

Supplementary Materials for

5 **The comparative strength of biodiversity trends in marine and terrestrial assemblages**

Shane A. Blowes and Sarah R. Supp, Laura H. Antão, Amanda Bates, Helge
Bruehlheide, Jonathan M. Chase, Faye Moyes, Anne Magurran, Brian McGill,
10 Isla Myers-Smith, Marten Winter, Anne D. Bjorkman, Diana Bowler, Jarrett
E.K. Byrnes, Andrew Gonzalez, Jes Hines, Forest Isbell, Holly Jones, Laetitia
M. Navarro, Patrick Thompson, Mark Vellend, Conor Waldock, Maria Dornelas

Correspondence to: sablowes@gmail.com, supps@dension.edu, maadd@st-andrews.ac.uk
15

This PDF file includes:

Materials and Methods
Figs. S1 to S18
20 Table S1

25 **Materials and Methods**

Data description and pre-processing

The BioTIME database represents the largest global effort mobilizing assemblage time series.
It includes 386 studies, and currently holds over 12 million records of abundance for over 45
30 thousand species across plants, invertebrates, fish, birds and mammals (37). Analyses
presented in this study used only time series of abundance data (i.e., studies that recorded
counts of the number of individuals for each species in an assemblage).

As we were interested in quantifying biodiversity change at the local scale, studies with multiple sampling locations and extents greater than 71.7 km² (n = 126) were partitioned into 96 km² equal area icosahedron hexagonal grid cells (39). This threshold was determined using the spatial extents of studies that reported sampling in only a single location, and was

5 calculated as the mean plus one standard deviation of the extent in these studies. Studies with a single location, and those with extents < 71.7 km² were assigned to the grid cell in which their centre latitude and longitude were located. The sample locations from all other studies were assigned to cells based on the latitude and longitude of individual samples, gridding the initial sampling extent of a study into multiple different cells. Each cell-level time series was

10 given a unique identifier that was the concatenation of the study ID and the cell reference number, allowing the integrity of each study within each grid cell to be retained for analyses. We then collated species within each unique study-cell combination for each year, resulting in new assemblage time series within grid cells. Most grid cells contained only a single time series (n = 32,878). For those that have more than one time series (n = 6,487), our

15 concatenation of study ID and the cell reference number means that each cell-level time series in our analysis was comprised of samples from only one study, and that important study-level considerations (e.g., sampling method) were consistent within each time series. In total, this process yielded 51,932 time series distributed among 39,365 cells.

20 To minimize the effect of unobserved species on our estimates of biodiversity change, we calculated the abundance-based coverage (60) of each (annual) sample (mean = 0.95, sd = 0.11) within each cell-level time series, and removed all samples with coverage less 0.85. This means that for the remaining time series included in our analysis, there was a <15% chance that another sample in any given year of one more individual would represent a new

25 species.

Finally, and before calculating our measures of biodiversity, we used sample-based rarefaction to standardize the number of samples per year within each cell-level time series. This procedure prevents temporal variation in sampling effort from affecting diversity estimates (61). Within each time series, we counted the number of samples taken in each year, and identified the minimum. This minimum was then used to randomly resample each year down to that number of samples, after which dissimilarity metrics of community composition and species richness were calculated. We repeated this process 199 times for each time series, recorded the values and took the median for dissimilarity and species richness in each year. Using the median instead of the mean reduced the effect of any outlier samples on our estimate, and meant we did not need to round to get an integer value for species richness. We calculated community dissimilarity using pairwise Jaccard dissimilarity measured between the first year and each subsequent year in the time series; and, species richness for each year of the time series. To assess if changes in community composition were driven by species replacement or changes in species richness, we partitioned total Jaccard dissimilarity in the additive components of turnover and nestedness (45, 62). Additionally, to examine whether any trends in compositional change were sensitive to our choice of comparisons with the initial assemblage, we also calculated pairwise turnover and nestedness components of dissimilarity between consecutive assemblages (i.e., t_1 compared with t_0 , t_2 compared with t_1 , t_3 compared with t_2 , etc.).

