Received: 30 January 2023 | Revised: 27 June 2023 | Accepted: 4 July 2023

DOI: 10.1111/geb.13735

DATA ARTICLE

Global Ecology and Biogeography A Journal of Macroecology

WILEY

BioDeepTime: A database of biodiversity time series for modern and fossil assemblages

Jansen Smith[1,2] | **Marina C. Rillo**[3] | **Ádám T. Kocsis**[1] | **Maria Dornelas**[4,5] | **David Fastovich**[6,7] | **Huai-Hsuan M. Huang**[8] | **Lukas Jonkers**[9] | **Wolfgang Kiessling**[1] | **Qijian Li**[10] | **Lee Hsiang Liow**[11] | **Miranda Margulis-Ohnuma**[12] | **Stephen Meyers**[13] | **Lin Na**[10] | **Amelia M. Penny**[4] | **Kate Pippenger**[12] | **Johan Renaudie**[14] | **Erin E. Saupe**[15] | **Manuel J. Steinbauer**[16] | **Mauro Sugawara**[17,18] | **Adam Tomašovỳch**[19] | **John W. Williams**[6] | **Moriaki Yasuhara**[20,21] | **Seth Finnegan**[22] | **Pincelli M. Hull**[12,23]

Correspondence
Jansen Smith, GeoZentrum Nordbayern, Friedrich-Alexander Universität Erlangen-Nürnberg (FAU), Loewenichstr. 28, 91054 Erlangen, Germany.
Email: jansen.smith@fau.de

Funding information
Chinese Academy of Sciences, Grant/Award Number: CAS-WX2021SF-0205; Deutsche Forschungsgemeinschaft, Grant/Award Number: 390741603; European Union, Grant/Award Number: 101044975; Leverhulme Centre for Anthropocene Biodiversity, Grant/Award Number: RC-2018-021; Leverhulme Trust, Grant/Award Number: RPG-201170 and RPG-2019-402; National Science Foundation, Grant/Award Number: DGE-2139841; National Science Research Council, Grant/Award Number: NE/V011405/1; Volkswagen Foundation, Grant/Award Number: Az 96 796; Youth Innovation Promotion Association, Grant/Award Number: 2019310; Leverhulme Prize; European Research Council, Grant/Award Number: 724324; Peter Buck Postdoc Fellowship; Smithsonian Institution; Slovak Research and Development Agency, Grant/Award Number: APVV22-0523; Slovak Scientific Grant Agency, Grant/Award Number: VEGA02/0106/23

Handling Editor: Jacquelyn Gill

Abstract

Motivation: We have little understanding of how communities respond to varying magnitudes and rates of environmental perturbations across temporal scales. BioDeepTime harmonizes assemblage time series of presence and abundance data to help facilitate investigations of community dynamics across timescales and the response of communities to natural and anthropogenic stressors. BioDeepTime includes time series of terrestrial and aquatic assemblages of varying spatial and temporal grain and extent from the present-day to millions of years ago.

Main Types of Variables Included: BioDeepTime currently contains 7,437,847 taxon records from 10,062 assemblage time series, each with a minimum of 10 time steps. Age constraints, sampling method, environment and taxonomic scope are provided for each time series.

Spatial Location and Grain: The database includes 8752 unique sampling locations from freshwater, marine and terrestrial ecosystems. Spatial grain represented by individual samples varies from quadrats on the order of several cm^2 to grid cells of ~100 km^2.

Time Period and Grain: BioDeepTime in aggregate currently spans the last 451 million years, with the 10,062 modern and fossil assemblage time series ranging in extent from years to millions of years. The median extent of modern time series is 18.7 years and for fossil series is 54,872 years. Temporal grain, the time encompassed by individual samples, ranges from days to tens of thousands of years.

Major Taxa and Level of Measurement: The database contains information on 28,777 unique taxa with 4,769,789 records at the species level and another 271,218 records known to the genus level, including time series of benthic and planktonic foraminifera, coccolithophores,

---

Jansen Smith, Marina C. Rillo and Ádám T. Kocsis contributed equally.

diatoms, ostracods, plants (pollen), radiolarians and other invertebrates and vertebrates. There are to date 7012 modern and 3050 fossil time series in BioDeepTime.

Software Format: SQLite, Comma-separated values.

KEYWORDS

deep-time, global, palaeobiology, species richness, temporal, timescale, turnover

1 | BACKGROUND

One of the greatest challenges of our time is disentangling the multiple dimensions of biodiversity change and their drivers (McGill et al., 2015) across a broad range of timescales and charting a best-course forward in a biosphere strongly impacted by human activity (IPBES, 2019). To begin addressing this challenge, the neontological BioTIME database was published in 2018 after nearly a decade of data compilation efforts (Dornelas et al., 2018). BioTIME includes more than 12 million records of species presence and/or abundance from time series of at least 2 years in extent (duration) and has transformed our understanding of the impacts of humanity on biodiversity (e.g. Blowes et al., 2019; Dornelas et al., 2014; Gotelli et al., 2017). Although the temporal scope of BioTIME is exceptional, its records typically span only the last several decades, and only occasionally reach centennial extents. However, humans have been altering ecosystems for centuries and millennia (Ellis et al., 2013; Jackson et al., 2001; Koch & Barnosky, 2006; Lotze et al., 2006; Yasuhara et al., 2012), and ecological change has occurred for as long as life itself. Without data spanning longer timescales, it is difficult to disentangle anthropogenic drivers of biodiversity change from processes inherent to communities in the absence of human influence (Barnosky et al., 2017; Dietl et al., 2019; Dillon et al., 2022; Kiessling et al., 2023; Lewandowska et al., 2020; Smith et al., 2020; Yasuhara et al., 2020).

A long temporal perspective is necessary to understand the full extent of humanity's impact on the biosphere (Barnosky et al., 2017; Blowes et al., 2019; Dietl et al., 2015; Dornelas et al., 2014; Kiessling et al., 2023). Widespread changes in the range and abundance of species in marine and terrestrial environments are well documented (e.g. Antão et al., 2020; Dirzo et al., 2014; McCauley et al., 2015; Pauly et al., 1998), and projected rates of future extinction are high (IPBES, 2019; Pimm et al., 2014). Yet, the number of confirmed recent extinctions is relatively low (IPBES, 2019) and the observed species richness of local assemblages is surprisingly constant through time (Dornelas et al., 2014; Vellend et al., 2013, 2017, but see Gonzalez et al., 2016; Primack et al., 2018), despite high rates of community turnover (Blowes et al., 2019; Dornelas et al., 2014). The fossil record provides temporal context for contemporary biodiversity trends (Barnosky et al., 2017; Dietl et al., 2015; Dillon et al., 2022; Wolkovich et al., 2014; Yasuhara & Deutsch, 2022). Moreover, modern assemblage time series spanning only several decades are insufficient to capture the full extent of climatic impacts on biodiversity (Antão et al., 2020; Chavez et al., 2003; Kuwae et al., 2017; Valdés et al., 2008), which can operate on timescales from a single year to millions of years (MacDonald et al., 2008; Mottl et al., 2021; Westerhold et al., 2020; Yasuhara et al., 2014; Yasuhara & Deutsch, 2023).

