ECOLOGY

Diversity dependence is a ubiquitous phenomenon across Phanerozoic oceans

Valentin Rineau[1]*, Jan Smyčka[1], David Storch[1,2]

Copyright © 2022 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. Distributed under a Creative Commons Attribution NonCommercial License 4.0 (CC BY-NC).

Biodiversity on Earth is shaped by abiotic perturbations and rapid diversifications. At the same time, there are arguments that biodiversity is bounded and regulated via biotic interactions. Evaluating the role and relative strength of diversity regulation is crucial for interpreting the ongoing biodiversity changes. We have analyzed Phanerozoic fossil record using public databases and new approaches for identifying the causal dependence of origination and extinction rates on environmental variables and standing diversity. While the effect of environmental factors on origination and extinction rates is variable and taxon specific, the diversity dependence of the rates is almost universal across the studied taxa. Origination rates are dependent on instantaneous diversity levels, while extinction rates reveal delayed diversity dependence. Although precise mechanisms of diversity dependence may be complex and difficult to recover, global regulation of diversity via negative diversity dependence of lineage diversification seems to be a common feature of the biosphere, with profound consequences for understanding current biodiversity crisis.

Downloaded from https://www.science.org at Universite Pierre Et Marie Curie - Paris 6 (Upmc) on October 31, 2022

INTRODUCTION

Understanding biodiversity dynamics, its regulation, and mechanisms of biodiversity maintenance is crucial for our ability to interpret current biodiversity changes. There has been a longstanding debate whether biological diversity on Earth during the Phanerozoic has steadily increased or whether it is bounded and regulated by biotic interactions (*1*–*3*). Biodiversity dynamics might have been influenced by abiotic factors such as environmental changes (*4*) or tectonics (*5*); by biotic interactions such as competition, leading to diversity dependence (higher diversity levels lead to decreasing origination or increasing extinction rates and vice versa) (*6*); or by a mixture of both abiotic and biotic effects (*7*). The entanglement of biotic and abiotic factors makes it essential to study these components together (*8*). Here, we evaluate the relative strength of the factors driving the large-scale biodiversity dynamics (*9*) across different taxa.

We focus on well-sampled marine fossil datasets through the Phanerozoic. The immense effort to fill fossil occurrence databases over the past decades has allowed us to analyze more complete biodiversity time series corrected using subsampling (*10*) and recent rate estimator (*11*) methods. In contrast to macroevolutionary studies that have typically focused on a single explanatory factor explored across short periods of time, we adopted a multifactorial approach to (i) reveal the universality of the processes responsible for diversity regulation across different taxonomic groups and time scales and (ii) explore causal links for biotic and abiotic factors simultaneously. Specifically, we tested for causal relationships between diversity, extinction rate, origination rate, temperature, and surrogates for productivity ($\delta^{13}C$) and nutrient input ($^{87}Sr/^{86}Sr$, $\delta^{34}S$) (figs. S4 and S5 and table S1) (*12*) coming from analyses on fossils and sedimentary rocks (*13*).

Diversity patterns in fossil time series may be affected by a complex interplay between gradual environmental changes, rapid disturbances causing mass extinctions, and biotic interactions of varying strength, which may be modulated by the environment itself. Under these circumstances, recovering causal links is notoriously difficult, and simple correlation analyses are typically insufficient (*8*). For this reason, we have adopted two novel statistical methods, recently developed for decomposing these complex interactions. Specifically, we tested for causal linkages using convergent cross-mapping (CCM) (*14*) and conditional transfer entropy (CTE) (*15*). These two methods detect causality between time series in the sense that one time series can be used for forecasting another one. They do not recourse to particular mechanisms of diversity dependence, neither to any particular functional links between the variables. This on one hand allows revealing the existence of diversity dependence and other, often nonlinear causal links using noisy paleontological data in nonparametric manner, but on the other hand, it precludes more detailed insights into the particular mechanisms of diversity dependence.

CCM is a method based on the theory of dynamical systems, which allows one to highlight nonlinear relationships and specify the directionality of causal relationships (*16*). It reveals the temporal signature of the relationship, specifically whether the interactions are synchronous or time-delayed and how long-lasting the effects are. In contrast, the specificity of CTE analysis is that it measures the direct causal link between two variables, understood as directional information flow between them, while explicitly taking into account the potential effect of additional conditioning variables. Adding conditioning variables allowed us to detect synergetic effects and to disentangle complex relationships because of the common driver effect (when $a \rightarrow b$ and $a \rightarrow c$, spurious causal links could be detected between b and c) (*17*). Although neither of the methods provides the ultimate test of causality (this is virtually impossible for observational data, without experiments), their focus on information flow and predictability allows recovering potential causal relationships even in the messy paleontological data with nonlinear causal links and variable time delays (*14*, *17*). Combination of both the methods may then illuminate the basic features of the causal structure of the

[1]Center for Theoretical Study, Charles University and the Academy of Sciences of the Czech Republic, Jilská 1, 110 00 Prague, Czech Republic. [2]Department of Ecology, Faculty of Science, Charles University, Viničná 7, 128 44 Prague, Czech Republic.
*Corresponding author. Email: valentin.rineau@sorbonne-universite.fr

system even when the exact mechanisms of diversity regulation are difficult or even impossible to be fully recovered.

RESULTS

Nine taxonomic datasets were selected, four from the Neptune Database (NDB) (diatoms, Coccolithophoridae, Foraminifera, and Radiolaria) (*18*) and five from the Paleobiology Database (PBDB) (https://paleobiodb.org) (Bivalvia, Gastropoda, Brachiopoda, Scleractinia, and a wider set of all marine metazoans). We analyzed the taxonomic datasets at generic rank [see (*12*) for sensitivity analyses using other taxonomic ranks], aggregated at a global scale (figs. S1 to S3).

Pairwise CCM analyses (Fig. 1, Table 1, figs. S6 and S7, and data S1) revealed that some causal links hold consistently among the explored taxa. In general, biotic factors play a more important role than abiotic factors in determining the long-term diversity dynamics. Both extinction and origination rates are diversity dependent in almost all datasets, while abiotic factors reveal no general pattern; instead, they are contingent and specific to each taxon. The additional autoregressive vector analysis suggests that the effect of diversity on extinction rate is either positive or nonmonotonic but never negative (i.e., increasing diversity never leads to lower extinction rates). Similarly, the effect of diversity on origination rate is always either negative or nonmonotonic (Table 1 and figs. S8 to S15). This is in agreement with the idea of equilibrium dynamics, which states that increasing diversity leads to the decrease of diversification (origination minus extinction) rate (*1*, *9*). All these effects are more pronounced for the Neptune datasets (Fig. 1, F to I), probably because of the shorter time steps and less biased sampling, and on the metazoan dataset, apparently because of its larger amount of sampled taxa. CCM analysis also revealed latency (several million years) of the effect of diversity on extinction rate in diatoms, radiolarians, foraminiferans, and coccoliths (Fig. 2), stressing the complexity of the mechanisms generating the diversity dependence.

