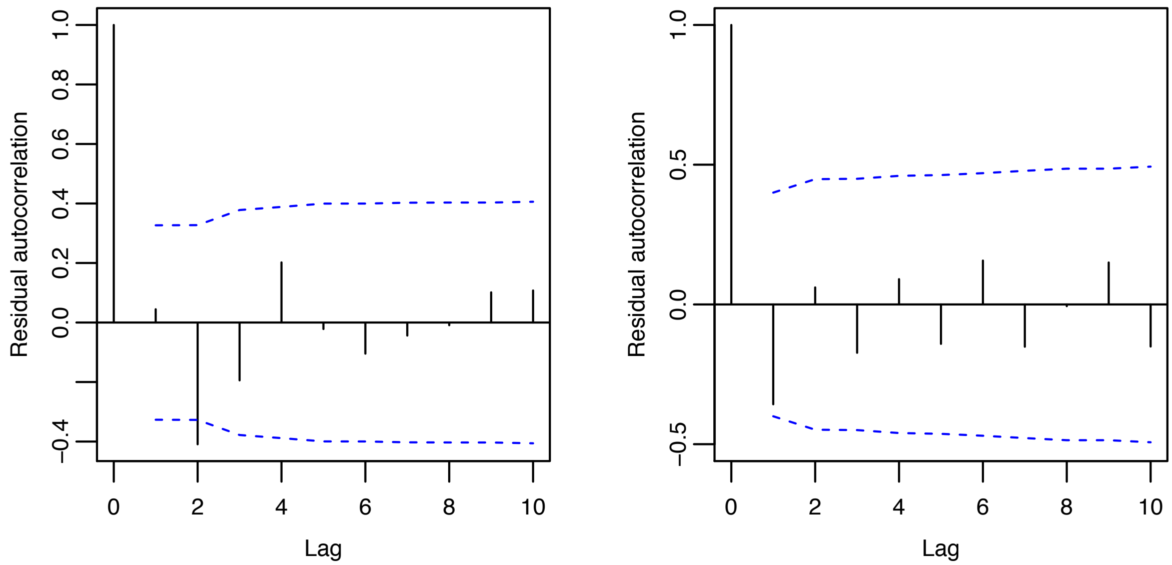

Figure S4. Weak temporal autocorrelation in residuals for the statistical model of Lagerstätten frequency through time. Plots shows autocorrelation among residuals of the model (including time and sea level as predictor variables) at different time lags for time bin sizes of 10 My (left) and 20 My (right). Dashed blue lines are 5% significance limits: black vertical bars extending outside the blue lines are significantly autocorrelated. Most time lags showed a non-significant amount of autocorrelation, and the overall amount of autocorrelation for bin size = 10 My was non-significant (Breusch-Godfrey test, p = 0.17). Lag = 2 in left panel shows significant negative autocorrelation, meaning that data points 20 My apart are negatively correlated. However, this effect disappeared with wider time bins (right panel).

(a) (b)

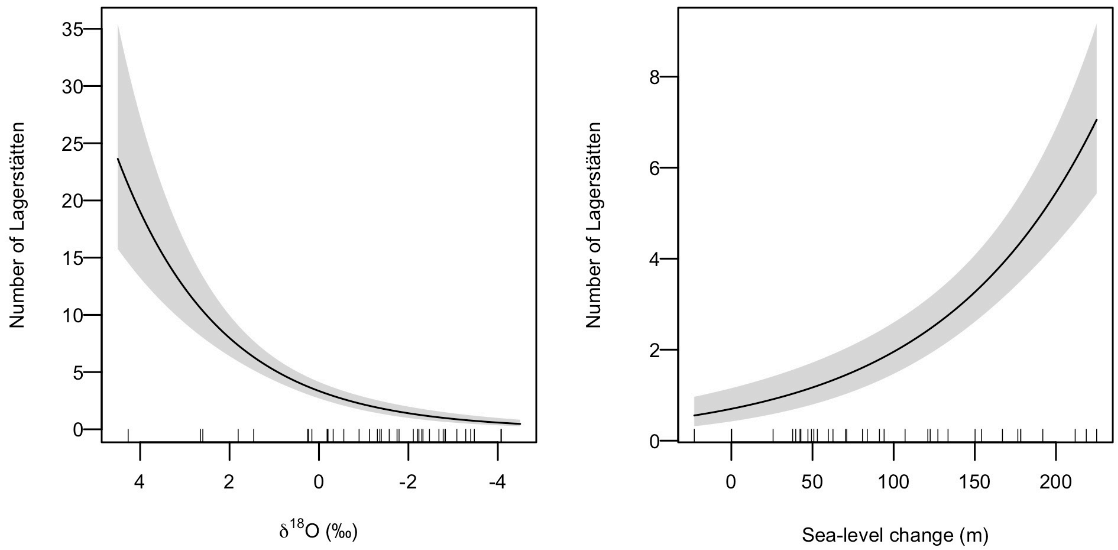

Figure S5. Number of Lagerstätten varies with $\delta^{18}O$ (proxy for temperature) (a) and global sea levels (b). Lines are conditional plots showing expected number of preserved Lagerstätten as a function of each predictor variable holding other variables constant at their median values. Overall quasi-Poisson model was highly significant ($P < 0.001$). Results are based on sea level datasets from [1-3]. Note different scale in y-axes.