Longer time series are needed to answer pressing questions including whether the rate, magnitude and direction of modern change are anomalously high relative to natural variability in the fossil record (Lewandowska et al., 2020; Williams et al., 2021). Such time series are also needed to address issues of scaling effects. Comparisons of biodiversity patterns between relatively short historical timescales and longer geological timescales are complicated by the fact that rates appear inherently slower when viewed through a coarser temporal lens (i.e. larger temporal grain) or when viewed across larger time steps, a phenomenon that is known to affect the perception of Earth system change (Sadler, 1981), climate change (Kemp et al., 2015), species diversification (Henao Diaz et al., 2019; Louca et al., 2022) and extinction (Foote, 1994; Spalding & Hull, 2021). Although the greatest strength of the fossil record is arguably the fact that it documents changes in species, communities and ecosystems through time—which can be used as temporal context for modern biodiversity change—there has yet to be a community-driven effort to compile fossil assemblage records and combine them with neontological time series in a single database that spans clades, environments and, critically, temporal scales across many orders of magnitude. Combining data on these various aspects, as in BioDeepTime, will enable progress in addressing these priority questions and associated challenges.

The BioDeepTime database combines taxonomically well-resolved assemblage time series of at least 10 time steps from existing modern and fossil databases and from the primary literature. BioDeepTime builds on efforts to synthesize and curate modern and fossil assemblage time series with a high degree of taxonomic resolution and temporal certainty. This version of the new database is focused primarily on taxonomic groups well represented in existing databases like BioTIME (Dornelas et al., 2018), Neotoma (Williams et al., 2018), Neptune Sandbox Berlin (Renaudie et al., 2020) and Triton (Fenton et al., 2021). The BioDeepTime database was created from available data filtered according to criteria for inclusion that enabled the integration and comparison of modern and fossil biodiversity dynamics.

2 | METHODS

BioDeepTime is a relational database (Smith et al., 2023b) created through a working group of the Paleosynthesis Project. Taxon records of presence and abundance are linked to additional tables providing context for each record (see Appendix S1, Figure S1.1). There are three nested levels of organization: record, assemblage sample and assemblage time series (Figure 1). Records are the finest resolution and indicate the observed presence, abundance (e.g. biomass, counts, percent cover) or relative abundance of a single taxon in a single sample of a time series. An assemblage sample (i.e. the observation of an assemblage) is composed of records that belong to the same sampling event with a constrained temporal duration such that all taxon records in a sample are considered to be of the same age. The amount of time contained in a single sample (i.e. temporal grain; Figure 1) may be on the scale of days or hours for some modern time series and up to tens of thousands of years in fossil time series (Table 1). The spatial grain of a sample is consistent within time series, ranging from cm^2 to ~100 km^2, but varies among time series due to factors including life history, underlying sampling processes and taphonomy.

Time series are internally consistent in sampling methodology, location (at time of sampling, rather than deposition) and taxonomy, but there is considerable variation across time series as sampling methods vary among practitioners working on different environments, timescales and taxonomic groups. At each organizational level, contextual variables are provided to facilitate the use of these data for community analyses (Figure S1). These additional variables include extended taxonomic information, references to original data sources, types of abundance units and uncertainties in age, among others. A congruence table is included in Appendix S1 to describe differences in variable names and content originating from the source databases (see 'Synonymy of sources' in Smith et al., 2023b).

2.1 | Data acquisition

Data were compiled from BioTIME (Dornelas et al., 2018, queried June 2021), Neptune Sandbox Berlin (Renaudie et al., 2020, queried October 2022), Neotoma (Williams et al., 2018, queried May 2022) and Triton (Fenton et al., 2021, queried December 2022; see Appendix S2 for information on database-specific queries). A single example time series each from the Paleobiology Database (PBDB, http://paleobiodb.org) and the Geobiodiversity Database (GBDB, http://www.geobiodiversity.com) were included to demonstrate future potential for database expansion. Additional data were aggregated from the literature, including modern planktonic foraminifera time series (from sediment traps—SedTraps, see Appendix S1: Appendix S2.2) and a new database called MARBEN (Tomasovych et al., 2022)—a compilation of marine benthic fossil assemblages collected in sediment cores containing ostracods, molluscs and foraminifera, mostly from the Pleistocene and Holocene (see Appendix S1: Appendix S2.3 for more information on MARBEN). Finally, two time series are included as 'Direct uploads' (see Appendix S1: Appendix S3.1), representing a pathway for inclusion of data that are not already tied to an existing database (e.g. from researchers in the scientific community). A list of the data sources is found in Appendix S5.

The primary criteria for inclusion of data in BioDeepTime were established to create a database of high-quality assemblage-level records for time-series analysis, including accurate age estimates that will enable future analyses that could incorporate environmental data. The criteria were: (i) internally consistent taxonomy within a time series, (ii) original sampling conducive to assemblage analysis and internally consistent, (iii) minimum of 10 samples with unique ages and (iv) age estimates for samples at each time step were reported, or could be inferred from age controls provided, for example, samples in a fossil time series (see Appendix S3 for more detail on criteria). The first criterion of internally consistent taxonomy was imposed to mitigate the effects of variable taxonomic resolution across groups. As the intended biodiversity analyses of BioDeepTime are specific to individual time series (and not comparisons of taxa across

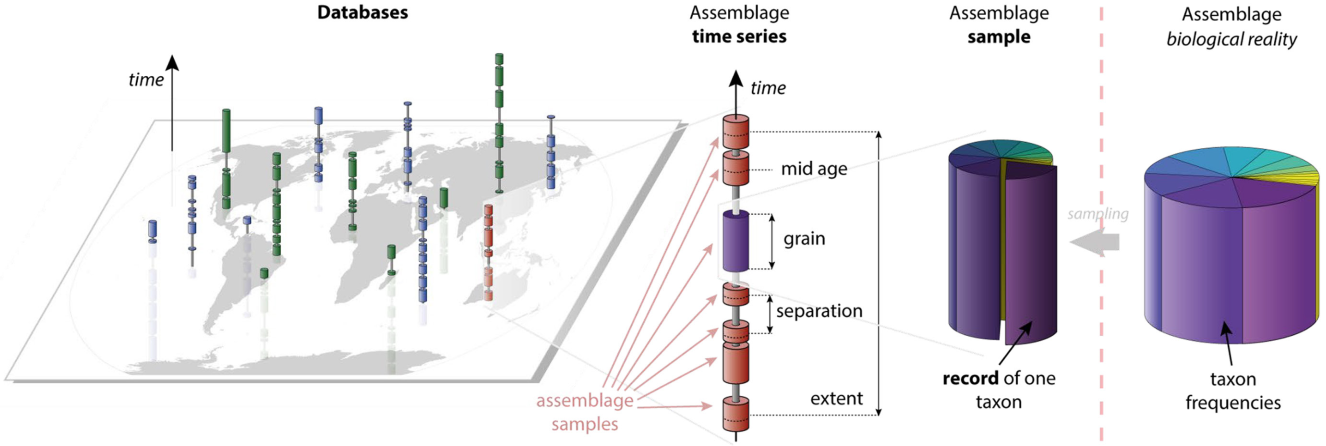

FIGURE 1 The BioDeepTime database includes assemblage time series from existing databases and the literature. Each assemblage time series is composed of multiple samples from biological assemblages, which are distorted by preservational, sampling and other processes. Within a sample, records (e.g. relative abundance, presence) for multiple taxa are included. Assemblage samples are characterized by an age and temporal grain (i.e. total time represented by the sample), with variable temporal separation between samples of different ages. Time series are characterized by their temporal extent (i.e. total duration) and other data including geographic location, taxonomic composition and environment.