The CCM analyses show only few significant causal links for scleractinian corals, bivalves, and brachiopods (Fig. 1, B to D), with no diversity dependence for brachiopods. It is possible that some of these nonconclusive results are due to the bivariate nature of the CCM method, not allowing the capture of more complex interactions

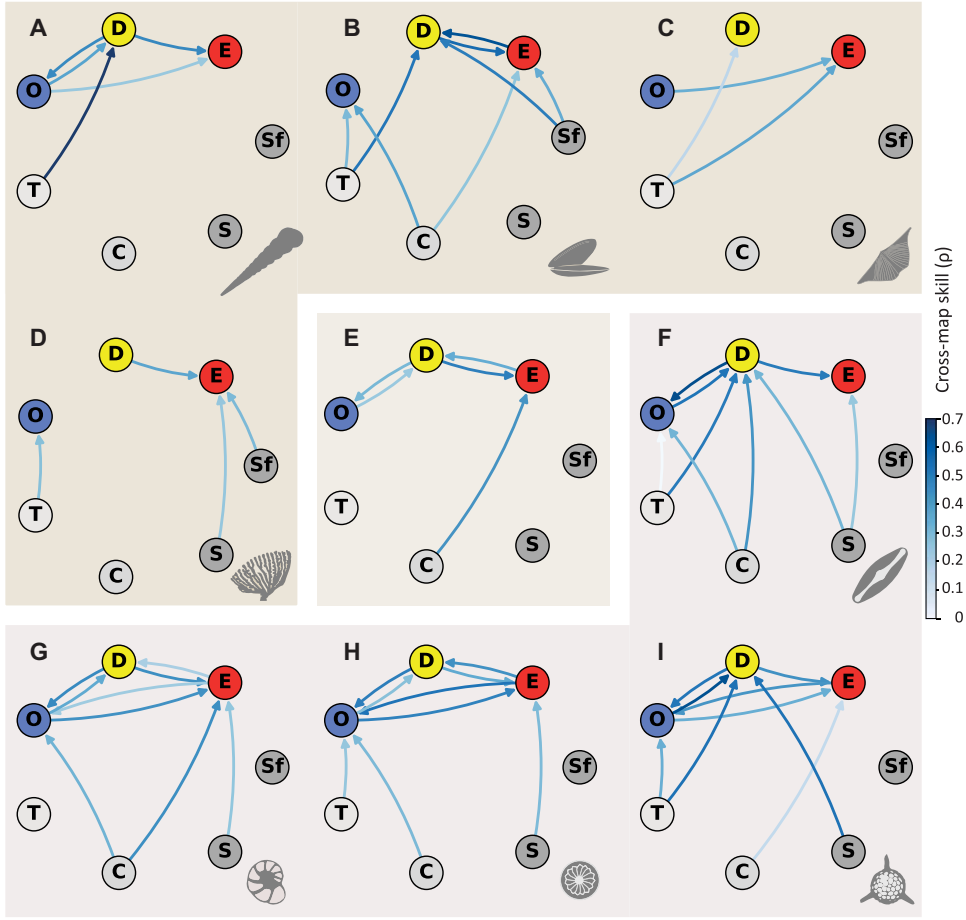

Fig. 1. Networks of causal relationships between abiotic and biotic factors for the nine taxa. Each arrow corresponds to a significant effect detected by the pairwise CCM analysis. The intensity of arrow colors corresponds to the intensity of the causality, expressed as cross-map skill in CCM (correlation between empirical and predicted values). Causal relationships between biotic variables (the rates and diversity) show the strongest effects, with diversity dependence of both rates as the key feature of diversity dynamics in most taxa. (**A**) Gastropoda, (**B**) Bivalvia, (**C**) Brachiopoda, (**D**) Scleractinia, (**E**) Metazoa, (**F**) diatoms, (**G**) Foraminifera, (**H**) Coccolithophoridae, and (**I**) Radiolaria. E, extinction rate; D, diversity; O, origination rate; T, temperature; C, $\delta^{13}C$; S, $^{87}Sr/^{86}Sr$; Sf, $\delta^{34}S$. (A) to (D) represent time series from the PBDB, except metazoans (E), which are in a separate box; (F) to (I) correspond to data taken from the NDB.

Downloaded from https://www.science.org at Universite Pierre Et Marie Curie - Paris 6 (Upmc) on October 31, 2022

Table 1. Synthetic table showing significance and direction of causal relationships affecting extinction and origination rates using CCMand CTE (○). (●) Gray cells highlight results significant in both CCM and CTE. Each cell includes the sign of the relationships between variables using additional information given by autoregressive vectors (CCM and CTE methods being unable to give directionality): +, positive relationship; −, negative relationship. No sign is given when the resulting model is more complex (e.g., nonmonotonic models cannot be characterized solely by a sign). In the case of nonambiguous results, diversity dependence of origination rates is always negative, while diversity dependence of extinction rates is always positive. Gas., Gastropoda; Biv., Bivalvia; Bra., Brachiopoda; Scl., Scleractinia; Met., Metazoa; Dia., diatoms; For., Foraminifera; Coc., Coccolithophorida; Rad., Radiolaria.

Causal relationship	Gas.	Biv.	Bra.	Scl.	Met.	Dia.	For.	Coc.	Rad.
Diversity → extinction	●+	●○	○+	●○	●○+	●○+	●○+	●○	●○
Diversity → origination	●−	○	○−	○	●○−	●○	●○	●○−	●○−
Temperature → extinction			●			○	○	○	○−
Temperature → origination		●+	+	●	○	●○	○	●○	●○
Carbon → extinction	+	●○+	○	○	+	○	●		
Carbon → origination	−	●○	+	○	○+	●	●	●+	○
Strontium → extinction		+		●○−	○	●○	●○	●○	+
Strontium → origination	+	○	○		○+	○	○	○	○
Sulfur → extinction		●○	○−	●	○	○	○	○	○
Sulfur → origination				○	+	○	○	○−	○

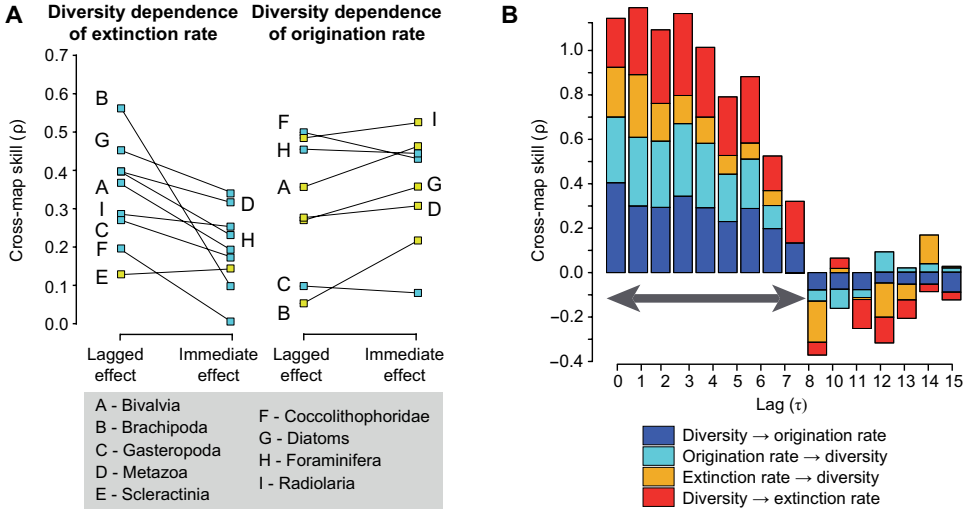

Fig. 2. Temporal signatures of causal relationships affecting the biodiversity dynamics. (**A**) Comparison of causality analyses (CCM) of diversity dependence with or without lag. Blue squares indicate increasing cross-map skill with increasing lag, and yellow squares correspond to decreasing values with increasing lag. In contrast to the origination rate that reveals immediate response to diversity change, the extinction rate responds to changing diversity levels more strongly with a lag. (**B**) Mean of the strength of relationships (CCM) between the biotic variables (diversity and the origination and extinction rates) across diatoms, Coccolithophoridae, Foraminifera, and Radiolaria, tested with various lags (in Ma). The strength of these relationships remains high for the first 7 Ma, highlighting long-lasting effects (gray arrow).