20

The times series resulting from the gridding, filtering and standardization processes had a mean duration of 5.5 years. To maximize the spatial and temporal coverage of the data, we chose to include time series with only two samples, and the minimum duration (i.e., the time elapsed between the samples) for these two-point time series was three years, but many have longer durations (median = 6 years, maximum = 55 years). The time series span from the late 1800s to the present, with most data collected in the past 40 years (Fig. S2). The data set

25

includes 9013 time series with duration spanning more than two decades and 27,619 spanning more than one. Temporal extent and start date varied substantially in the data, and because time series have such varied start dates, the data includes more than 1500 time series at any one point since the mid 1960s (Fig. S2).

5

Models of biodiversity change

We examined geographic patterns of biodiversity change using three complementary hierarchical linear models. All models nested the cell-level time series into the original studies from which they originated at the lowest levels, but differed in how these studies were grouped geographically. Grouping the cell-level time series into studies accounts for the non-independence of time series from within studies, and the higher-level groupings (described below) allowed us to characterize biodiversity change for different levels of the data. For concision, in the main text we focus on the biome-taxon model that has the richest detail and the realm-latitude-taxon that is the simplest. However, we used all three models to examine the robustness of our results to how the spatial groupings were defined.

Our first and most detailed model nested cells and studies first into taxonomic-habitat groups, and secondly into ecological biomes; we refer to this model as the biome-taxon (BT) model. The taxonomic-habitat groupings reflect the BioTIME metadata for each study and were amphibians, benthos, birds, fish, invertebrates, mammals, marine plants/invertebrates, plants, and multiple taxa (indicating studies that measured more than one taxonomic group). Cell-level time series were assigned ecological biomes using the geographic center of the samples within each cell. Specifically, where samples within a cell came from a single location, that coordinate was used to assign the biome; for cells with samples from multiple locations, biome was assigned based on the center of a convex hull drawn around all the samples within the cell. We used biomes from the published Ecoregions of the World (EOW) datasets

available from *The Nature Conservancy* website (http://maps.tnc.org/gis_data.html, (40-43)). Specifically, for the terrestrial realm, we used the EOW biome classification that is based largely on vegetation types (40). For freshwater regions, we used the Major Habitat Types (MHTs) that are considered to be roughly equivalent to the EOW biomes for terrestrial systems (42). Our marine biomes used the province-level from the Marine Ecoregions of the World; these are coastal and shelf areas expected to be of relatively distinct biota (41). Where cell locations fell outside these EOW classifications, biomes were assigned using the nearest appropriate (terrestrial biome, freshwater MHT, marine province) neighboring EOW group, specifically: terrestrial: $n_{\text{study}} = 11$, $n_{\text{time series}} = 27$; freshwater: $n_{\text{study}} = 3$, $n_{\text{time series}} = 52$; marine: $n_{\text{study}} = 12$, $n_{\text{time series}} = 4526$). Our data span 10 of the 14 terrestrial EOW biomes defined globally, 5 of the 12 freshwater MHTs, and 33 of the 62 marine provinces. This model allowed us to characterize variation at the level of biomes, taxon groups within biomes, and among studies of the same taxon group within biomes.

Our model of intermediate complexity again first nested cells into studies. However, as consistent biome classifications across realms were not available (e.g., both terrestrial and freshwater biomes incorporate more detail of specific habitats compared to the marine biomes that are based more strongly on geography), we replaced the terrestrial and freshwater biomes with broader spatial groupings. Specifically, studies in terrestrial and freshwater systems were grouped into continents, and we retained the marine biomes for studies in the marine realm. Cells and studies were nested into a new grouping covariate that was the concatenation of realm, region and taxon, and we refer to the model as the realm-region-taxon (RRT) model. Realm was one of marine, terrestrial or freshwater; region was a continent (i.e., Africa, Asia, Australia, Europe, North America, and South America) for studies in the terrestrial and freshwater realms, and biome (defined above) for marine studies; taxonomic-habitat groupings were the same as the BT model described above. This model allows to us to

characterize variation among realm-region-taxon groups, as well as among studies within these groups.

Finally, our simplest model also nested cells into studies. These were subsequently nested into a concatenation of realm, latitude and taxon, and we refer to this model as the realm-latitude-taxon (RLT) model. Realm was one of marine, terrestrial or freshwater; latitude was simplified to three latitudinal bands (polar: $|\text{latitude}| > 60^\circ$, temperate: $23.5^\circ < |\text{latitude}| < 60^\circ$, and tropical: $|\text{latitude}| < 23.5^\circ$), and the taxonomic-habitat groupings were the same as the BT model (see above). These groupings allowed us to characterize variation in biodiversity change for taxon groups within latitudinal bands for each realm separately, as well as studies within these realm-latitude-taxon groups.