TABLE 1 Approximate temporal grain (the amount of time represented in a sample) for time series, number of time series and number of samples from source databases included in BioDeepTime.

Database	Approximate temporal grain	Comments	Number of time series	Number of samples
BioTIME	Minutes to days	Most records were collected in real time, with little temporal averaging	6982	797,425
Direct Upload	–	This group of time series will have variable temporal grain, as they are not derived from a single source (e.g. an existing database)	2	239
GBDB	10^3–10^4 years	Fossil records are based on section data and provide quality control at the level of bed-by-bed rock units	1	33
MARBEN	1–10^4 years	Some marine records are reworked more than others, resulting in a large amount of variation in the temporal grain in samples from this depositional environment	257	13,812
Neotoma	1–10^3 years	Most pollen records used here are from lake and mire sediments from the Quaternary, which tend to have sediment deposition times on the order of 0.1–50 years per cm, and a mode of 10 years per cm (Goring et al., 2012). Typical sediment sample size is 1 cm^3	1358	96,164
Neptune Sandbox Berlin	10^3–10^4 years	Marine records are often reworked and low sedimentation rates in the deeper ocean can result in relatively large temporal grain in this depositional environment	609	31,634
Paleobiology Database	10^3–10^4 years	Records are from deeper in the fossil record where there are greatest uncertainties in temporal grain	1	19
SedTraps	Days to months	Time series are from moored sediment traps, which record export flux integrated over precisely controlled intervals	30	1295
Triton	10^2–10^4 years	Marine records are often reworked and low sedimentation rates in the deeper ocean can result in relatively large temporal grain in this depositional environment	822	44,685

time series), internal consistency in nomenclature within a time series was considered sufficient for inclusion—taxonomic nomenclature of the same group might differ among time series. Time series from most databases have ongoing updates and synonymizations of their taxonomy (e.g. BioTIME, Neotoma, Triton), and future versions of BioDeepTime will incorporate these as they become available.

The sampling standards used to ensure high-quality data for assemblage analyses likewise varied by taxonomic group, environment and timescale, leading us to use a broad criterion for inclusion—conducive to assemblage analysis (i.e. individuals in a sample counted and identified)—to allow for this heterogeneity. For example, with planktonic foraminifera from Triton, a random sample of ~300 individuals from a larger sample are counted and identified to the species level. However, in fish trawls from BioTIME, for example, all individuals in a trawl are counted and identified to the lowest possible taxonomic level (e.g. genus level). Time series were required to have at least 10 samples with unique ages. Some sediment core samples had different depth values but identical age estimates; in these cases, we allowed for 10 samples with unique depths (instead of ages). There was no constraint on the temporal extent of a time series, which ranged from years to millions of years (Figure 3a,c). For modern samples extracted from BioTIME, date of observation is recorded. For fossil samples, ages were inferred using a variety of age constraints and age-depth modelling approaches (see below and Appendix S4).

Version 1.0 of BioDeepTime (Smith et al., 2023b) represents the first milestone of the data gathering process. Data acquisition is ongoing and future versions of the database with significant additions will periodically be released. In addition to new queries of the databases drawn from here, future data integration efforts will (1) aim to expand the taxonomic coverage of the database and (2) incorporate data from more ancient periods beyond the Cenozoic, with a particular emphasis on quantifying the uncertainty of age-depth models in these times (Figure 2).

2.2 | Data curation and quality control

Field names (i.e. data table columns) from existing databases were synonymized for inclusion in BioDeepTime (see Appendix S1, Tables S1.1–S1.17 and 'Synonymy of sources' in Smith et al. (2023b)). To include as many time series as possible, we allowed the methods and types of data used to infer sample ages to vary among data sets, as long as the methods were consistent within each time series. For several databases (BioTIME, Neptune Sandbox Berlin, Triton, SedTraps), the ages assigned to samples in the original database or publication are the ages used in BioDeepTime, as are the uncertainties around those estimates, when available. For Neotoma samples, we generated updated age-depth models for most time series (1281 of 1362) using *Bchron* (Haslett & Parnell, 2008) based

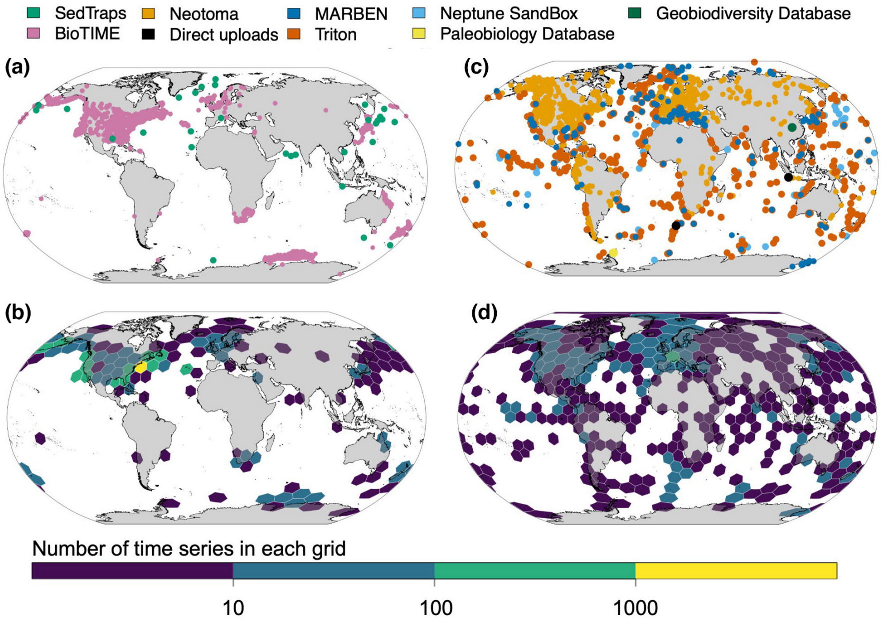

FIGURE 2 Distribution of assemblage time series from modern, neontological (a, b) and fossil, palaeontological (c, d) sources. Specific time-series locations are indicated in (a, c), with spatially binned data shown in (b, d) to visualize the density of time series. The distribution of time series also corresponds to the coverage of environments (e.g. Marine, Terrestrial) in BioDeepTime. Landmass polygons from NaturalEarth (Robinson projection; https://www.naturalearthdata.com/) and grid cells were created using *icosa* (Kocsis, 2017).

on chronological controls stored in Neotoma for each site, which were commonly uncalibrated radiocarbon ages but included biostratigraphic markers (e.g. European settlement in North America), calibrated radiocarbon ages and laminated varves. *Bchron* generates an ensemble of possible age models ($n=1000$) in a Bayesian framework considering the uncertainty in age-depth constraints, including differences in radiocarbon calibration curves (see Appendix S4), with the median age estimate for each sample used as the sample age in BioDeepTime version 1.0 (see 'neotoma_bchron' in Smith et al., 2023b). Although not currently implemented for all time series, this Bayesian approach provides an example of how age models might be improved and standardized across BioDeepTime in future releases and highlights the importance of storing fossil data with the original age controls used for age-depth modelling. Details for age estimates and models used are in Appendix S4.