Downloaded from https://www.science.org at Universite Pierre Et Marie Curie - Paris 6 (Upmc) on October 31, 2022

between the variables. In this respect, the CTE analysis that allows for mutual conditioning between multiple causal links may be more informative. We performed CTE analyses of the effect of diversity on origination and extinction rates while conditioning environmental variables and then the reverse, i.e., the analyses comprising the effect of environmental variables on the rates while conditioning the effect of diversity. The results are largely congruent with the results of CCM analyses, as they support the dominant role of biotic factors in diversity dynamics, revealing especially strong effect of diversity on origination rates [Fig. 3 and figs. S16 and S17; see (*12*) and figs. S18 to S24 for sensitivity analyses]. CTE analyses were able to capture the signal of diversity dependence also for taxa showing nonconclusive results in the CCM analyses.

DISCUSSION

Venditti *et al*. (*19*) has suggested that the major evolutionary patterns are determined by rare stochastic events and unpredictable environmentally driven crises and opportunities, rather than by biotic interactions. Our results suggest that, on the contrary, there has been a persistent regulation of biodiversity in Phanerozoic oceans by biotic processes. This finding is robust enough across our datasets, so that it is unlikely to be generated by statistical artifacts. At the same time, it complements previous studies showing that bounded models may represent appropriate characterization of diversity dynamics even when diversity fluctuates and, for most of the history, is out of equilibrium (*4*, *6*). The evidence of diversity dependence of extinction and origination rates in most explored taxa (Table 1) is congruent with the hypothesis of a carrying capacity of the environment for diversity (*20*). However, such a carrying capacity does not necessarily represent a hard ceiling, but a dynamic equilibrium between diversity-dependent extinction and origination rates that can be modulated by environmental factors themselves. Carrying capacity would then be an emergent property of biotic interactions independent

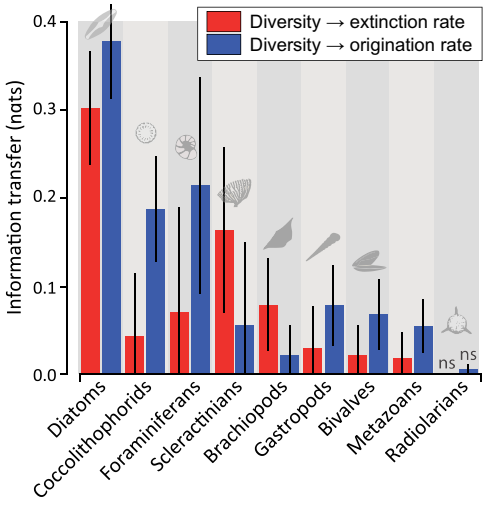

Fig. 3. Differential strength of diversity dependence of origination (blue) and extinction (red) rates while accounting for fluctuations of environmental variables using CTE. Each bar (means ± SD) represents the mean of four CTE analyses with different conditioning variables: $T_{\text{diversity}\rightarrow\text{extinction}|\text{temperature}}$, $T_{\text{diversity}\rightarrow\text{extinction}|\delta 13C}$, $T_{\text{diversity}\rightarrow\text{extinction}|87Sr/86Sr}$, and $T_{\text{diversity}\rightarrow\text{extinction}|\delta 34S}$. Only radiolarians show nonsignificant (ns) values relative to a null model without diversity dependence (*13*). nats, natural unit of information.

of particular histories of different evolutionary lineages, and its level is set by environmental properties affecting the rates (*20*, *21*).

We show that biodiversity dynamics may be diversity dependent but, simultaneously, can be affected by environmental changes. The carrying capacity (diversity equilibrium) may thus change itself in time. This has potentially important implications for analytical tests of diversity regulation. Because the environment has not been constant during the Phanerozoic and the functional form of diversity dependence is not known, the use of simple logistic models (or any models based on a priori–chosen functional forms of causal links) (*1*, *4*) may not be appropriate. The reason is that these models are able to reveal only a specific form of diversity regulation, constrained by the model structure. In contrast, our methods cover a broad class of possible diversity regulation mechanisms, allowing the detection of nonlinearities, simultaneous effects of multiple variables, and time-varying effects. However, this generality is achieved at the expense of an inevitable agnosticism as to the exact mechanisms and functional properties of respective processes. In other words, CCM and CTE analyses do not allow us to identify a particular mechanism generating paleontological diversity time series and its parametric values. Instead, they only show that a subset of these mechanisms that comprises diversity dependence is more likely than a subset without diversity dependence.

Further analyses of paleobiodiversity that would compare various spatial, temporal, or taxonomic scales and take into account a broad list of possible particular mechanisms would enable better understanding of the exact processes responsible for diversity regulation. Process-based approaches also have potential to control for various statistical artifacts that might, even if slightly, distort the presented information flow analyses [such as the regression to the mean sensu (*22*); see Materials and Methods for detailed discussion). To explore biodiversity dynamics in paleontological series in a mechanistic and parametric manner, we need further development of quantitative theories of biodiversity dynamics applicable to large spatial and temporal scales (*20*, *23*), coupled with analytical tools able to deal with noise, various biases, and artifacts that are inherently present in paleontological time series (*3*, *10*, *24*).

Diversity dependence of extinction and origination rates may be mediated by several classes of mechanisms. One possibility is that it is due to changing abundances of individual species (*23*). Increasing diversity at the global scale results, if not complemented with an equal increase of available resources, in a decrease in average species abundance. This is predicted to affect extinction rates because the probability of extinction increases with decreasing population size. Origination rate may be also affected by diversity-dependent species population abundances if species with smaller populations have lower probability to speciate (*20*, *21*). In addition, if origination rates are linked to niche differentiation (*24*), then they may decrease with increasing diversity because of niche filling. These two possibilities—diversity dependence mediated by changing species abundances and by changing niche availability—are not exclusive, as niche availability and niche filling almost certainly affect abundances. Therefore, species abundance may represent a mediator of diversity dependence regardless of the exact role of species niches.