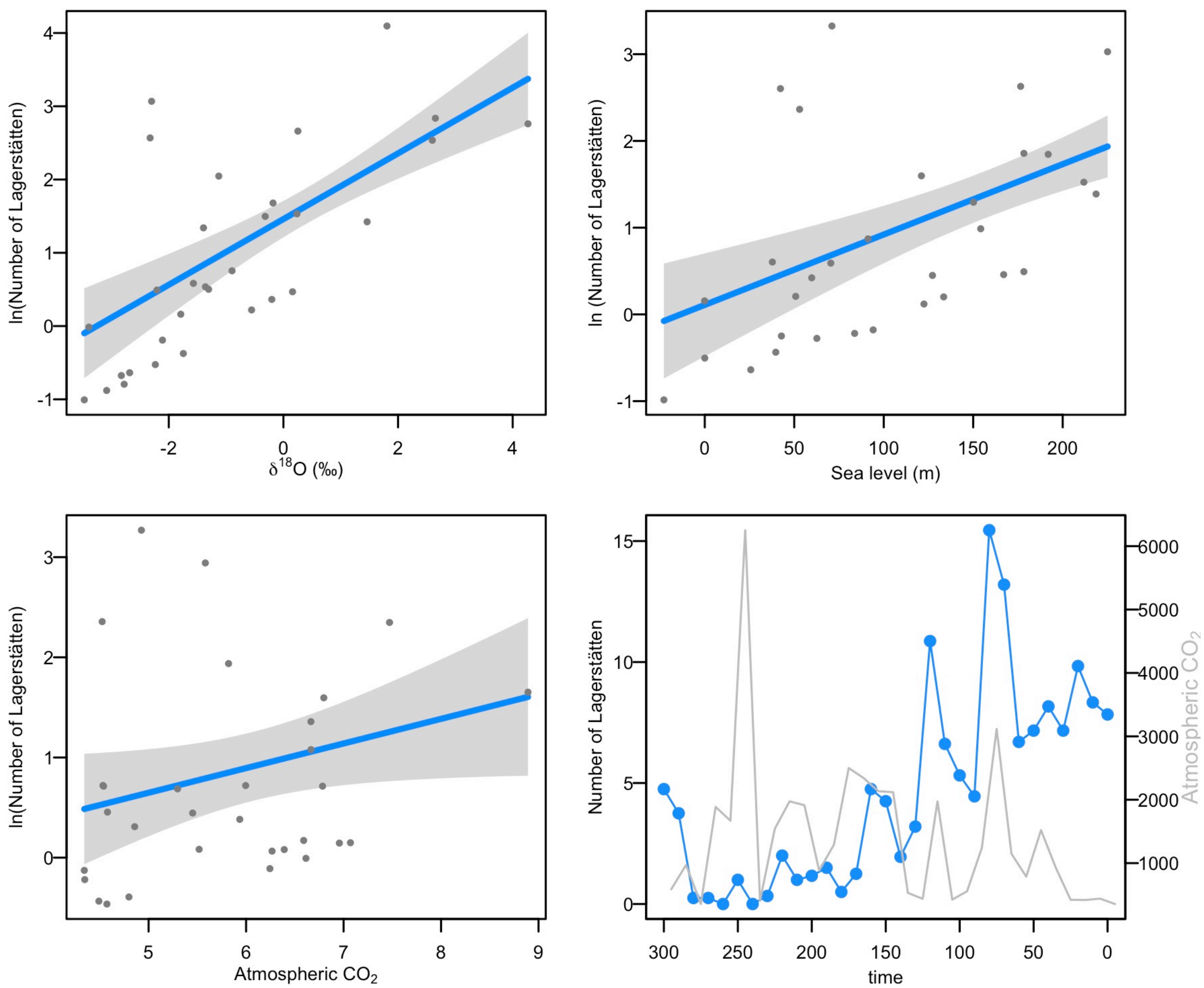

Figure S6. Analysis using atmospheric CO_2 data from [1]. CO_2 data was fourth-root transformed [4] to meet assumptions of normality. Only $\delta^{18}O$ and sea level significantly predicted the number of Lagerstätten through time. The lack of support for CO_2—a previously suggested driver of Lagerstätten in marine systems [2,3,5]—is likely because of the high accumulation of Lagerstätten in the Cenozoic despite decreasing CO_2 levels.