All models were fit to data that comprised at least 3 cell-level time series per biome-taxon (BT model), realm-region-taxon (RRT model) and realm-latitude-taxon group (RLT model), respectively. This threshold resulted in different numbers of studies being included in the final analyses: 239 for the BT model, 252 for the RRT model, and 271 for the RLT model. In the main text, we discuss trends at the biome, taxon and study levels for the BT model, and for the realm-latitude-taxon and study levels for the RLT model, as the analytic technique is not well suited to describing trends at the cell-level where the data are sparse. All models were specified as linear models, and fit with year as global, fixed effect slope that allows us to describe the overall trend across all of the data. Additionally, both year (i.e., the slope parameter) and the intercept varied for each of the hierarchical levels in the models, allowing us to quantify biogeographic, taxonomic and study-level variation around the overall trend. Species richness was modelled assuming a Poisson error distribution and a log link function. We discuss change in species richness using the slope coefficients, meaning (due to the link

function) that we are discussing change in the natural logarithm of species richness. The BT model had the form:

$$y_{l,k,j,i,t} \sim \text{poisson}(\mu_{l,k,j,i,t}),$$

$$5 \quad \mu_{l,k,j,i,t} = \beta_0 + \beta_{0l} + \beta_{0l,k} + \beta_{0l,k,j} + \beta_{0l,k,j,i} + (\beta_1 + \beta_{1l} + \beta_{1l,k} + \beta_{1l,k,j} + \beta_{1l,k,j,i})x_{l,k,j,i,t},$$

where $y_{l,k,j,i,t}$ is the (rarefied) species richness in year t of the i th cell in the j th study for the k th taxonomic group within the l th biome. $x_{l,k,j,i,t}$ is the time in years, β_0 and β_1 are the global intercept and slope (often termed fixed effects), β_{0l} and β_{1l} are the biome-level departures from β_0 and β_1 (respectively; biome-level random effects), $\beta_{0l,k}$ and $\beta_{1l,k}$ are taxon-level departures (i.e., taxon-level random effects, nested within biome) from β_0 and β_1 , $\beta_{0l,k,j}$ and $\beta_{1l,k,j}$ are study-level departures (study-level random effects, nested within taxon and biome) from β_0 and β_1 , $\beta_{0l,k,j,i}$ and $\beta_{1l,k,j,i}$ are the cell-level departures from β_0 and β_1 (cell-level random effects, nested within study, taxon and biome).

15 The species richness RRT model had the form:

$$y_{k,j,i,t} \sim \text{poisson}(\mu_{k,j,i,t}),$$

$$20 \quad \mu_{k,j,i,t} = \beta_0 + \beta_{0k} + \beta_{0k,j} + \beta_{0k,j,i} + (\beta_1 + \beta_{1k} + \beta_{1k,j} + \beta_{1k,j,i})x_{k,j,i,t},$$

where $y_{k,j,i,t}$ is the (rarefied) species richness in year t of the i th cell in the j th study of the k th combination of realm-region-taxon. $x_{k,j,i,t}$ is the time in years, β_0 and β_1 are the global intercept and slope (fixed effects), β_{0k} and β_{1k} are the departures for each realm-region-taxon group from β_0 and β_1 (respectively; random effects), $\beta_{0k,j}$ and $\beta_{1k,j}$ are the study-level departures from β_0 and β_1 (study-level random effects, nested within realm-region-taxon), and $\beta_{0k,j,i}$ and $\beta_{1k,j,i}$ are the cell-level departures from β_0 and β_1 (cell-level random effects, nested within studies and realm-region-taxon).