For a compact data format in BioDeepTime, each entry is a presence (or abundance/biomass, indicating presence) record. However, these can be converted to 'presence-absence' format for those taxa that are not observed in a particular sample but are otherwise observed in at least one other sample in the same time series. That is, observed zeros can be restored by converting from long to wide format (i.e. assuming zeros where data are otherwise lacking).

3 | DESCRIPTION OF DATA

BioDeepTime currently includes 7,437,847 taxon records distributed in 985,306 samples from 10,062 time series collectively spanning 451 million years (Figure 3). These records are from terrestrial, freshwater and marine environments globally (Figure 2), although there is limited coverage in the global south (Figure 2b,d). These geographical gaps reflect documented biases in global data distribution in ecology and palaeontology, which are underlain by inequities that include funding, research infrastructure and geopolitical histories (e.g. Dornelas et al., 2018; Nuñez et al., 2021; Raja et al., 2022). Addressing these geographical biases in the data distribution is a critical and ongoing research goal that is essential for gaining a comprehensive understanding of biodiversity change.

The mean duration of a time series in BioDeepTime is 1.63 million years, with a median duration of 29 years. As BioTIME contributed the largest number of time series ($n=6982$, 69%) and samples ($n=797{,}425$, 81%), the difference in mean and median duration is primarily due to the distribution of samples and ages across the BioTIME data (mean duration = 22.6 years; mean age of modern samples = 25 years old) as compared to the fossil time series (mean duration = 5.38 million years; mean age of samples = 5.09 million years; Figure 3a). The fossil data can be further divided into 'near-time' paleontological data (here defined as having their oldest samples since the start of the Quaternary, 2.58 million years ago) and 'deep-time' data (with oldest samples >2.58 million years). Of the fossil data, 60.1% are near-time, with 649 (21.3%) of the 3050 fossil time series having a maximum age that falls in the Holocene (last 12,000 years) and another 1184 (38.8%) having their oldest records further back in the Quaternary. The remaining 1217 (39.9%) extend into the 'deep time' palaeontological record (>2.58 million years). All but 67 (2.2%) of the fossil time series have a maximum age within the Cenozoic (<66 million years). The paucity of time series from the older fossil record reflects the propensity for increased uncertainty in age estimates and taxonomic assignments in deeper time, which resulted in exclusion from BioDeepTime based on our current criteria.

BioDeepTime includes more than 28,700 taxon entries. When possible, taxa are included at the level of species ($n=4{,}769{,}789$; 64.1%) or genus ($n=271{,}218$; 3.6%). These taxa were subsequently grouped into 16 taxonomic categories and a 'Mixed' category, when a time series includes taxa from multiple groups (Figure 3d), to

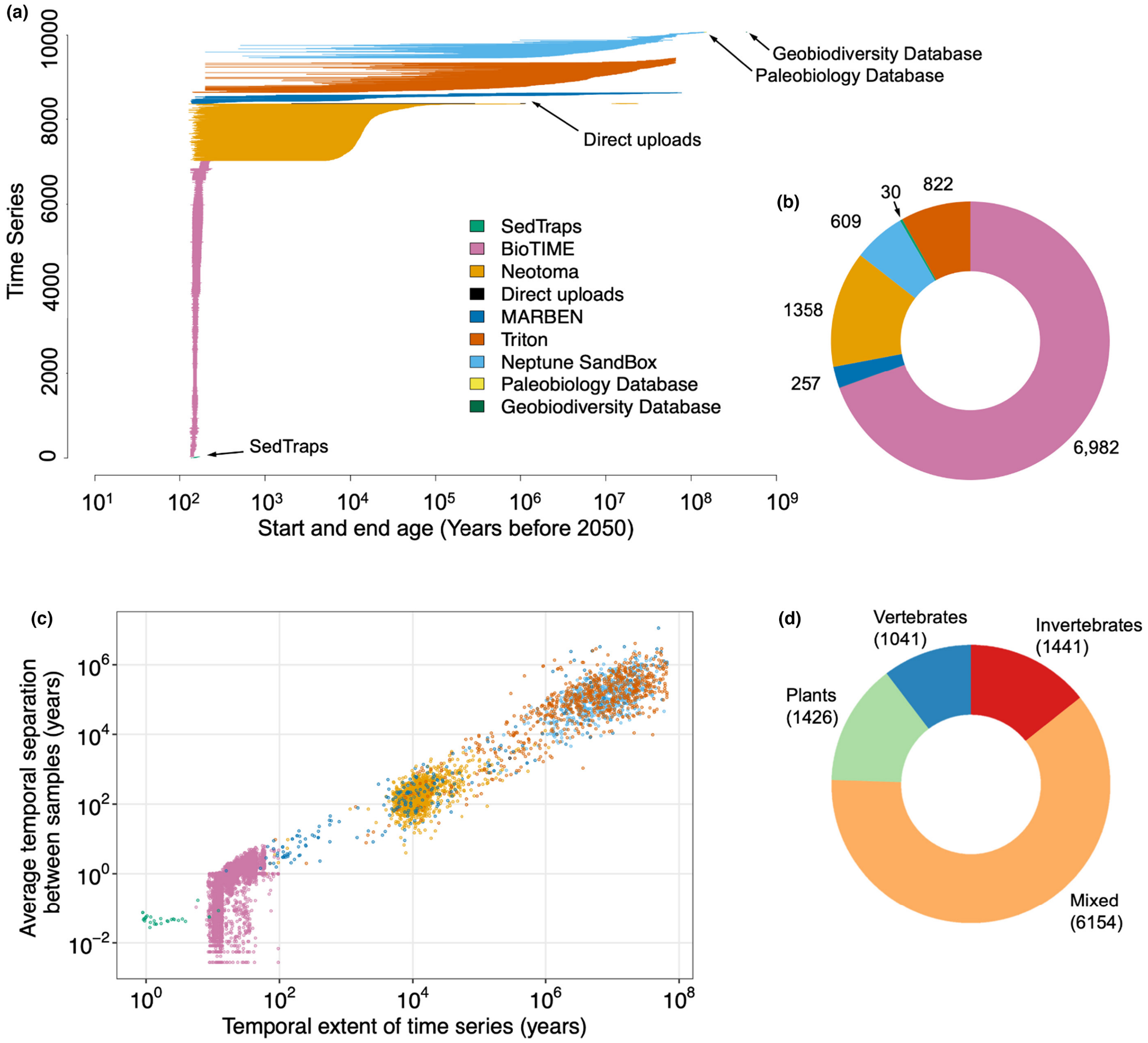

FIGURE 3 (a) Temporal distribution of assemblage time series in geological time, by database. (b) Number of assemblage time series in BioDeepTime drawn from each database, using the same legend as panel (a). (c) Average temporal separation between samples of a time-series plotted against its temporal extent (total duration). (d) Distribution of taxonomic groups in BioDeepTime in broad categories: mixed, invertebrates (e.g. foraminifera, ostracods), plants (e.g. pollen, bryophytes), and vertebrates (e.g. amphibians, mammals). To avoid negative age values in panels (a, c), 100 years were added to the age of each sample so that all dates represent years before 2050 (instead of years before 1950 and negative values from 1950 onwards). Assemblage time series from direct uploads ($n=2$), Paleobiology Database ($n=1$) and Geobiodiversity Database ($n=1$) are present in panels (a–c) but hard to visualize. Note: the colour scheme used in panels (a–c) is consistent and refers to the database from which time series originated; panel d uses an independent colour scheme.

facilitate time-series analysis based on potential groups of interest. These categories are variable in specificity, with higher levels of detail retained when possible (e.g. 'Vertebrates' and 'Amphibians' are both used). Apart from the Mixed category, vascular plants (i.e. pollen) and foraminifera account for the majority of time series (62.8%; Figure 3c). This taxonomic composition reflects, in part, our choice to include the highest quality fossil time series available, as these groups and the method by which they are sampled (i.e. sediment cores) allow for the construction of the most reliable age models.