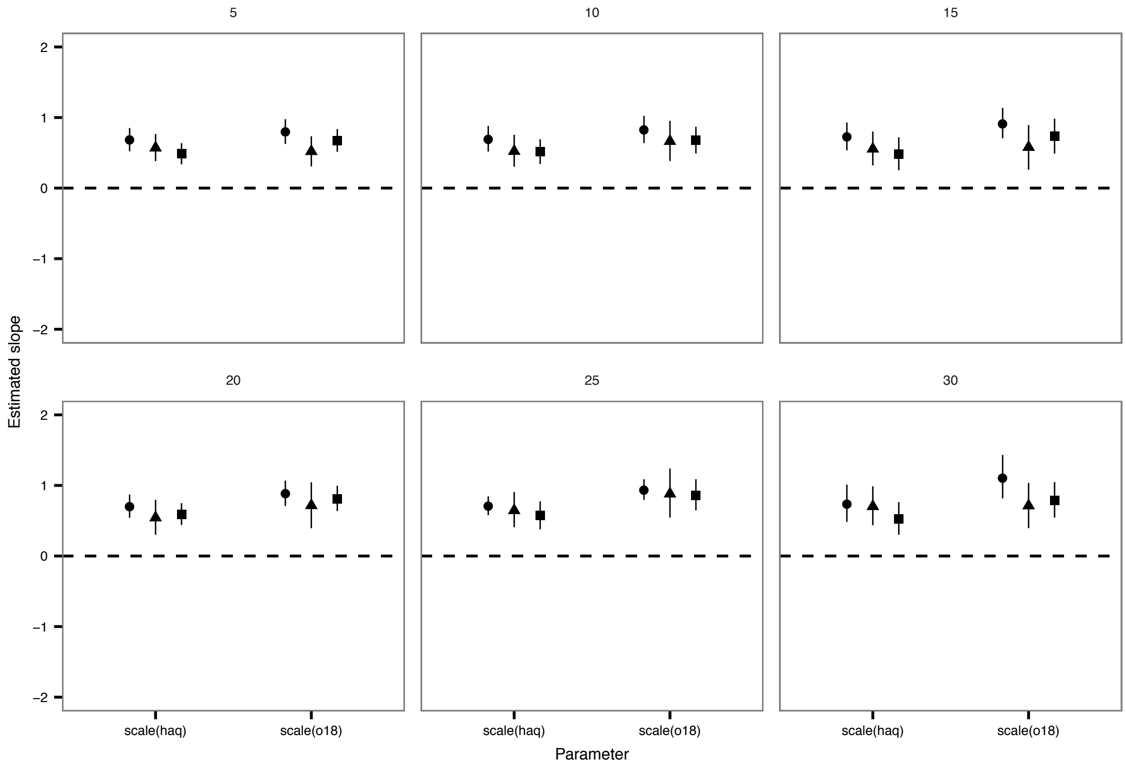

Figure S7. Source of sea level data and time bin width has negligible effect on the statistical relationship between number of Lagerstätten and environmental variables. Points show quasi-Poisson regression parameter estimates for various datasets [circles: Vail 1977 dataset [6]; triangles: full composite dataset of Haq et al. [2,3]; squares: truncated dataset of Haq et al. 1987 [2], with sea level data only for the last 250 My]. Vertical lines are 95% confidence intervals. Panels show results for different time bin widths (5, 10, 15, 20, 25, and 30 My); the mean error in age in our fossil database was 7.0 My. Both parameters ($\delta^{18}O$ and sea level change) were significantly different from zero for all datasets and time bin sizes (i.e. confidence intervals do not overlap zero, dashed horizontal lines).