CTE revealed that diversity affects the origination rate more often than the extinction rate (Fig. 3). This has already been demonstrated for smaller taxonomic groups (*25*, *26*), but here, we show that it is a common pattern across multiple phylogenetically distant taxa (Fig. 2A). This may have been influenced by a non-negligible

Downloaded from https://www.science.org at Universite Pierre Et Marie Curie - Paris 6 (Upmc) on October 31, 2022

proportion of undetected taxa because of ephemeral speciation that can blur the distinction between origination and extinction (*27*). If diversity-dependent extinction comprises mostly incipient taxa that were not able to reach sufficiently large abundances to leave imprints in the paleontological record, then it would be reflected in our analyses by the apparent diversity dependence of origination rates. Moreover, in contrast to origination rates, diversity dependence of extinction rates revealed substantial time delay (Fig. 2A and figs. S6 and S7). As discussed above, diversity-related extinction may be driven by shrinking population sizes because of increasing competition for resources. Because extinction of small populations is a probabilistic process, it may take a long time. In contrast, speciation may be relatively fast if it is promoted by an increase of population and range size after a decrease of diversity. Our finding that diversity regulation might be common during the whole Phanerozoic has profound consequences on understanding the current biodiversity crisis. First, if global diversity is bounded and regulated by universal diversity dependence mediated by resource and/or niche availability, then any human-driven effects on the resource base and habitat availability may have profound effects on the maintenance of future biodiversity on Earth. Second, our analyses revealed long-lasting effects of diversity on origination and especially extinction rates. The likely reason for this prolonged effect is that global-scale dynamics results from a complex array of microevolutionary processes that have long-term consequences (*22*). These long-term responses are consistent with the finding that the rebound from an extinction event, such as the one we are experiencing today, would globally extend for the duration of a geological period (*28*). A marked increase in extinction rate in the oceans as we are currently experiencing could thus influence global diversity dynamics for the next several million years.

MATERIALS AND METHODS

Fossil occurrence databases and data cleaning

The taxa were selected on the basis of the following criteria: high abundance and good preservation in the fossil record, sufficient time extent, and ease of identification. Nine taxonomic datasets were retained following these criteria. Four were extracted from the NDB version 30/07/2020 (*18*) and five from the PBDB version 24/08/2020 (https://paleobiodb.org).

From the NDB, our analyses were conducted on Foraminifera [157,960 occurrences; 123 to 0 million years (Ma)], Coccolithophoridae (280,935 occurrences; 150 to 0 Ma), diatoms (114,745 occurrences; 46 to 0 Ma), and Radiolaria (116,529 occurrences; 64 to 0 Ma) (fig. S1). From the PBDB, our analyses were conducted on Scleractinia (32,420 occurrences; 242 to 0 Ma), Brachiopoda (137,695 occurrences; 477.7 to 0 Ma), Bivalvia (144,507 occurrences; 425.6 to 0 Ma), and Gastropoda (108,153 occurrences; 470 to 0 Ma) (fig. S2). The last taxonomic dataset was generated to handle all metazoans (705,129 occurrences; 477.7 to 0 Ma). For Scleractinia, the taxonomic dataset was not extracted directly from the PBDB; instead, we used a corrected PBDB dataset (*29*) extracted using the R package divDyn 0.8.0 (*30*).

All the analyses were conducted at the genus level. All occurrences of taxa without known generic affiliation have been deleted. For PBDB taxonomic datasets, the following criteria have been applied: lump taxa by genus (only one record is reported for each occurrence of a given genus in a given collection); regular taxa only (deletion of ichnotaxa and form taxa); worldwide; exclusion of occurrences with "aff.," "ex. Gr.," "sensu lato," and "informal"; question marks; or quotation marks. We also removed all nonmarine taxa using the command envtype=!terr,lacust,fluvial,terrother) to analyze taxa coming from ecologically homogeneous environments and avoid major fluctuations of potentially inhabitable areas, for example, with the vertebrate land invasion.

For PBDB, the records are dated using the ages of the geological time scale (e.g., Pleistocene from 2.7 to 0 Ma corresponds to one time bin). For NDB for which occurrences have absolute ages, we tested various bin sizes. First, we tested all bin sizes from 0.1 to 1 Ma with 0.1-Ma increment, and then we tested bin sizes from 1 to 10 Ma with 1-Ma increment. We kept equal bin sizes of 1 Ma to avoid empty bins while keeping the highest temporal resolution.

Diversity and turnover rates estimate

Diversity curves were computed using the R package divDyn (*30*) using the shareholder quorum subsampling (SQS) (*31*), which gives the best results when evenness is low (*32*). SQS is an interpolation method that computes how many taxa are expected to be found given a fixed coverage q of the underlying occurrence distribution. SQS deals with the heterogeneity of the fossil record by estimating the diversity of samples with similar levels of sample completeness. We kept the highest q while avoiding empty bins for each taxon (table S1). The curve was constructed using sampled-in-bin values (i.e., sample-in-bin measure of raw taxonomic diversity that does not extrapolate when there are Lazarus taxa present, with occurrences before and after, but not inside a bin). We corrected the sampled-in-bin values by using the three-timer method [corrected sampled-in-bin (cSIB) hereafter] (*28*), which detects sampling variations, and log-transformed the data.

We used the "second for third" method (2f3) to accurately compute origination and extinction rates (*11*). 2f3 is an accurate estimator except when turnover rates are very high or if sampling is very poor, which is not the case here because of the spatial and temporal range of the analyses and because we selected only well-sampled marine taxa. The 2f3 method computes the extinction rate (E2f3) on a bin i_1 as

$$\text{E2f3} = \frac{s_1 - s_3}{t_2 - p}$$

with s_1 as the number of taxa sampled in i_0 and i_1 but not after, s_3 as the number of taxa sampled in i_0 and i_3 but not i_1 and i_2, t_2 as the number of two-timers (taxa sampled in i_0 and i_1), and p the number of part-timers (taxa sampled in i_0 and i_2 but not i_1). The 2f3 method can be corrected if specific counts are too small (*11*, *28*). Last, the rate is computed as log(1/(1 − E2f3)). We also estimated the rates using boundary crossers (*22*), per capita rates (*33*), and gap filler (*34*), which always gave very similar results. We computed all rates using divDyn (*30*).

Stationarity of the time series is a prerequisite of both the CCM and CTE methods to remove spurious relationships (*17*, *35*). We performed Kwiatkowski-Phillips-Schmidt-Shin (KPSS) (*36*) and augmented Dickey–Fuller (*37*) tests to analyze whether the generated time series are stationary or not. The trends were then removed using the LOESS method (fig. S3). The span of the smoothing function was determined using the ADF and KPSS tests (table S1). This procedure was applied for each time series and for each taxonomic dataset because of series-specific time ranges and bin sizes. Last, we performed an ordered quantile transformation on diversity, origination rate, and extinction rate time series to suppress variance changes over time (*38*) and normalized it to zero mean and unit SD before the analyses.