The species richness RLT model had the form:

$$y_{k,j,i,t} \sim \text{poisson}(\mu_{k,j,i,t}),$$

$$\mu_{k,j,i,t} = \beta_0 + \beta_{0k} + \beta_{0k,j} + \beta_{0k,j,i} + (\beta_1 + \beta_{1k} + \beta_{1k,j} + \beta_{1k,j,i})x_{k,j,i,t},$$

5

where $y_{k,j,i,t}$ is the (rarefied) species richness in year t of the i th cell in the j th study of the k th combination of realm-latitude-taxon. $x_{k,j,i,t}$ is the time in years, β_0 and β_1 are the global intercept and slope (fixed effects), β_{0k} and β_{1k} are the departures for each realm-latitude-taxon group from β_0 and β_1 (respectively; random effects), $\beta_{0k,j}$ and $\beta_{1k,j}$ are the study-level departures from β_0 and β_1 (study-level random effects, nested within realm-latitude-taxon), and $\beta_{0k,j,i}$ and $\beta_{1k,j,i}$ are the cell-level departures from β_0 and β_1 (cell-level random effects, nested within studies and realm-latitude-taxon).

15 We modelled the dissimilarity metrics assuming Gaussian error and an identity link function.

We made this choice as it allows us to discuss the rate of change in dissimilarity using the slope coefficients. However, we also assessed the sensitivity of our results to the assumption of Gaussian error when modelling a response that is one the [0-1] interval (see below).

Gaussian error and an identity link resulted in BT models of the form:

20

$$y_{l,k,j,i,t} \sim N(\mu_{l,k,j,i,t}, \sigma^2),$$

$$\mu_{l,k,j,i,t} = \beta_0 + \beta_{0l} + \beta_{0l,k} + \beta_{0l,k,j} + \beta_{0l,k,j,i} + (\beta_1 + \beta_{1l} + \beta_{1l,k} + \beta_{1l,k,j} + \beta_{1l,k,j,i})x_{l,k,j,i,t},$$

where $y_{l,k,j,i,t}$ is the value of the dissimilarity metric (total Jaccard dissimilarity, or one of the components) in year t of the i th cell in the j th study of the k th taxonomic group within the l th biome. $x_{l,k,j,i,t}$ is the time in years, β_0 and β_1 are the global intercept and slope, β_{0l} and β_{1l} are the biome-level departures from β_0 and β_1 (respectively), $\beta_{0l,k}$ and $\beta_{1l,k}$ are the taxon-level departures from β_0 and β_1 (taxon-level random effects, nested within biome), $\beta_{0l,k,j}$ and $\beta_{1l,k,j}$ are the study-level departures from β_0 and β_1 (study-level random effects, nested within taxon and biome),

and $\beta_{0l,k,i}$ and $\beta_{1l,k,i}$ are the cell-level departures from β_0 and β_1 (cell-level random effects, nested within study, taxon and biome). The dissimilarity metric was set to equal zero (perfectly similarity) for the first year of each time series.

5 The RRT models of dissimilarity with Gaussian error and an identity link had the form:

$$y_{k,j,i,t} \sim N(\mu_{k,j,i,t}, \sigma^2),$$

$$\mu_{k,j,i,t} = \beta_0 + \beta_{0k} + \beta_{0k,j} + \beta_{0k,j,i} + (\beta_1 + \beta_{1k} + \beta_{1k,j} + \beta_{1k,j,i})x_{l,k,j,i,t},$$

10 where $y_{j,i,t}$ is the value of the dissimilarity metric in year t of the i th cell in the j th study of the k th combination of realm-region-taxon. $x_{k,i,t}$ is the time in years, β_0 and β_1 are the global intercept and slope, β_{0k} and β_{1k} are the departures for each realm-region-taxon group from β_0 and β_1 (respectively), $\beta_{0k,j}$ and $\beta_{1k,j}$ are the study-level departures from β_0 and β_1 (study-level random effects, nested within realm-latitude-taxon), $\beta_{0k,j,i}$ and $\beta_{1k,j,i}$ are the cell-level departures
 15 from β_0 and β_1 (cell-level random effects, nested within study and realm-region-taxon). The dissimilarity metric was set to equal zero (perfectly similarity) for the first year of each time series.