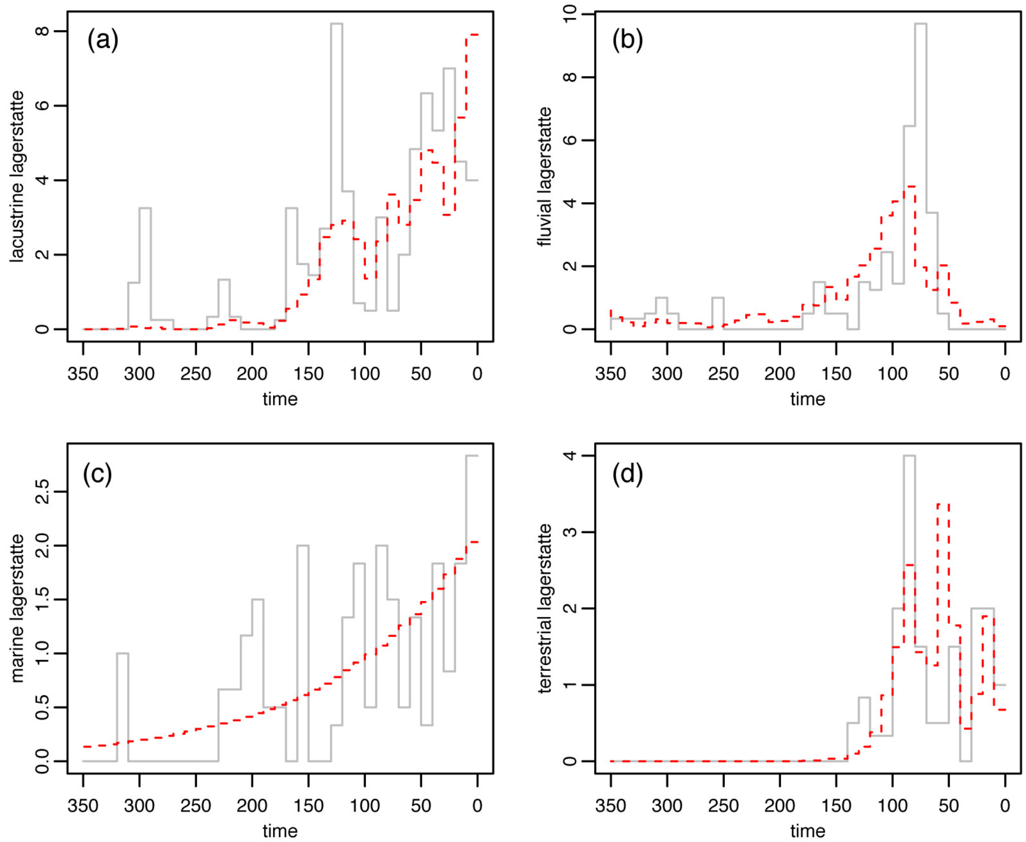

Figure S8. Observed and predicted number of Lagerstätten found in different depositional environments. Plots show actual (grey lines) and predicted number of Lagerstätten occurrences (dashed red lines) in 10-My bins under the best-fitting multiple regression models (see table S2).

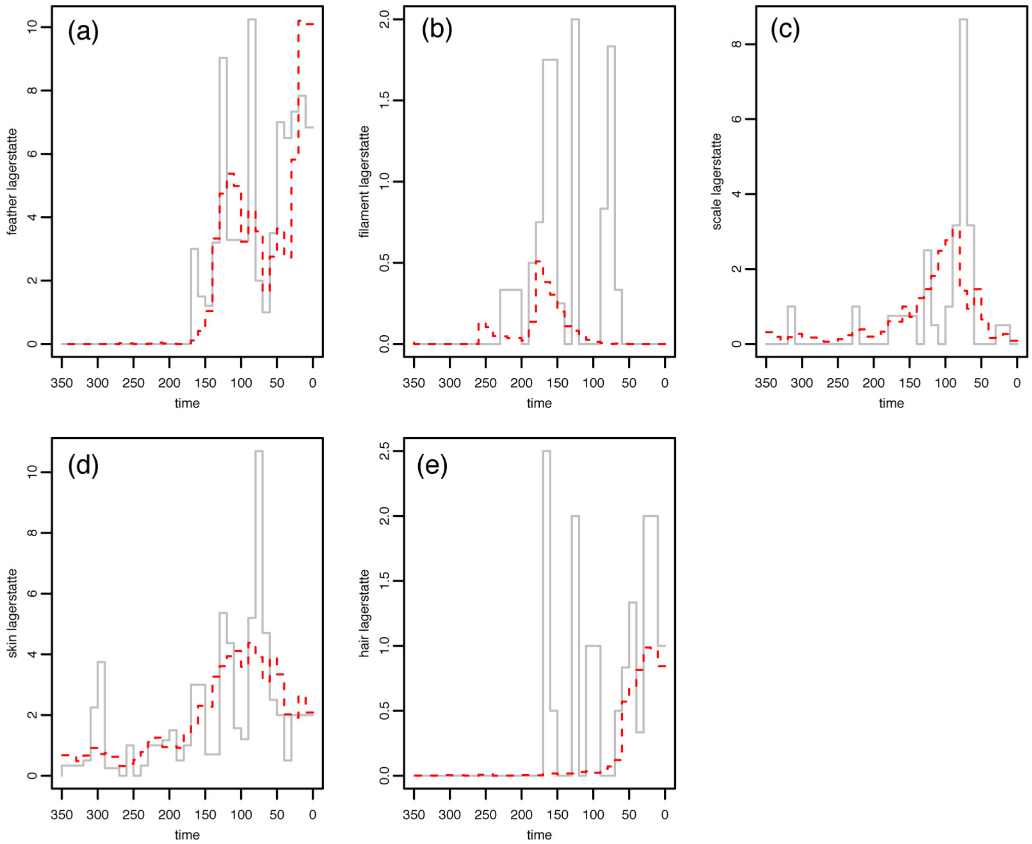

Figure S9. Observed and predicted number of Lagerstätten for different integument types. Plots show actual (grey lines) and predicted number of Lagerstätten (dashed red lines) in 10-My bins under the best-fitting multiple regression model (see table S4).