Downloaded from https://www.science.org at Universite Pierre Et Marie Curie - Paris 6 (Upmc) on October 31, 2022

Paleoenvironmental surrogate datasets

Four paleoenvironmental surrogate datasets have been used in our study (fig. S4). The first one was temperature reconstructed through the Phanerozoic using $\delta^{18}O$ and latitudinal belt reconstruction using geological evidence (*39*). The three other paleoenvironmental time series used were $\delta^{13}C$, $^{87}Sr/^{86}Sr$, and $\delta^{34}S$ isotopic ratios, following Cárdenas and Harries (*13*). $\delta^{13}C$ reflects variations in not only primary productivity and burial of organic matter (photosynthetic reduction of CO_2 in organic matter) but also weathering, hydrothermal degassing, clathrates, and volcanism (*40*–*42*). Here, we used the $\delta^{13}C$ time series as an environmental proxy of oceanic productivity. The dataset was taken from the recent compilation made for the Geological Time Scale 2016 (*43*). The $^{87}Sr/^{86}Sr$ time series is a proxy of nutrient input from the weathering of continental rocks. It is taken from McArthur *et al.* (*44*). $\delta^{34}S$ reflects nutrient input from recycling of organic material in ocean sediments. This time series is taken from Prokoph *et al.* (*45*). The four paleoenvironmental datasets were first smoothed using LOESS method (smoothing span parameter = 0.01) and bin-averaged to the same time bins as for the diversity data. The span was determined to get the most precise time series while smoothing fluctuations inside each bin. For all curves, the ages were calibrated to the Geological Time Scale 2016 (*43*) following the procedure of Wei and Peleo-Alampay (*46*). Then, specific detrending (table S1), ordered quantile transformation, and normalization to zero mean and unit SD were performed on all paleoenvironmental time series same as in the case of diversity time series (fig. S4E). We then checked that the four paleoenvironmental time series were not excessively correlated to each other (fig. S5).

Bivariate causality analyses using CCM

The CCM (*14*) is a nonparametric method for inferring causal relationships between variables directly from the data by generating empirical models in the form of dynamic systems. It can perform analyses with uneven data, such as time series computed from PBDB, but requires stationarity (*17*). The CCM allows one to distinguish causality from spurious correlation, based on the inference of predictability. Specifically, CCM uses the concept of causality in the sense of the study by Sugihara *et al.* (*14*), which says that, in the framework of dynamical systems, if a variable x (the source variable) has an influence on another variable y (the target variable), then y can be used to reconstruct the states of x, i.e., the information about the source must be recorded in the target (*47*). The cross-map skill ρ is the measure of the strength of causality between pairs of variables. It is calculated by a Pearson correlation between the observed and predicted values. If the causality is unidirectional, e.g., $x \rightarrow y$, then the dynamical system modeled from y is efficient in predicting the values of x but not the reverse. If a causal relationship links two given variables, then ρ should increase and converge by increasing the size of the subsamples of the time series used for the reconstruction. CCM analyses have been repeatedly used for paleontological time series (*48*–*50*).

For each taxonomic dataset, 18 bivariate CCM analyses were performed (Fig. 1 and Table 1). All biotic-to-biotic relationships between diversity, extinction rate, and origination rate were tested. For environmental-to-biotic causality relationships, we tested the effect of each of the four environmental variables (temperature, $\delta^{13}C$, $^{87}Sr/^{86}Sr$, and $\delta^{34}S$) on each of the three biotic variables (diversity, extinction rate, and origination rate). To perform simplex projection and CCM analyses, we used the R package rEDM 0.7.5. This package implements time delays with CCM that identifies a causal relationship from x to y by testing for different time lags. Time delay analyses can be performed with different lags to identify the best informative time lags in terms of cross-map skill (*16*). For PBDB taxonomic datasets, bivariate analyses were performed with all lags between 0 and 5 (Fig. 2A and fig. S6). For taxonomic datasets coming from the NDB, all bivariate analyses were performed with all lags between 0 and 15 (Fig. 2, A and B, and fig. S7). To assess significance of the results, we compared the results with 1000 randomly generated time series using the Ebisuzaki method (*51*). The Ebisuzaki method allows generating time series by randomizing the phases of a Fourier transform, preserving the power spectra of the original time series. Alternatively, we also assessed the significance of the results using 1000 randomly reshuffled surrogate time series, but this method gave less conservative results than the Ebisuzaki method. Kendall correlation tests were also performed to assess convergence of the results with increasing sample size (*52*). Results were considered significant when $P < 0.05$.

To explore the signs of the relationships between significantly linked variables, we performed additional analyses. The first was a lagged cross-correlation (figs. S8 to S15). The second set of analyses aimed to fit a vector autoregression model for pairs of variables using an Akaike information criterion (AIC) to retain the adequately lagged variables (Table 1). Here, the sign of the causation is given by the coefficient of each lagged explanatory variable. However, several coefficients for various lags concerning the same explanatory variable may occur, which thus precludes straightforward interpretation. For both cross-correlations and vector autoregressions analyses, the bivariate analyses and the number of minimum and maximum lags were the same as for CCM.

CTE for testing drivers of the rates of extinction and origination

In addition to bivariate CCM analyses, we used the CTE to control for spurious causal relationships caused by other variables (*17*). Similar to CCM, CTE infers causality from predictability using stationary time series (*15*, *17*). CTE is a measure of information transfer between a source variable (the causal variable) and a target variable while conditioning a third variable. Unlike CCM, CTE allows for modeling simultaneous causal effects of multiple variables, can be used for causality partitioning, and can be used specifically for extracting the net effect of one causal relationship in a system where multiple causal links take place. Within the information theory framework (*53*, *54*), the transfer entropy approach (*55*, *56*) is a method to compute a nonsymmetric measure of causality, the information transfer, from one variable to another.

When computing the transfer entropy between a source variable x and a target variable y, an additional variable z can lead to spurious causation signals if only simple bivariate analyses are carried out between x and y. Common driver effects lead to spurious information flow from x to y caused by the information flow of a third variable z on x and y (i.e., $z \rightarrow x$ and $z \rightarrow y$ instead of $x \rightarrow y$). Similarly, cascade effects appear where we actually have $x \rightarrow z \rightarrow y$ instead of $x \rightarrow y$. These cases can be distinguished by conditioning on z, which allows the direct information flow from y to x to be computed. CTE has been applied to paleontological and geological time series analyses to test for causality and/or to detect common driver effects (*3*, *48*, *57*, *58*).

We used the CTE to test the effect of various potential causal variables on rates of origination and extinction while accounting for

Downloaded from https://www.science.org at Universite Pierre Et Marie Curie - Paris 6 (Upmc) on October 31, 2022

the potential interplay with other variables (Fig. 3 and Table 1). We performed a CTE analysis for each taxonomic dataset and for each environmental time series. Thus, for each taxonomic dataset, we used four sets of time series: diversity, extinction rates, origination rates, and temperature; diversity, extinction rates, origination rates, and carbon; diversity, extinction rates, origination rates, and strontium; and diversity, extinction rates, origination rates, and sulfur. For each set of each taxonomic dataset, we analyzed, using CTE, the following relationships: diversity → extinction rate, diversity → origination rate, temperature/carbon/strontium/sulfur → origination rate, and temperature/carbon/strontium/sulfur → extinction rate, while accounting for the effect of other time series from the set taken as conditional variables.