The RLT models of dissimilarity with Gaussian error and an identity link had the form:

20
$$y_{k,j,i,t} \sim N(\mu_{k,j,i,t}, \sigma^2),$$

$$\mu_{k,j,i,t} = \beta_0 + \beta_{0k} + \beta_{0k,j} + \beta_{0k,j,i} + (\beta_1 + \beta_{1k} + \beta_{1k,j} + \beta_{1k,j,i})x_{l,k,j,i,t},$$

where $y_{j,i,t}$ is the value of the dissimilarity metric in year t of the i th cell in the j th study of the
 25 k th combination of realm-latitude-taxon. $x_{k,i,t}$ is the time in years, β_0 and β_1 are the global intercept and slope, β_{0k} and β_{1k} are the departures for each realm-latitude-taxon group from β_0 and β_1 (respectively), $\beta_{0k,j}$ and $\beta_{1k,j}$ are the study-level departures from β_0 and β_1 (study-level random effects, nested within realm-latitude-taxon), $\beta_{0k,j,i}$ and $\beta_{1k,j,i}$ are the cell-level departures

from β_0 and β_i (cell-level random effects, nested within study and realm-latitude-taxon). The dissimilarity metric was set to equal zero (perfectly similarity) for the first year of each time series.

- 5 We used weakly regularizing normally-distributed priors for the global intercept and slope. For all models of composition (the turnover and nestedness components of Jaccard's dissimilarity) and the species richness RLT model they were specified as:

$$\beta_0 \sim N(0,2),$$

$$\beta_1 \sim N(0,1),$$

10

the global intercept and slope of the BT and RRT species richness models were specified as:

$$\beta_0 \sim N(0,1),$$

$$\beta_1 \sim N(0,0.2).$$

15

Group-level parameters were all assumed to be $N(0, \sigma)$, and priors on the σ were the same for all models of composition (i.e., as follows, with the l subscript grouping dropped for the RRT and RLT models):

20

$$\sigma_{\beta_{0l}} = \sigma_{\beta_{0l,k}} = \sigma_{\beta_{0l,k,j}} = \sigma_{\beta_{0l,k,j,i}} \sim \text{half Cauchy}(0, 2),$$

$$\sigma_{\beta_{1l}} = \sigma_{\beta_{1l,k}} = \sigma_{\beta_{1l,k,j}} = \sigma_{\beta_{1l,k,j,i}} \sim \text{half Cauchy}(0, 2).$$

The group-level parameters of the BT and the RLT of model species richness were also assumed to be $N(0, \sigma)$, but the priors for σ in the BT and RRT models were drawn from the

25

exponential distribution:

$$\sigma_{\beta_{0l}} = \sigma_{\beta_{0l,k}} = \sigma_{\beta_{0l,k,j}} = \sigma_{\beta_{0l,k,j,i}} \sim \text{exponential}(1),$$

$$\sigma_{\beta_{1l}} = \sigma_{\beta_{1l,k}} = \sigma_{\beta_{1l,k,j}} = \sigma_{\beta_{1l,k,j,i}} \sim \text{exponential}(1).$$

and the priors for the RLT model of species richness were drawn from the student-t

30

distribution:

$$\sigma_{\beta_{0k}} = \sigma_{\beta_{0k,j}} = \sigma_{\beta_{0k,j,i}} \sim \text{student } t(3, 0, 10),$$

$$\sigma_{\beta_{1k}} = \sigma_{\beta_{1k,j}} = \sigma_{\beta_{1k,j,i}} \sim \text{student } t(3, 0, 10).$$

5 Correlations between levels of the grouping-factors (e.g., taxa with biomes) are estimated using the Cholesky decomposition (L) of the correlation matrix, with a Lewandowski-Dorota-Joe (LKJ) prior (63), here set as:

$$L \sim \text{LKJ}(2).$$

10

Model convergence and goodness of fit were assessed using a combination of statistics (Gelman–Rubin diagnostic; (64, 65)) and visual inspection of the Markov chains.

All data manipulation and analysis were conducted in R (66). Models were coded using the

15

‘brms’ package (version 1.5.1 or greater; 63), which fits models with the probabilistic programming language Stan (67).