4 | USAGE NOTES

BioDeepTime version 1.0 can be downloaded from Zenodo (Smith et al., 2023b). All users of BioDeepTime data should cite this present paper, original data sources and papers presenting the contributing databases. In all cases, data users should also follow the data use agreements set by contributing databases (see also Appendix S1: Appendix S3.2). The procedurally compiled relational database is available as an SQLite file, which was denormalized to the level of records (i.e. taxon presence or abundance) and is provided as a comma-separated values file so users familiar with other database products can readily explore and analyse the compiled data.

Users of BioDeepTime should note several features of the database that may affect their analyses. Most notably, the database was compiled for the purpose of examining biodiversity patterns (e.g. observed richness, turnover) within time series, and in such a way that future analyses can readily evaluate these patterns relative to environmental data. Variation across time series in the underlying data included in BioDeepTime—in time (age models), space (sampling area), taxonomic resolution, collection methods and other biases (e.g. preservational biases)—can be a strength of the database or can become a source of biases. In general, as the temporal extent of a time series increases, so does the distance between samples (Figure 3c) and the temporal grain of the sample. All ages use 1950 as a reference point, with positive values for samples collected before 1950 and negative values for those collected after 1950.

When a fossil assemblage has accumulated over an extended period of time, the sampled assemblage will be time-averaged and its observed richness and evenness often will be higher than found in a living assemblage observed at any given point during the same period of time (e.g. as observed in a 'snapshot' neontological survey; Adler et al., 2005; Olszewski & Kidwell, 2007; Tomašových & Kidwell, 2010). Increases in temporal grain and time averaging result in a prolonged sampling window. The effects of prolonging this sample window are multifarious and have well-known effects on diversity and turnover (e.g. Jarzyna & Jetz, 2018; Olszewski & Kidwell, 2007; Tomašových & Kidwell, 2010). Depending on the analysis, these effects can be accounted for as an uncertainty on diversity estimates (e.g. Smith et al., 2021) and a modifying term for rate-of-change analysis (e.g. Lotter et al., 1992). Temporal grain in fossil samples, approximated by the inverse of sedimentation rate (i.e. the amount of time represented in a sample; not accounting for bioturbation), ranges from days to tens of thousands of years (Table 1).

Some data sets included here have information on sampling effort, but many do not and this information is important for macroscale biodiversity analyses that correct for the effect of sampling effort on observed species richness by, for example, standardization of sampling using rarefaction methods (Chao & Jost, 2012; Roswell et al., 2021). We encourage authors to include data related to sampling effort in future published data sets to enable a broader range of cross-community analyses.

5 | MOVING FORWARD

The BioDeepTime database enables integrated biodiversity analyses across a far greater range of temporal scales than has previously been possible. It can be used to provide critical insights into how natural systems will respond to ongoing and future environmental changes as well as new opportunities for theoretical insights into the temporal scaling of biodiversity dynamics. The taxonomic and environmental scope of BioDeepTime are broad but still encompass only a small fraction of the potentially suitable data. The database will continue to grow and we welcome data contributions from the broader scientific community to expand the utility of this community resource (see Appendix S3).

AFFILIATIONS

[1]GeoZentrum Nordbayern, Friedrich-Alexander Universität Erlangen-Nürnberg (FAU), Erlangen, Germany
[2]Paleontological Research Institution, Ithaca, New York, USA
[3]ICBM – Institute for Chemistry and Biology of the Marine Environment, University of Oldenburg, Wilhelmshaven, Germany
[4]University of St Andrews, St Andrews, UK
[5]Guia Marine Lab, MARE, Faculty of Science of the University of Lisbon, Cascais, Portugal
[6]Department of Geography, University of Wisconsin-Madison, Madison, Wisconsin, USA
[7]Department of Earth and Environmental Sciences, Syracuse University, Syracuse, New York, USA
[8]Department of Paleobiology, National Museum of Natural History, Smithsonian Institution, Washington DC, USA
[9]MARUM – Center for Marine Environmental Sciences University of Bremen, Bremen, Germany
[10]State Key Laboratory of Palaeobiology and Stratigraphy, Center for Excellence in Life and Paleoenvironment, Chinese Academy of Sciences, Nanjing Institute of Geology and Palaeontology, Nanjing, China
[11]Natural History Museum, University of Oslo, Oslo, Norway
[12]Department of Earth and Planetary Sciences, Yale University, New Haven, Connecticut, USA
[13]Department of Geoscience, University of Wisconsin – Madison, Madison, Wisconsin, USA
[14]Museum für Naturkunde, Berlin, Germany
[15]Department of Earth Sciences, University of Oxford, Oxford, UK
[16]Bayreuth Center of Ecology and Environmental Research, University of Bayreuth, Bayreuth, Germany
[17]University of British Columbia, Vancouver, British Columbia, Canada
[18]Mário Schenberg Institute, São Paulo, São Paulo, Brazil
[19]Slovak Academy of Sciences, Earth Science Institute, Bratislava, Slovakia
[20]Area of Ecology and Biodiversity, School of Biological Sciences, Swire Institute of Marine Science, Institute for Climate and Carbon Neutrality, Musketeers Foundation Institute of Data Science, The University of Hong Kong, Kadoorie Biological Sciences Building, Hong Kong, Special Administrative Regions of China, China
[21]State Key Laboratory of Marine Pollution, City University of Hong Kong, Hong Kong, Special Administrative Regions of China, China
[22]Department of Integrative Biology, Museum of Paleontology, University of California, Berkeley, Valley Life Sciences Building, Berkeley, California, USA
[23]Yale Peabody Museum, New Haven, Connecticut, USA

ACKNOWLEDGEMENTS

We thank the Paleosynthesis Project and the Volkswagen Stiftung for funding that supported this project (Az 96 796), as well as Dr. Barbara Seuss for her indefatigable support. We also thank Jacquelyn Gill (editor) and two anonymous reviewers for their comments on an earlier version of the manuscript, which improved its quality. M.C.R. acknowledges the German Research Foundation (DFG) for funding through the Cluster of Excellence 'The Ocean Floor – Earth's Uncharted Interface' (EXC 2077, grant no. 390741603). E.E.S. acknowledges funding from Leverhulme Trust grant RPG-201170, the Leverhulme Prize and the National Science Research Council grant NE/V011405/1. Q.J.L. and L.N. acknowledge support from the Youth Innovation Promotion Association (2019310) and the Chinese