CTE analyses were performed using the Python 3 toolkit IDTxl (*59*). This library implements the CTE with a selection of variables and source lags using an algorithm to infer the minimum set of selected source and target variables with specific lag. Potential source and target variables with specific lags are iteratively added to the set of selected sources and target to optimize a criterion of uncertainty reduction, quantified as conditional mutual information (*53*), and stop when uncertainty cannot be lower. In our analyses, any variable present in the set of time series (a to d; see above) could be used under these criteria. Moreover, several statistical tests (*15*) are implemented in IDTxl to assess statistical significance, including a false discovery rate correction (*60*) to control for false positives. As IDTxl allows enforcing the inclusion of specific variables in the conditioning set of the CTE analysis, we forced the inclusion of a variable's informative (i.e., with positive cross-map skill) lags detected using lagged CCM analyses (figs. S6 and S7 and data S1) while allowing the algorithm to detect and include other informative lags in the analysis. For the analysis of diversity dependence, we enforced the use of environmental time series informative lags, and we enforced the use of diversity informative lags when analyzing the interaction between environmental variables and the origination and extinction rates. The candidate target and source sets were set with a lag of between 0 and 5 time bins for every time series of the PBDB analyses and between 0 and 15 time bins for the NDB analyses. We explored several sets of parameters of enforced variables (without enforcing anything; with all variables enforced; and with all source and target variables with a lag of 0, then with lag of 1, then with a lag of 0 and 1) to ensure the consistency of the results. The information transfer values (expressing the strength of the relationships) were always similar, and only the number of significant results varied. Analyses were performed by using both the Gaussian estimator, equivalent to Granger causality when used for transfer entropy estimation (*61*), and the Kraskov-Stögbauer-Grassberger (KGS) estimator for linear and nonlinear data (*62*). We retained only the results from the analyses that used the Gaussian estimator (fig. S16), as the time series were too short to obtain reliable results with the KGS estimator, which generally requests more data. The use of the Gaussian estimator is made under the assumption that the set of time series is jointly Gaussian, i.e., their joint distribution is a multivariate normal distribution. We checked this assumption for each set of time series corresponding to a taxonomic dataset using Henze-Zirkler's tests. Last, for each analysis, a joint transfer entropy of all selected sources to the target was calculated by IDTxl (fig. S17) using an omnibus test (*15*). The joint transfer entropy test takes all the variables together to quantify a global amount of information transfer.

Parametric null model

We simulated fossil occurrence datasets using a null birth-death model to meet two objectives: (i) to show the significance of our results compared to a simplistic but biologically interpretable model without diversity dependence and (ii) to show that regression to the mean (*10*, *28*, *63*) is not likely to distort the results obtained by the analyses presented here (see the "Effect of the regression to the mean" section for details). We ran the null model simulations 100 times across discrete time axis with 100 time bins, roughly corresponding to our empirical series that range between 42 (Scleractinia) and 153 (Coccolitophoridae) time bins. In every time bin τ, the number of genera going extinct in time bin $\tau + 1$ was drawn from a binomial distribution with the number of trials equal to the total richness in time bin τ and the probability of success equal to per-genus extinction rate. The number of genera originating in $\tau + 1$ was generated in similar way, with the probability of success equal to per-genus origination rate. The per-genus origination and extinction rates were kept constant across the time bins. We used a value of 0.1 for both origination and extinction, roughly corresponding to the origination (0.05 to 0.26 per time bin) and extinction (0.03 to 0.20 per time bin) rates in our empirical datasets. We also performed simulations with stochastic origination an extinction rates using a β distribution ($\alpha = 10$, $\beta = 90$), which gave similar results. The initial number of species in the simulations was 5000. The simulated series were then subsampled to 0.1 and 0.5 proportion of genera in every time bin. The null model simulations were processed with the same analytical pipelines as the real-world data, resulting in the cross-map skills and CTEs based on CCM and CTE analyses, respectively. All simulations were performed in R, and the code used for running simulations and filtering (data S1) is available in the Data and materials availability statement.

Overall, the results of our analyses of diversity dependence are almost all significant relatively to those obtained with the null model described here (fig. S18). This constitutes additional evidence that our results significantly differ not only from nonparametric null approaches such as Ebisuzaki time series or reshuffling presented above but also from a parametric null model, albeit simplistic one.

Effect of the regression to the mean

We assessed whether the regression to the mean (*10*, *28*, *63*) might cause important biases in our results. We proceed in two ways, concluding that the reported results of CCM and CTE analyses are very likely not driven by this artifact.

First, we explored the temporal patterns of link strengths in the CCM analyses. The regression to the mean is known to generate artificial linkage between diversity and origination or extinction rates, which has the strongest effect when comparing neighboring time bins (i.e., the immediate effect), and the strength of the effect is predicted to uniformly decrease with the time lag (*22*). This pattern is exactly opposite to the comparison of immediate and lagged effects of diversity on extinction rates presented in Fig. 2A. Moreover, Fig. 2B and figs. S6 and S7 suggest that the effect of diversity on both the rates does not uniformly decrease with time lag but instead is generally comparable for time lags between 0 and 6 and only then sharply drops. This evidence suggests that, although the regression to the mean may play a certain role in the presented CCM results, the core part of the results, including the strong effect of diversity on the rates with large time lags, is extremely unlikely to result from this artifact.

Downloaded from https://www.science.org at Universite Pierre Et Marie Curie - Paris 6 (Upmc) on October 31, 2022

In addition to this evidence, we also explored the effect of regression to the mean using the null birth-death model described above. The effect of regression to the mean should increase with stronger censoring of the data from the generating diversification process (*64*). We thus simulated the fossil occurrences with different sampling probabilities $P = 1$, $P = 0.5$, and $P = 0.1$ and analyzed them with the same pipeline as the real-world data. If our analytical workflow was strongly biased by the regression to the mean, then the obtained cross-map skills and CTEs should be highest for the most censored simulations ($P = 0.1$). However, the simulation results show the opposite pattern (fig. S19). This suggests that our analytical workflow of LOESS detrending of the time series followed by CCM and CTE analyses is not vulnerable to the regression to the mean artifact, at least under the birth-death model used for these simulations.

Sensitivity analyses

We ran a set of analyses to test the sensitivity of our results to the choices of data cleaning and taxonomic rank. To evaluate the sensitivity of our results to fossilization biases, we produced sensitivity analyses for PBDB datasets by excluding occurrences with specific taphonomic parameters, such as occurrences with soft tissue or original aragonite, occurrences of fossils replaced with silica, and occurrences coming from unlithified or poorly lithified and sieved deposits. The results were similar to those of the main analyses, confirming the robustness of our results.

The analyses shown in the Materials and Methods section used data detrended using a LOESS smoothing function, removing long-term trends and making time series stationary. To determine the impact of detrending strategies, we performed other CCM analyses on the nondetrended time series (fig. S20) and on the time series that were detrended by differencing (fig. S21) that removes not only the long-term trends but also the autocorrelation. The differencing in first order or second order was determined visually for each time series by looking at the autocorrelation plots. The results are generally more significant, with higher cross-map skills, when the time series are not detrended, suggesting that detrending is a conservative choice. We also performed CTE analyses with alternative detrending strategies (fig. S22). They show that the diversity dependence of the origination rates is stronger than that of the extinction rates regardless of the type or absence of detrending. The diversity dependence is present regardless of the method (CCM and CTE) and regardless of the type of detrending.