Robustness checks and sensitivity analyses

To examine whether overall trends and the geographic and taxonomic patterns we found were

20

sensitive to the differences in the grouping structures of the three model, we examined estimates from comparable levels of the models. The estimates of biodiversity change from the models were remarkably consistent in terms of the overall trends they estimated, and at the study- (Fig. S3A, B) and cell-levels (Fig. S3C, D). This suggests that our overall trend estimates were largely robust to the specifics of the model. We did, however, find some

25

differences between the taxon estimates from BT model when they were compared to the taxon estimates from the RRT model for turnover, though not for changes in species richness (Fig. S4-S6). When terrestrial and freshwater taxon were grouped into regions instead of

biomes, we found more variation in the departures around the overall turnover trend in the RRT model when compared with BT model. This meant that the differences between the realms that we focus on in the main text, whereby rates of turnover were faster and more variable in marine realm, were exaggerated in the RRT model (Fig. S4, Fig. S5), though the rank in terms of turnover rates were similar between the two models (Spearman's $\rho = 0.6$).

A recurrent criticism of existing analyses of biodiversity time series is the lack of an appropriate baseline from which to detect change (12, 68). Although baselines can be defined when assessing the effect of a particular event, it is not clear how the concept applies more broadly to detect and document change. Therefore, obtaining baselines for all of the datasets in the BioTIME database was unrealistic, but we assessed whether the rates of change are themselves changing through time by quantifying biodiversity change for different time periods (since the 1950's). To do this, we subset the data into three periods: 1951-1970, 1971-1990, and 1990-2010, and refit each of the models to each of these subsets (Fig. S17-S18 show results for richness and the turnover component of dissimilarity). We did not find large differences in the global slope among the different time periods. The 90% credible interval (CI) overlapped zero for the period 1950-1970 (global slope = 0.02; 90% CI = -0.003 - 0.05) and 1971-1990 (global slope = -0.001; 90% CI = -0.016 - 0.15); the period 1991-2010 had a weakly positive slope (global slope = 0.007; 90% CI = 0.0005 - 0.01), though the 95% CI did overlap zero (95% CI: -0.001 - 0.01)

We assessed how our estimates of biodiversity change varied with the length of the time series, the number of discrete years sampled, and the starting year of each time series (Fig. S7-S8). We also examined the estimates of change as a function of the initial assemblage species richness (i.e., the number of species in the first year of each assemblage time series; Fig. S7-S8). We did not detect systematic effects of any of these variables on our estimate of rates of change of species richness, turnover or nestedness.

We used simulations to examine whether our finding of directional trends in compositional dissimilarity could be due to repeated random sampling from a regional species pool. Each simulation ($n_{sim} = 1000$) consisted of a time series randomly drawn from a regional species pool; time series duration was approximately matched to the distribution of duration in the empirical data (Fig. S9 inset), and was a random draw from a log-normal distribution with $\text{mean}(\log(\text{duration})) = 2.3$, and $\text{sd}(\log(\text{duration})) = 0.65$ that we rounded down to an integer value; the size of the regional species pool was a uniformly distributed integer value between 200 and 1000; sampling effort (i.e., the number of samples at each time point) was a uniformly distributed integer value between 2 and 50. For each simulated time series we calculated the turnover and nestedness components of Jaccard's dissimilarity between each time point and the initial assemblage, and estimated the rate of change in turnover and nestedness as the slope coefficient of a linear model that assumed Gaussian error and an identity link, where either turnover or nestedness was modelled as a function of time. Median values for both turnover and nestedness change were zero, with the 95th percentile equal to zero for nestedness change and 0.01 for turnover (Fig. S10), suggesting our finding of turnover rates greater than zero are very unlikely to have arisen simply from random sampling through time of a constant regional species pool.

We further examined the robustness of the results of our compositional change analyses in two additional ways. To examine whether our results were sensitive to quantifying dissimilarity between the initial assemblage and each subsequent year, we refit the BT models with Gaussian error to dissimilarities (i.e., the turnover and nestedness components of Jaccard dissimilarity) calculated between assemblages at consecutive time points. We also examined the robustness of assuming Gaussian error and using an identity link when modelling dissimilarity, for the BT model only, by fitting two alternative models that assumed Beta