Academy of Sciences (CAS-WX2021SF-0205). A.M.P. acknowledges funding from the Leverhulme Trust through research grant RPG-2019-402, and also thanks Dr Faye Moyes (University of St Andrews) for assistance with taxonomic work on data from the BioTIME database. M.D. acknowledges funding from Leverhulme Trust through the Leverhulme Centre for Anthropocene Biodiversity (RC-2018-021) and a research grant (RPG-2019-402), and the European Union (ERC coralINT, 101044975). L. H. L. acknowledges funding from the European Research Council (macroevolution.abc ERC grant no. 724324). K.H.P acknowledges funding from the National Science Foundation Graduate Research Fellowship Program (DGE-2139841). H.H.M.H. acknowledges support from Peter Buck Postdoc Fellowship, Smithsonian Institution. A.T. acknowledges funding from the Slovak Research and Development Agency (APVV 22-0523) and the Slovak Scientific Grant Agency (VEGA 02/0106/23). Neotoma data were obtained from the Neotoma Paleoecology Database (http://www.neotomadb.org) on 25 May 2022 and its constituent database(s) the African Pollen Database, Alpine Palynological Database, European Pollen Database, Indo-Pacific Pollen Database, Japan Pollen Database, Latin American Pollen Database and North American Pollen Database. The work of data contributors, data stewards and the Neotoma community is gratefully acknowledged. We are grateful to the BioTIME core team and all contributors to the BioTIME consortium. The views and opinions expressed in this manuscript are those of the authors only and do not necessarily reflect those of the funding bodies (e.g. European Union, European Research Council, National Science Foundation, Chinese Academy of Sciences) and those funding bodies and associated organizations cannot be held responsible for them. Open Access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST STATEMENT

None.

DATA AVAILABILITY STATEMENT

The BioDeepTime database and the code to reproduce the data compilation are accessible from the Zenodo repository (Smith et al., 2023b). A second repository with code required to reproduce the figures and summary statistics in this paper is present at Smith et al. (2023a).

ORCID

Jansen Smith https://orcid.org/0000-0003-3771-2039
Marina C. Rillo https://orcid.org/0000-0002-2471-0002
Ádám T. Kocsis https://orcid.org/0000-0002-9028-665X
Maria Dornelas https://orcid.org/0000-0003-2077-7055
David Fastovich https://orcid.org/0000-0002-0340-9819
Huai-Hsuan M. Huang https://orcid.org/0000-0003-2764-5995
Lukas Jonkers https://orcid.org/0000-0002-0253-2639
Wolfgang Kiessling https://orcid.org/0000-0002-1088-2014
Lee Hsiang Liow https://orcid.org/0000-0002-3732-6069
Stephen Meyers https://orcid.org/0000-0003-4422-720X
Lin Na https://orcid.org/0000-0001-5290-6152
Amelia M. Penny https://orcid.org/0000-0002-4392-8090
Johan Renaudie https://orcid.org/0000-0002-9107-1984
Erin E. Saupe https://orcid.org/0000-0002-0370-9897
Manuel J. Steinbauer https://orcid.org/0000-0002-7142-9272
Mauro Sugawara https://orcid.org/0000-0003-0403-8586
Adam Tomašových https://orcid.org/0000-0002-0471-9480
John W. Williams https://orcid.org/0000-0001-6046-9634
Moriaki Yasuhara https://orcid.org/0000-0003-0990-1764
Seth Finnegan https://orcid.org/0000-0002-6175-6173
Pincelli M. Hull https://orcid.org/0000-0001-8607-4817

REFERENCES

Adler, P. B., White, E. P., Lauenroth, W. K., Kaufman, D. M., Rassweiler, A., & Rusak, J. A. (2005). Evidence for a general species-time-area relationship. *Ecology*, *86*, 2032–2039.

Antão, L. H., Bates, A. E., Blowes, S. A., Waldock, C., Supp, S. R., Magurran, A. E., Dornelas, M., & Schipper, A. M. (2020). Temperature-related biodiversity change across temperate marine and terrestrial systems. *Nature Ecology & Evolution*, *4*, 927–933.

Barnosky, A. D., Hadly, E. A., Gonzalez, P., Head, J., Polly, P. D., Lawing, A. M., Eronen, J. T., Ackerly, D. D., Alex, K., Biber, E., Blois, J., Brashares, J., Ceballos, G., Davis, E., Dietl, G. P., Dirzo, R., Doremus, H., Fortelius, M., Greene, H. W., … Zhang, Z. (2017). Merging paleobiology with conservation biology to guide the future of terrestrial ecosystems. *Science*, *355*, eaah4787.

Blowes, S. A., Supp, S. R., Antão, L. H., Bates, A., Bruelheide, H., Chase, J. M., Moyes, F., Magurran, A., McGill, B., & Myers-Smith, I. H. (2019). The geography of biodiversity change in marine and terrestrial assemblages. *Science*, *366*, 339–345.

Chao, A., & Jost, L. (2012). Coverage-based rarefaction and extrapolation: Standardizing samples by completeness rather than size. *Ecology*, *93*, 2533–2547.

Chavez, F. P., Ryan, J., Lluch-Cota, S. E., & Ñiquen, C. M. (2003). From anchovies to sardines and back: Multidecadal change in the Pacific Ocean. *Science*, *299*, 217–221.

Dietl, G. P., Kidwell, S. M., Brenner, M., Burney, D. A., Flessa, K. W., Jackson, S. T., & Koch, P. L. (2015). Conservation paleobiology: Leveraging knowledge of the past to inform conservation and restoration. *Annual Review of Earth and Planetary Sciences*, *43*, 79–103.

Dietl, G. P., Smith, J. A., & Durham, S. R. (2019). Discounting the past: The undervaluing of paleontological data in conservation science. *Frontiers in Ecology and Evolution*, *7*, 108.

Dillon, E. M., Pier, J. Q., Smith, J. A., Raja, N. B., Dimitrijević, D., Austin, E. L., Cybulski, J. D., De Entrambasaguas, J., Durham, S. R., & Grether, C. (2022). What is conservation paleobiology? Tracking 20 years of research and development. *Frontiers in Ecology and Evolution*, *10*, 1117.

Dirzo, R., Young, H. S., Galetti, M., Ceballos, G., Isaac, N. J., & Collen, B. (2014). Defaunation in the Anthropocene. *Science*, *345*, 401–406.

Dornelas, M., Antao, L. H., Moyes, F., Bates, A. E., Magurran, A. E., Adam, D., Akhmetzhanova, A. A., Appeltans, W., Arcos, J. M., & Arnold, H. (2018). BioTIME: A database of biodiversity time series for the Anthropocene. *Global Ecology and Biogeography*, *27*, 760–786.

Dornelas, M., Gotelli, N. J., McGill, B., Shimadzu, H., Moyes, F., Sievers, C., & Magurran, A. E. (2014). Assemblage time series reveal biodiversity change but not systematic loss. *Science*, *344*, 296–299.