To check the robustness of our choice to work with data aggregated by genera, we performed CTE analyses on the effect of diversity on extinction and origination rates (while conditioning on environmental proxies) using different taxonomic ranks. The results using families, genera, and species are shown for each environmental proxy taken as a conditioning variable (figs. S23 and S24). For the PBDB analyses at the family level, we used the familial attributions in the database. For the NDB analyses, we used the familial determinations of the PBDB. Quorum values used for species and families are given in table S1. Comparison of the results shows a high variance of the results using species, which is not present at the genus or at the family level, probably because of the high turnover at this taxonomic level that affects diversity, origination, and extinction rate calculations.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at https://science.org/doi/10.1126/sciadv.add9620

REFERENCES AND NOTES

1. J. J. Sepkoski, A kinetic model of Phanerozoic taxonomic diversity I. Analysis of marine orders. *Paleobiology* **4**, 223–251 (1978).
2. R. A. Close, R. B. Benson, E. E. Saupe, M. E. Clapham, R. J. Butler, The spatial structure of Phanerozoic marine animal diversity. *Science* **368**, 420–424 (2020).
3. B. Hannisdal, S. E. Peters, Phanerozoic earth system evolution and marine biodiversity. *Science* **334**, 1121–1124 (2011).
4. P. Cermeño, C. García-Comas, A. Pohl, S. Williams, M. Benton, G. L. Gland, R. D. Muller, A. Ridgwell, S. Vallina, Post-extinction recovery of the Phanerozoic oceans and biodiversity hotspots. *Nature* **607**, 507–511 (2022).
5. A. Zaffos, S. Finnegan, S. E. Peters, Plate tectonic regulation of global marine animal diversity. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 5653–5658 (2017).
6. J. Alroy, The shifting balance of diversity among major marine animal groups. *Science* **329**, 1191–1194 (2010).
7. M. Benton, The Red Queen and the Court Jester: Species diversity and the role of biotic and abiotic factors through time. *Science* **323**, 728–732 (2009).
8. B. Hannisdal, L. H. Liow, Causality from palaeontological time series. *Palaeontology* **61**, 495–509 (2018).
9. D. L. Rabosky, Diversity-dependence, ecological speciation, and the role of competition in macroevolution. *Annu. Rev. Ecol. Evol. Syst.* **44**, 481–502 (2013).
10. J. Alroy, Geographical, environmental and intrinsic biotic controls on Phanerozoic marine diversification: Controls on phanerozoic marine diversification. *Palaeontology* **53**, 1211–1235 (2010).
11. J. Alroy, A more precise speciation and extinction rate estimator. *Paleobiology* **41**, 633–639 (2015).
12. Supplementary Materials can be found on *Science Advances* Online.
13. A. L. Cárdenas, P. J. Harries, Effect of nutrient availability on marine origination rates throughout the Phanerozoic eon. *Nat. Geosci.* **3**, 430–434 (2010).
14. G. Sugihara, R. May, H. Ye, C.-H. Hsieh, E. Deyle, M. Fogarty, S. Munch, Detecting causality in complex ecosystems. *Science* **338**, 496–500 (2012).
15. L. Novelli, P. Wollstadt, P. Mediano, M. Wibral, J. T. Lizier, Large-scale directed network inference with multivariate transfer entropy and hierarchical statistical testing. *Netw. Neurosci.* **3**, 827–847 (2019).
16. H. Ye, E. R. Deyle, L. J. Gilarranz, G. Sugihara, Distinguishing time-delayed causal interactions using convergent cross mapping. *Sci. Rep.* **5**, 14750 (2015).
17. T. Bossomaier, L. Barnett, M. Harré, J. T. Lizier, *An Introduction to Transfer Entropy* (Springer International Publishing, 2016).
18. J. Renaudie, D. Lazarus, P. Diver, NSB (Neptune Sandbox Berlin): An expanded and improved database of marine planktonic microfossil data and deep-sea stratigraphy. *Palaeontol. Electron.* **23**, a11 (2020).
19. C. Venditti, A. Meade, M. Pagel, Phylogenies reveal new interpretation of speciation and the Red Queen. *Nature* **463**, 349–352 (2010).
20. D. Storch, J. G. Okie, The carrying capacity for species richness. *Glob. Ecol. Biogeogr.* **28**, 1519–1532 (2019).
21. M. L. Rosenzweig, *Species Diversity in Space and Time* (Cambridge Univ. Press, 1995).
22. M. Foote, R. A. Cooper, J. S. Crampton, P. M. Sadler, Diversity-dependent evolutionary rates in early Palaeozoic zooplankton. *Proc. R. Soc. B* **285**, 20180122 (2018).
23. D. Storch, E. Bohdalková, J. Okie, The more-individuals hypothesis revisited: The role of community abundance in species richness regulation and the productivity-diversity relationship. *Ecol. Lett.* **21**, 920–937 (2018).
24. J. T. Flannery-Sutherland, D. Silvestro, M. J. Benton, Global diversity dynamics in the fossil record are regionally heterogeneous. *Nat. Commun.* **13**, 2751 (2022).
25. D. Moen, H. Morlon, Why does diversification slow down? *Trends Ecol. Evol.* **29**, 190–197 (2014).
26. J. Alroy, Constant extinction, constrained diversification, and uncoordinated stasis in North American mammals. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **127**, 285–311 (1996).
27. E. B. Rosenblum, B. A. J. Sarver, J. W. Brown, S. Des Roches, K. M. Hardwick, T. D. Hether, J. M. Eastman, M. W. Pennell, L. J. Harmon, Goldilocks meets Santa Rosalia: An ephemeral speciation model explains patterns of diversification across time scales. *Evol. Biol.* **39**, 255–261 (2012).
28. J. Alroy, Dynamics of origination and extinction in the marine fossil record. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 11536–11542 (2008).
29. W. Kiessling, Á. T. Kocsis, Biodiversity dynamics and environmental occupancy of fossil azooxanthellate and zooxanthellate scleractinian corals. *Paleobiology* **41**, 402–414 (2015).
30. Á. T. Kocsis, C. J. Reddin, J. Alroy, W. Kiessling, The *R* package divdyn for quantifying diversity dynamics using fossil sampling data. *Methods Ecol. Evol.* **10**, 735–743 (2019).
31. J. Alroy, Fair sampling of taxonomic richness and unbiased estimation of origination and extinction rates. *Paleontol. Soc. Pap.* **16**, 55–80 (2010).

Downloaded from https://www.science.org at Universite Pierre Et Marie Curie - Paris 6 (Upmc) on October 31, 2022