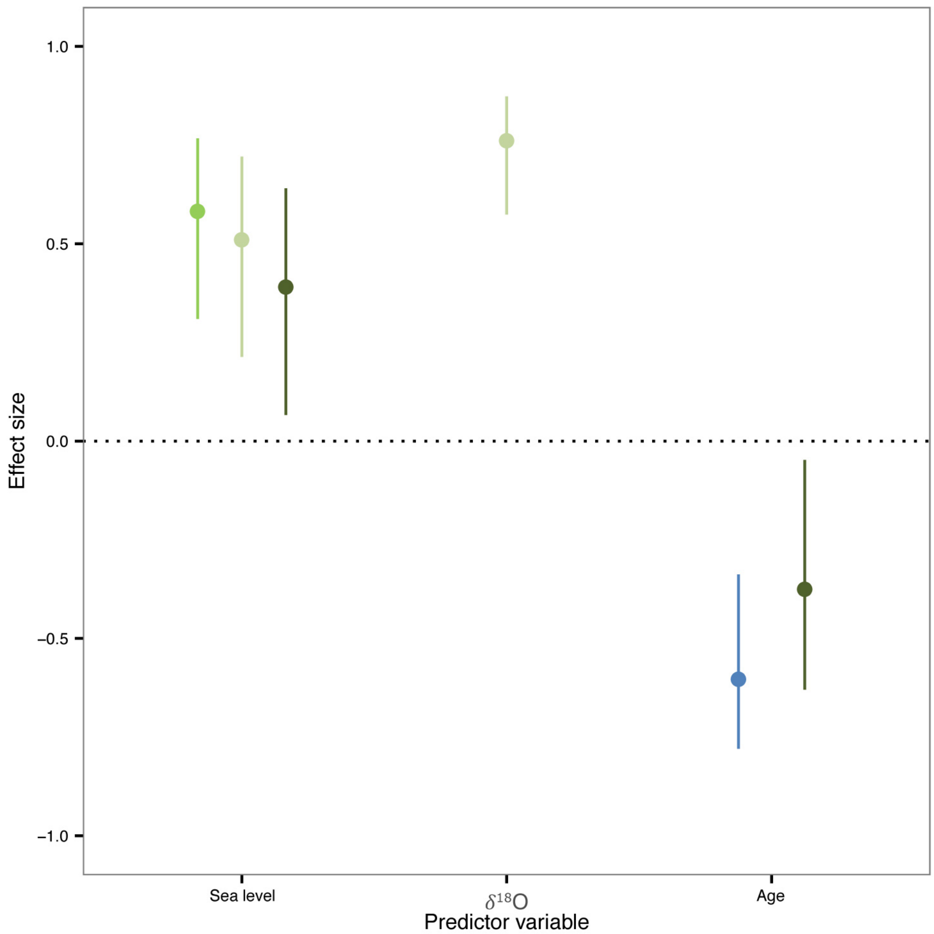

Figure S10. Abiotic factors predict the number of Lagerstätten in different depositional environments. Lines show standardized effect sizes (points) and 95% confidence intervals (vertical lines) calculated following [7]. Colours correspond to depositional environments: estuarine/near-shore marine (blue), lacustrine (light green), fluvial (medium green), and other terrestrial (dark green). Only significant variables are shown: global sea levels (m), $\delta^{18}O$ values, and Lagerstätten age (Ma).