Ellis, E. C., Kaplan, J. O., Fuller, D. Q., Vavrus, S., Klein Goldewijk, K., & Verburg, P. H. (2013). Used planet: A global history. *Proceedings of the National Academy of Sciences of the United States of America*, *110*, 7978–7985.

Fenton, I. S., Woodhouse, A., Aze, T., Lazarus, D., Renaudie, J., Dunhill, A. M., Young, J. R., & Saupe, E. E. (2021). Triton, a new species-level database of Cenozoic planktonic foraminiferal occurrences. *Scientific Data*, *8*, 160.

Foote, M. (1994). Temporal variation in extinction risk and temporal scaling of extinction metrics. *Paleobiology*, *20*, 424–444.

Gonzalez, A., Cardinale, B. J., Allington, G. R. H., Byrnes, J., Arthur Endsley, K., Brown, D. G., Hooper, D. U., Isbell, F., O'Connor, M. I., & Loreau, M. (2016). Estimating local biodiversity change: A critique of papers claiming no net loss of local diversity. *Ecology*, *97*, 1949–1960.

Goring, S., Williams, J. W., Blois, J. L., Jackson, S. T., Paciorek, C. J., Booth, R. K., Marlon, J. R., Blaauw, M., & Christen, J. A. (2012). Deposition times in the northeastern United States during the Holocene: Establishing valid priors for Bayesian age models. *Quaternary Science Reviews*, *48*, 54–60.

Gotelli, N. J., Shimadzu, H., Dornelas, M., McGill, B., Moyes, F., & Magurran, A. E. (2017). Community-level regulation of temporal trends in biodiversity. *Science Advances*, *3*, e1700315.

Haslett, J., & Parnell, A. (2008). A simple monotone process with application to radiocarbon-dated depth chronologies. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *57*, 399–418.

Henao Diaz, L. F., Harmon, L. J., Sugawara, M. T. C., Miller, E. T., & Pennell, M. W. (2019). Macroevolutionary diversification rates show time dependency. *Proceedings of the National Academy of Sciences of the United States of America*, *116*, 7403–7408.

IPBES (2019). In E. S. Brondizio, J. Settele, S. Díaz, & H. T. Ngo (Eds.), *Global assessment report on biodiversity and ecosystem services of the intergovernmental science-policy platform on biodiversity and ecosystem services*. IPBES Secretariat.

Jackson, J. B., Kirby, M. X., Berger, W. H., Bjorndal, K. A., Botsford, L. W., Bourque, B. J., Bradbury, R. H., Cooke, R., Erlandson, J., & Estes, J. A. (2001). Historical overfishing and the recent collapse of coastal ecosystems. *Science*, *293*, 629–637.

Jarzyna, M. A., & Jetz, W. (2018). Taxonomic and functional diversity change is scale dependent. *Nature Communications*, *9*, 2565.

Kemp, D. B., Eichenseer, K., & Kiessling, W. (2015). Maximum rates of climate change are systematically underestimated in the geological record. *Nature Communications*, *6*, 8890.

Kiessling, W., Smith, J. A., & Raja, N. B. (2023). Improving the relevance of paleontology to climate change policy. *Proceedings of the National Academy of Sciences of the United States of America*, *120*, e2201926119.

Koch, P. L., & Barnosky, A. D. (2006). Late quaternary extinctions: State of the debate. *Annual Review of Ecology, Evolution, and Systematics*, *37*, 215–250.

Kocsis, Á.T. (2017). The R package icosa: Coarse resolution global triangular and penta-hexagonal grids based on tessellated icosahedra.

Kuwae, M., Yamamoto, M., Sagawa, T., Ikehara, K., Irino, T., Takemura, K., Takeoka, H., & Sugimoto, T. (2017). Multidecadal, centennial, and millennial variability in sardine and anchovy abundances in the western North Pacific and climate–fish linkages during the late Holocene. *Progress in Oceanography*, *159*, 86–98.

Lewandowska, A. M., Jonkers, L., Auel, H., Freund, J. A., Hagen, W., Kucera, M., & Hillebrand, H. (2020). Scale dependence of temporal biodiversity change in modern and fossil marine plankton. *Global Ecology and Biogeography*, *29*, 1008–1019.

Lotter, A. F., Ammann, B., & Sturm, M. (1992). Rates of change and chronological problems during the late-glacial period. *Climate Dynamics*, *6*, 233–239.

Lotze, H. K., Lenihan, H. S., Bourque, B. J., Bradbury, R. H., Cooke, R. G., Kay, M. C., Kidwell, S. M., Kirby, M. X., Peterson, C. H., & Jackson, J. B. C. (2006). Depletion, degradation, and recovery potential of estuaries and coastal seas. *Science*, *312*, 1806–1809.

Louca, S., Henao-Diaz, L. F., & Pennell, M. (2022). The scaling of diversification rates with age is likely explained by sampling bias. *Evolution*, *76*, 1625–1637.

MacDonald, G. M., Bennett, K. D., Jackson, S. T., Parducci, L., Smith, F. A., Smol, J. P., & Willis, K. J. (2008). Impacts of climate change on species, populations and communities: Palaeobiogeographical insights and frontiers. *Progress in Physical Geography*, *32*, 139–172.

McCauley, D. J., Pinsky, M. L., Palumbi, S. R., Estes, J. A., Joyce, F. H., & Warner, R. R. (2015). Marine defaunation: Animal loss in the global ocean. *Science*, *347*, 1255641.

McGill, B. J., Dornelas, M., Gotelli, N. J., & Magurran, A. E. (2015). Fifteen forms of biodiversity trend in the Anthropocene. *Trends in Ecology & Evolution*, *30*, 104–113.

Mottl, O., Flantua, S. G. A., Bhatta, K. P., Felde, V. A., Giesecke, T., Goring, S., Grimm, E. C., Haberle, S., Hooghiemstra, H., Ivory, S., Kuneš, P., Wolters, S., Seddon, A. W. R., & Williams, J. W. (2021). Global acceleration in rates of vegetation change over the past 18,000 years. *Science*, *372*, 860–864.

Nuñez, M. A., Chiuffo, M. C., Pauchard, A., & Zenni, R. D. (2021). Making ecology really global. *Trends in Ecology & Evolution*, *36*, 766–769.

Olszewski, T. D., & Kidwell, S. M. (2007). The preservational fidelity of evenness in molluscan death assemblages. *Paleobiology*, *33*, 1–23.

Pauly, D., Christensen, V., Dalsgaard, J., Froese, R., & Torres, F., Jr. (1998). Fishing down marine food webs. *Science*, *279*, 860–863.

Pimm, S. L., Jenkins, C. N., Abell, R., Brooks, T. M., Gittleman, J. L., Joppa, L. N., Raven, P. H., Roberts, C. M., & Sexton, J. O. (2014). The biodiversity of species and their rates of extinction, distribution, and protection. *Science*, *344*, 1246752.

Primack, R. B., Miller-Rushing, A. J., Corlett, R. T., Devictor, V., Johns, D. M., Loyola, R., Maas, B., Pakeman, R. J., & Pejchar, L. (2018). Biodiversity gains? The debate on changes in local- vs global-scale species richness. *Biological Conservation*, *219*, A1–A3.