32. R. A. Close, S. W. Evers, J. Alroy, R. J. Butler, How should we estimate diversity in the fossil record? Testing richness estimators using sampling-standardised discovery curves. *Methods Ecol. Evol.* **9**, 1386–1400 (2018).
33. M. Foote, Morphological diversity in the evolutionary radiation of Paleozoic and post-Paleozoic crinoids. *Paleobiology* **25**, 1–115 (1999).
34. J. Alroy, Accurate and precise estimates of origination and extinction rates. *Paleobiology* **40**, 374–397 (2014).
35. A. A. Tsonis, E. R. Deyle, H. Ye, G. Sugihara, Convergent cross mapping: Theory and an example, in *Advances in Nonlinear Geosciences*, A. A. Tsonis, Ed. (Springer, 2018), pp. 587–600.
36. D. Kwiatkowski, P. C. Phillips, P. Schmidt, Y. Shin, Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *J. econom.* **54**, 159–178 (1992).
37. D. A. Dickey, W. A. Fuller, Distribution of the estimators for autoregressive time series with a unit root. *J. Am. Stat. Assoc.* **74**, 427–431 (1979).
38. R. A. Peterson, J. E. Cavanaugh, Ordered quantile normalization: A semiparametric transformation built for the cross-validation era. *J. Appl. Stat.* **47**, 2312–2327 (2020).
39. C. R. Scotese, H. Song, B. J. W. Mills, D. G. van der Meer, Phanerozoic paleotemperatures: The earth's changing climate during the last 540 million years. *Earth Sci. Rev.* **215**, 103503 (2021).
40. T. E. Cerling, J. M. Harris, B. J. MacFadden, M. G. Leakey, J. Quade, V. Eisenmann, J. R. Ehleringer, Global vegetation change through the Miocene/Pliocene boundary. *Nature* **389**, 153–158 (1997).
41. R. F. Sage, D. A. Wedin, M. Li, The biogeography of C4 photosynthesis: Patterns and controlling factors, in *C4 Plant Biology*, R. F. Sage, R. K. Monson Eds. (Elsevier, 1999), pp. 313–373.
42. J. R. Reinfelder, A. M. L. Kraepiel, F. M. M. Morel, Unicellular C4 photosynthesis in a marine diatom. *Nature* **407**, 996–999 (2000).
43. J. G. Ogg, G. Ogg, F. M. Gradstein, *A Concise Geologic Time Scale: 2016* (Elsevier, ed. 1, 2016).
44. J. M. McArthur, R. J. Howarth, G. A. Shields, Strontium isotope stratigraphy, in *The Geologic Time Scale 2012*, F. M. Gradstein, J. G. Ogg, M. D. Schmitz, G. M. Ogg, Eds. (Elsevier, ed. 1, 2012), vol. 1, chap. 7.
45. A. Prokoph, G. A. Shields, J. Veizer, Compilation and time-series analysis of a marine carbonate $\delta^{18}O$, $\delta^{13}C$, $^{87}Sr/^{86}Sr$ and $\delta^{34}S$ database through Earth history. *Earth Sci. Rev.* **87**, 113–133 (2008).
46. W. Wei, A. Peleo-Alampay, Updated Cenozoic nannofossil magnetobiochronology. *Int. Nannofossil Assoc. Newsl.* **15**, 15–17 (1993).
47. F. Takens, Detecting strange attractors in turbulence, in *Dynamical Systems and Turbulence*, D. Rand, L.-S. Young, Eds. (Springer, 1981), vol. 898 of *Lecture Notes in Mathematics*, pp. 366–381.
48. B. Hannisdal, K. A. Haaga, T. Reitan, D. Diego, L. H. Liow, Common species link global ecosystems to climate change: Dynamical evidence in the planktonic fossil record. *Proc. R. Soc. B* **284**, 20170722 (2017).
49. P. Cermeño, M. J. Benton, Ó. Paz, C. Vérard, Trophic and tectonic limits to the global increase of marine invertebrate diversity. *Sci. Rep.* **7**, 15969 (2017).
50. K. L. Cramer, A. O'Dea, C. Carpenter, R. D. Norris, A 3000 year record of Caribbean reef urchin communities reveals causes and consequences of long-term decline in *Diadema antillarum*. *Ecography* **41**, 164–173 (2018).
51. W. Ebisuzaki, A method to estimate the statistical significance of a correlation when the data are serially correlated. *J. Climate* **10**, 2147–2153 (1997).
52. C.-W. Chang, M. Ushio, C.-H. Hsieh, Empirical dynamic modeling for beginners. *Ecol Res.* **32**, 785–796 (2017).
53. T. M. Cover, J. A. Thomas, *Elements of Information Theory* (Wiley, ed. 2, 2006).
54. C. E. Shannon, A mathematical theory of communication. *Bell Sys. Tech. J.* **27**, 379–423 (1948).
55. M. Paluš, V. Albrecht, I. Dvořák, Information theoretic test for nonlinearity in time series. *Phys. Lett.* **175**, 203–209 (1993).
56. T. Schreiber, Measuring information transfer. *Phys. Rev. Lett.* **85**, 461–464 (2000).
57. B. Hannisdal, Non-parametric inference of causal interactions from geological records. *Am. J. Sci.* **311**, 315–334 (2011).
58. A. M. Dunhill, B. Hannisdal, N. Brocklehurst, M. J. Benton, On formation-based sampling proxies and why they should not be used to correct the fossil record. *Palaeontology* **61**, 119–132 (2018).
59. P. Wollstadt, J. Lizier, R. Vicente, C. Finn, M. Martinez-Zarzuela, P. Mediano, L. Novelli, M. Wibral, IDTxl: The information dynamics toolkit xl: A python package for the efficient analysis of multivariate information dynamics in networks. *J. Open Source Softw.* **4**, 1081 (2019).
60. Y. Benjamini, Y. Hochberg, Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. B Stat. Methodol.* **57**, 289–300 (1995).
61. A. García-Medina, J. B. Hernandez C, Network analysis of multivariate transfer entropy of cryptocurrencies in times of turbulence. *Entropy* **22**, 760 (2020).
62. A. Kraskov, H. Stögbauer, P. Grassberger, Estimating mutual information. *Phys. Rev. E* **69**, 066138 (2004).
63. A. M. Bush, R. K. Bambach, Paleoecologic megatrends in marine Metazoa. *Annu. Rev. Earth Planet. Sci.* **39**, 241–269 (2011).
64. A. G. Barnett, J. C. Van Der Pols, A. J. Dobson, Regression to the mean: What it is and how to deal with it. *Int. J. Epidemiol.* **34**, 215–220 (2005).

Acknowledgments: We thank J. McArthur, J. Ogg, G. Ogg, and M. Saltzman for sharing datasets; H. Ye (rEDM) and P. Wollstadt (IDTxl) for advice about the software; and L. Villier for insightful discussions. **Funding:** This work was supported by the Czech Science Foundation grant 20-29554X (DS). **Author contributions:** Writing—conceptualization: V.R. Methodology: V.R. Investigation: V.R., J.S., and D.S. Visualization: V.R., J.S., and D.S. Funding acquisition: D.S. Project administration: D.S. Supervision: D.S. Writing—original draft: V.R., J.S., and D.S. Writing—review and editing: V.R., J.S., and D.S. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** Script files and all datasets (data S1) used here are available at https://doi.org/10.5281/zenodo.6241515. All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials.

Submitted 15 July 2022
Accepted 12 September 2022
Published 28 October 2022
10.1126/sciadv.add9620

Downloaded from https://www.science.org at Universite Pierre Et Marie Curie - Paris 6 (Upmc) on October 31, 2022

ScienceAdvances

Diversity dependence is a ubiquitous phenomenon across Phanerozoic oceans

Valentin RineauJan Smy#kaDavid Storch

Sci. Adv., 8 (43), eadd9620. • DOI: 10.1126/sciadv.add9620

View the article online
https://www.science.org/doi/10.1126/sciadv.add9620
Permissions
https://www.science.org/help/reprints-and-permissions

Downloaded from https://www.science.org at Universite Pierre Et Marie Curie - Paris 6 (Upmc) on October 31, 2022

Use of this article is subject to the Terms of service

Science Advances (ISSN) is published by the American Association for the Advancement of Science. 1200 New York Avenue NW, Washington, DC 20005. The title *Science Advances* is a registered trademark of AAAS.
Copyright © 2022 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. Distributed under a Creative Commons Attribution NonCommercial License 4.0 (CC BY-NC).