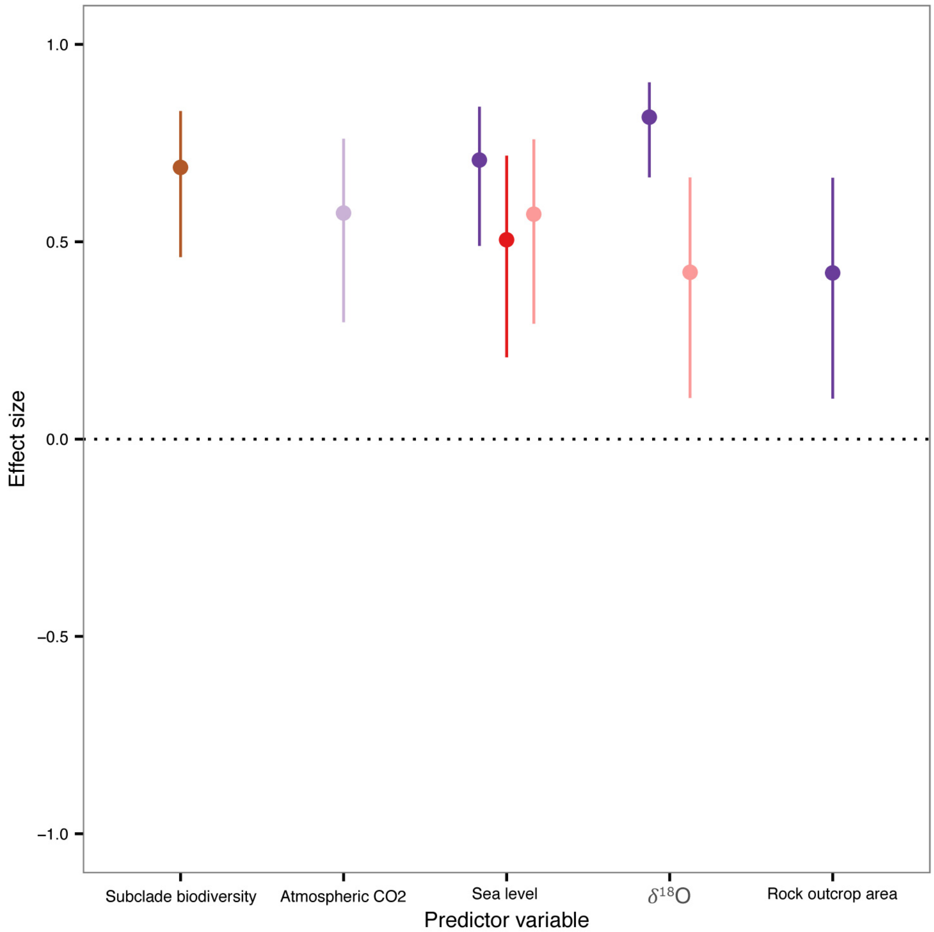

Figure S11. Abiotic and biotic variables significantly predict integument preservation. Lines show standardized effect sizes (points) and 95% confidence intervals (vertical lines) calculated following [7]. Colours correspond to integument type: feathers (dark purple), filaments (light purple), scales (red), skin (pink), and hair (brown). Only significant variables are shown: subclade biodiversity, atmospheric carbon dioxide levels, global sea levels (m), $\delta^{18}O$ values, and terrestrial rock outcrop area.

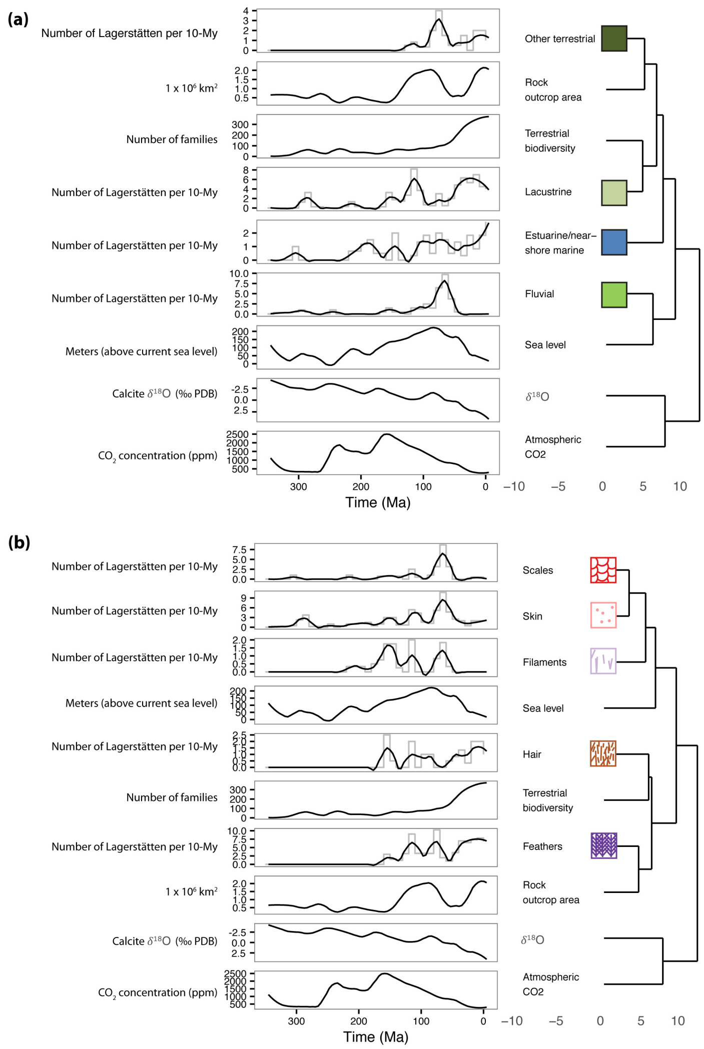

Figure S12. Comparative time series analysis of Lagerstätten preservation through time. Panels show time series of paleoenvironmental data and integrated number of Lagerstätten for different depositional environments (a) and integument types (b). Dendrograms show hierarchical clustering of time series data with Euclidean distances on the x-axis (e.g., two variables directly connected by a short path indicate similar trends with time). Variables were scaled to have a standard deviation of unity prior to computing Euclidean distances and running a clustering analysis. Lagerstätten counts were integrated over 10-My bins (see supplementary figure S1). Sea level estimates from composite of [2] and [3], as presented in [8]. Global terrestrial rock outcrop area from [9]. Atmospheric CO_2 data taken from [10] and represents an average of global proxies (see figure S6 for an alternative CO_2 dataset taken from [1]).