Raja, N. B., Dunne, E. M., Matiwane, A., Khan, T. M., Nätscher, P. S., Ghilardi, A. M., & Chattopadhyay, D. (2022). Colonial history and global economics distort our understanding of deep-time biodiversity. *Nature Ecology & Evolution*, *6*, 145–154.

Renaudie, J., Lazarus, D., & Diver, P. (2020). NSB (Neptune Sandbox Berlin): An expanded and improved database of marine planktonic microfossil data and deep-sea stratigraphy. *Palaeontologia Electronica*, *23*, a11.

Roswell, M., Dushoff, J., & Winfree, R. (2021). A conceptual guide to measuring species diversity. *Oikos*, *130*, 321–338.

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

Smith, J. A., Dietl, G. P., & Durham, S. R. (2020). Increasing the salience of marine live–dead data in the Anthropocene. *Paleobiology*, *46*, 279–287.

Smith, J. A., Handley, J. C., & Dietl, G. P. (2021). Accounting for uncertainty from zero inflation and overdispersion in paleoecological studies of predation using a hierarchical Bayesian framework. *Paleobiology*, *48*, 65–82.

Smith, J. A., Rillo, M. C., Kocsis, Á. T., Dornelas, M., Fastovich, D., Huang, H.-H. M., Jonkers, L., Kiessling, W., Li, Q., Liow, L. H., Margulis-Ohnuma, M., Meyers, S., Na, L., Penny, A., Pippenger, K., Renaudie, J., Saupe, E. E., Steinbauer, M. J., Sugawara, M. T. C., … Hull, P. M. (2023a). *Code and data used for the study: 'BioDeepTime: A database of biodiversity time series for modern and fossil assemblages'*. Zenodo.

Smith, J. A., Rillo, M. C., Kocsis, Á. T., Dornelas, M., Fastovich, D., Huang, H.-H. M., Jonkers, L., Kiessling, W., Li, Q., Liow, L. H., Margulis-Ohnuma, M., Meyers, S., Na, L., Penny, A. M., Pippenger, K., Renaudie, J., Saupe, E. E., Steinbauer, M., Sugawara, M., … Hull, P. (2023b). *BioDeepTime: Database and compilation code*. Zenodo.

Spalding, C., & Hull, P. M. (2021). Towards quantifying the mass extinction debt of the Anthropocene. *Proceedings of the Royal Society B: Biological Sciences*, *288*, 20202332.

Tomasovych, A., Huang, H.-H. M., & Yasuhara, M. (2022). MARBEN database (Version 01a) [Data set]. Zenodo.

Tomašových, A., & Kidwell, S. M. (2010). Predicting the effects of increasing temporal scale on species composition, diversity, and rank-abundance distributions. *Paleobiology*, *36*, 672–695.

Valdés, J., Ortlieb, L., Gutierrez, D., Marinovic, L., Vargas, G., & Sifeddine, A. (2008). 250 years of sardine and anchovy scale deposition record in Mejillones Bay, northern Chile. *Progress in Oceanography*, *79*, 198–207.

Vellend, M., Baeten, L., Myers-Smith, I. H., Elmendorf, S. C., Beauséjour, R., Brown, C. D., De Frenne, P., Verheyen, K., & Wipf, S. (2013). Global meta-analysis reveals no net change in local-scale plant biodiversity over time. *Proceedings of the National Academy of Sciences of the United States of America*, *110*, 19456–19459.

Vellend, M., Dornelas, M., Baeten, L., Beauséjour, R., Brown, C. D., De Frenne, P., Elmendorf, S. C., Gotelli, N. J., Moyes, F., & Myers-Smith, I. H. (2017). Estimates of local biodiversity change over time stand up to scrutiny. *Ecology*, *98*, 583–590.

Westerhold, T., Marwan, N., Drury, A. J., Liebrand, D., Agnini, C., Anagnostou, E., Barnet, J. S. K., Bohaty, S. M., Vleeschouwer, D. D., Florindo, F., Frederichs, T., Hodell, D. A., Holbourn, A. E., Kroon, D., Lauretano, V., Littler, K., Lourens, L. J., Lyle, M., Pälike, H., … Zachos, J. C. (2020). An astronomically dated record of Earth's climate and its predictability over the last 66 million years. *Science*, *369*(6509), 1383–1387.

Williams, J. W., Grimm, E. C., Blois, J. L., Charles, D. F., Davis, E. B., Goring, S. J., Graham, R. W., Smith, A. J., Anderson, M., & Arroyo-Cabrales, J. (2018). The Neotoma Paleoecology Database, a multiproxy, international, community-curated data resource. *Quaternary Research*, *89*, 156–177.

Williams, J. W., Ordonez, A., & Svenning, J.-C. (2021). A unifying framework for studying and managing climate-driven rates of ecological change. *Nature Ecology & Evolution*, *5*, 17–26.

Wolkovich, E. M., Cook, B. I., McLauchlan, K. K., & Davies, T. J. (2014). Temporal ecology in the Anthropocene. *Ecology letters*, *17*, 1365–1379.

Yasuhara, M., & Deutsch, C. A. (2022). Paleobiology provides glimpses of future ocean: Fossil records from tropical oceans predict biodiversity loss in a warmer world. *Science*, *375*, 25–26.

Yasuhara, M., & Deutsch, C. A. (2023). Tropical biodiversity linked to polar climate. *Nature*, *614*, 626–628.

Yasuhara, M., Hunt, G., Breitburg, D., Tsujimoto, A., & Katsuki, K. (2012). Human-induced marine ecological degradation: Micropaleontological perspectives. *Ecology and Evolution*, *2*, 3242–3268.

Yasuhara, M., Okahashi, H., Cronin, T. M., Rasmussen, T. L., & Hunt, G. (2014). Response of deep-sea biodiversity to abrupt deglacial and Holocene climate changes in the North Atlantic Ocean. *Global Ecology and Biogeography*, *23*, 957–967.

Yasuhara, M., Wei, C.-L., Kucera, M., Costello, M. J., Tittensor, D. P., Kiessling, W., Bonebrake, T. C., Tabor, C. R., Feng, R., Baselga, A., Kretschmer, K., Kusumoto, B., & Kubota, Y. (2020). Past and future decline of tropical pelagic biodiversity. *Proceedings of the National Academy of Sciences of the United States of America*, *117*, 12891–12896.

BIOSKETCH

The BioDeepTime working group (funded by the Paleosynthesis Project) seeks to address one of the central challenges in biodiversity science by compiling and harmonizing ecological time series from modern and fossil sources to investigate the scaling and drivers of biodiversity and community dynamics across timescales ranging from months to millions of years. The working group formed in 2020 and continues to pursue questions related to temporal dynamics of biodiversity using neontological and palaeontological records in the BioDeepTime database and beyond.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Smith, J., Rillo, M. C., Kocsis, Á. T., Dornelas, M., Fastovich, D., Huang, H.-H. M., Jonkers, L., Kiessling, W., Li, Q., Liow, L. H., Margulis-Ohnuma, M., Meyers, S., Na, L., Penny, A. M., Pippenger, K., Renaudie, J., Saupe, E. E., Steinbauer, M. J., Sugawara, M. … Hull, P. M. (2023). BioDeepTime: A database of biodiversity time series for modern and fossil assemblages. *Global Ecology and Biogeography*, *32*, 1680–1689. https://doi.org/10.1111/geb.13735