Supplementary Tables

Table S1. Results of quasi-Poisson regressions fitting number of Lagerstätten found in different depositional environments as a function of sea level and formation age. Numbers are estimated coefficients and standard errors in parentheses. Models were fit using significant terms from the most parsimonious pooled model (sea level and $\delta^{18}O$). Significant effects ($p < 0.05$) indicated in bold.

	Lacustrine	**Fluvial**	**Marine**	**Terrestrial**
(Intercept)	-0.001	**-2.281****	**-0.722***	**-3.808****
	(0.350)	(0.761)	(0.321)	(1.069)
$\delta^{18}O$	**0.462*****	0.108	**0.355*****	**0.909****
	(0.087)	(0.189)	(0.081)	(0.258)
Sea level	**0.007****	**0.017*****	**0.005***	**0.023*****
	(0.002)	(0.004)	(0.002)	(0.006)
Deviance	40.999	43.845	20.851	15.131
Num. obs.	36	36	36	36

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table S2. Results of quasi-Poisson regressions fitting number of Lagerstätten found in different depositional environments as a function different abiotic and biotic predictors. Numbers are estimated coefficients and standard errors in parentheses. Models were fit using backward stepwise regression, dropping non-significant terms (p > 0.05) from the model. Significant effects (p < 0.05) indicated in bold.

	Lacustrine	**Fluvial**	**Marine**	**Terrestrial**
(Intercept)	**0.781*****	**-2.393****	**0.709****	**-5.395***
	(0.118)	(0.734)	(0.238)	(2.553)
$\delta^{18}O$	**0.210*****			
	(0.031)			
Sea level	**0.003****	**0.017*****		**1.391***
	(0.001)	(0.004)		(0.579)
Time			**-0.008*****	**-4.318***
			(0.002)	(1.886)
Zero model: (Intercept)				-1.460
				(1.055)

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table S3. Results of quasi-Poisson regressions fitting number of Lagerstätten by integument as a function of sea level and formation age. Numbers are estimated coefficients and standard errors in parentheses. Models were fit using significant terms from the most parsimonious pooled model (sea level and $\delta^{18}O$). Significant effects ($p < 0.05$) indicated in bold.

	Skin	**Scales**	**Feathers**	**Filaments**	**Hair**
(Intercept)	-0.2051	**-2.3318****	**-1.0865***	**-2.6516****	**-1.7862***
	(0.3201)	(0.7984)	(0.4836)	(0.7381)	(0.6939)
$\delta^{18}O$	**0.2158***	0.3162	**0.7628*****	0.0122	**0.5313****
	(0.0815)	(0.1980)	(0.1181)	(0.1873)	(0.1718)
Sea level	**0.0086*****	**0.0164****	**0.0155*****	**0.0120****	0.0091
	(0.0021)	(0.0048)	(0.0030)	(0.0042)	(0.0047)
Deviance	36.9873	41.6069	33.3385	22.9341	22.1529
Num. obs.	36	36	36	36	36

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table S4. Results of stepwise multiple regressions fitting number of Lagerstätten for different integument types as a function different abiotic and biotic predictors. Numbers are estimated coefficients and standard errors in parentheses. Models were fit using backward stepwise regression, dropping non-significant terms ($p > 0.05$) from the model. Significant effects ($p < 0.05$) are indicated in bold.

	Skin	**Scales**	**Feathers**	**Filaments**	**Hair**
(Intercept)	-0.2051	**-2.4317****	**0.3380****	-0.1072	-0.2124
	(0.3201)	(0.7984)	(0.1185)	(0.1201)	(0.1172)
$\delta^{18}O$	**0.2158***		**0.2285*****		
	(0.0815)		(0.0286)		
Sea level	**0.0086*****	**0.0158****	**0.0042*****		
	(0.0021)	(0.0046)	(0.0007)		
Rock outcrop area			**0.2448***		
			(0.0931)		
CO_2				**0.0004*****	
				(0.0001)	
Synapsid biodiversity					**0.3228*****
					(0.0583)

***$p < 0.001$, **$p < 0.01$, *$p < 0.05$

Table S5. Co-distribution of integument types through time and in different Lagerstätten localities. Table shows binary correlations (Phi statistic) for co-occurrence of integument types in a given Lagerstätten (lower triangle) and detrended time series correlations (Spearman's rho; upper triangle). Numbers in bold are significant (Bonferroni adjusted $p < 0.05$), as assessed with Chi-square tests (lower triangle) or Spearman rank sum correlation tests (upper triangle).

	Feathers	Filaments	Skin	Scales	Hair
Feathers		0.06	0.19	0.14	**0.61**
Filaments	-0.06		**0.58**	**0.71**	0.10
Skin	**-0.55**	0.15		**0.51**	0.13
Scales	**-0.37**	0.07	**0.31**		0.15
Hair	0.03	0.16	0.04	-0.10	

Table S6. Testing the effects of bin size on the relationship between the number of Lagerstätten through time and abiotic or biotic factors. Each statistical model (see tables S2, S4) was re-fit using resampled proxy and fossil occurrence data at different bin sizes (5 - 30, in 5 My increments). Predictors marked "n.s." indicate variables that were no longer significant at a given bin size. At bin sizes larger than the mean resolution of our fossil occurrence data (i.e. > 10 My), some variables were no longer significant.

			bin size (My)					
	response variable	predictor variable	5	10	15	20	25	30
paleoenvironment	near-shore marine	formation age						
	lacustrine	sea level						
		delta-O-18						
	fluvial	sea level						n.s.
	terrestrial	sea level					n.s.	n.s.
		formation age						
integument type	skin	delta-O-18						
		sea level						
	scales	sea level						n.s.
	feathers	delta-O-18						
		sea level						
		rock outcrop area			n.s.		n.s.	n.s.
	filaments	atmospheric CO2						
	hair	Synapsid biodiversity						

Table S7. Gap analysis results using minimum and maximum estimate of fossil age ranges. Values show upper and lower ends of 50% confidence intervals for different integument types based on gap analysis using three different approaches. Confidence intervals were calculated based on minimum and maximum of age ranges for individual Lagerstätten (maximum values in parentheses). All values are in Ma.

Integument type	N	Upper	Lower		
			Strauss and Sadler 1989, Marshall 1990	Marshall 1994	Wang et al. 2016
Feathers	80	164 (165)	165 (167)	164 – 166 (166 – 167)	175 (172)
Filaments	11	209 (227)	219 (238)	212 - 243 (230 - 261)	253 (283)
Skin	67	331 (347)	335 (351)	332 – 334 (348 – 351)	388 (394)
Scales	25	316 (317)	325 (325)	316 – 330 (317 – 329)	420 (419)
Hair	15	164 (165)	172 (173)	165 – 189 (166 – 190)	204 (195)

Supplementary References

1. Retallack, G. J. 2009 Greenhouse crises of the past 300 million years. *Geological Society of America Bulletin* **121**, 1441–1455. (doi:10.1130/B26341.1)

2. Haq, B. U., Hardenbol, J. & Vail, P. R. 1987 Chronology of fluctuating sea levels since the triassic. *Science* **235**, 1156–1167. (doi:10.1126/science.235.4793.1156)

3. Haq, B. U. & Al-Qahtani, A. M. 2005 Phanerozoic cycles of sea-level change on the Arabian Platform. *GeoArabia* **10**, 127–160.

4. Zar, J. H. 1999 Biological statistics. Englewood Cliffs, NJ: Prentice-Hall.

5. Retallack, G. J. 2011 Exceptional fossil preservation during CO_2 greenhouse crises? *Palaeogeography, Palaeoclimatology, Palaeoecology* **307**, 59–74. (doi:10.1016/j.palaeo.2011.04.023)

6. Vail, P. R., Mitchum, R. M., Jr & Thompson, S., III 1977 Seismic stratigraphy and Global Changes of Sea Level: Part 4. Global Cycles of Relative Changes of Sea Level. *American Association of Petroleum Geologists Memoir*, **27**, 83–97.

7. Nakagawa, S. & Cuthill, I. C. 2007 Effect size, confidence interval and statistical significance: a practical guide for biologists. *Biol. Rev. Camb. Philos. Soc.* **82**, 591–605. (doi:10.1111/j.1469-185X.2007.00027.x)

8. Miller, K. G. et al. 2005 The Phanerozoic record of global sea-level change. *Science* **310**, 1293–1298. (doi:10.1126/science.1116412)

9. Wall, P. D., Ivany, L. C. & Wilkinson, B. H. 2011 Impact of outcrop area on estimates of Phanerozoic terrestrial biodiversity trends. *Geological Society, London, Special Publications* **358**, 53–62. (doi:10.1144/SP358.5)

10. Berner, R. A. 2006 GEOCARBSULF: A combined model for Phanerozoic atmospheric O_2 and CO_2. *Geochimica et Cosmochimica Acta* **70**, 5653–5664. (doi:10.1016/j.gca.2005.11.032)