error. As dissimilarity was dominated by the turnover component, we present these sensitivity analyses for the turnover component of Jaccard's dissimilarity only. Our turnover data include many zeroes and ones: all communities start from perfect similarity and many underwent complete change (i.e., no shared species with the initial community). Whilst these zeros and ones are informative for the biodiversity change estimates, and thus need to be included, the Beta distribution does not include values of zero or one. In addition to problematic dissimilarity values of zero or one, assuming Beta error also requires a link function, most commonly the logit-link. The logit-link function means that we can no longer interpret the slope coefficients as a rate of change. In a linear model with a logit-link, intercept α and slope β and a single covariate x , the derivative with respect to x is $\beta e^{(\alpha + \beta x)} / [(1 + e^{(\alpha + \beta x)})^2]$; the rate of change is no longer simply the β coefficient as it is when we assume Gaussian error and an identity link. With the logit-link, the rate of change additionally depends on the intercept, and more critically, the value of the covariate x . We took two approaches to dealing with the first problem (i.e., the zeros and ones) with two different models that assumed Beta error and a logit-link. And as the rate of change in both of these new models is no longer independent of year (our x covariate), we simply examined whether our slope coefficients were qualitatively consistent between models that assumed Gaussian versus Beta error, rather than making comparisons of the rate of change in dissimilarity estimated by the different error distributions and their corresponding link functions.

We first modelled turnover using a zero-one inflated beta regression with a logit-link function. The probability density function of the zero-one inflated beta distribution is:

$$betainfl(y; \alpha, \gamma, \mu, \phi) = \begin{cases} \alpha(1 - \gamma), & y = 0 \\ \alpha\gamma, & y = 1 \\ (1 - \alpha)\gamma f(y; \mu, \phi), & 0 < y < 1 \end{cases}$$

where α is the probability that an observation is a zero or one, γ is the probability that an observation is a one (given that it is a zero or a one), and μ and ϕ are the mean and precision

(1/variance) of the beta distribution, respectively. Preliminary inspection showed that a higher proportion of marine assemblages underwent complete turnover compared to terrestrial and freshwater realms, and so we modelled both α and γ assuming a Bernoulli distribution and a logit link as a function of realm. The Beta distribution part of the model (i.e., for $0 < \text{turnover} < 1$) was the same for both this model and the next and is described below.

For our second approach to modelling turnover assuming Beta error and a logit-link, we made small adjustments to all observations that were zero (+ 1e-6) and one (-1e-3). This model (and the Beta component of the zero-one inflated model described above) had the form:

$$y_{l,k,j,i,t} \sim \text{Beta}(\mu_{l,k,j,i,t}, \phi),$$

$$\text{logit}(\mu_{l,k,j,i,t}) = \beta_0 + \beta_{0l} + \beta_{0l,k} + \beta_{0l,k,j} + \beta_{0l,k,j,i} + (\beta_1 + \beta_{1l} + \beta_{1l,k} + \beta_{1l,k,j} + \beta_{1l,k,j,i})x_{l,k,j,i,t},$$

where $y_{l,k,j,i,t}$ is the value of the turnover component of Jaccard's dissimilarity in year t of the i th cell in the j th study of the k th taxonomic group within the l th biome; and μ and ϕ are the mean and precision (inverse of the variance) of the Beta distribution. $x_{l,k,j,i,t}$ is the time in years, β_0 and β_1 are the global intercept and slope, β_{0l} and β_{1l} are the biome-level departures from β_0 and β_1 (respectively), $\beta_{0l,k}$ and $\beta_{1l,k}$ are the taxon-level departures from β_0 and β_1 (taxon-level random effects, nested within biome), $\beta_{0l,k,j}$ and $\beta_{1l,k,j}$ are the study-level departures from β_0 and β_1 (study-level random effects, nested within biome and taxon), and $\beta_{0l,k,j,i}$ and $\beta_{1l,k,j,i}$ are the cell-level departures from β_0 and β_1 (cell-level random effects, nested within biome, taxon and study).

We found that the conditional one inflation (i.e., the probability of perfect dissimilarity, given an observation was either a zero or a one) differed between realms, and was highest in the marine realm (Fig. S13). This pattern contributed to our main finding of higher rates of turnover in marine assemblages. Additionally, we found that our second model assuming Beta error (where the slope estimates were allowed to be influenced by values very close to zero and one) showed qualitatively similar results to those we present in the main text (Fig. S14;

Spearman's rank correlations > 0.9 for the slope estimates at all levels of the model: cell, study, taxon and biome). Finally, our results for change in the turnover and nestedness components of Jaccard's dissimilarity were qualitatively similar regardless of whether dissimilarities were calculated between the initial assemblage and each subsequent year (as presented in the main text), or calculated between assemblages in consecutive years (Fig. S15, S16).

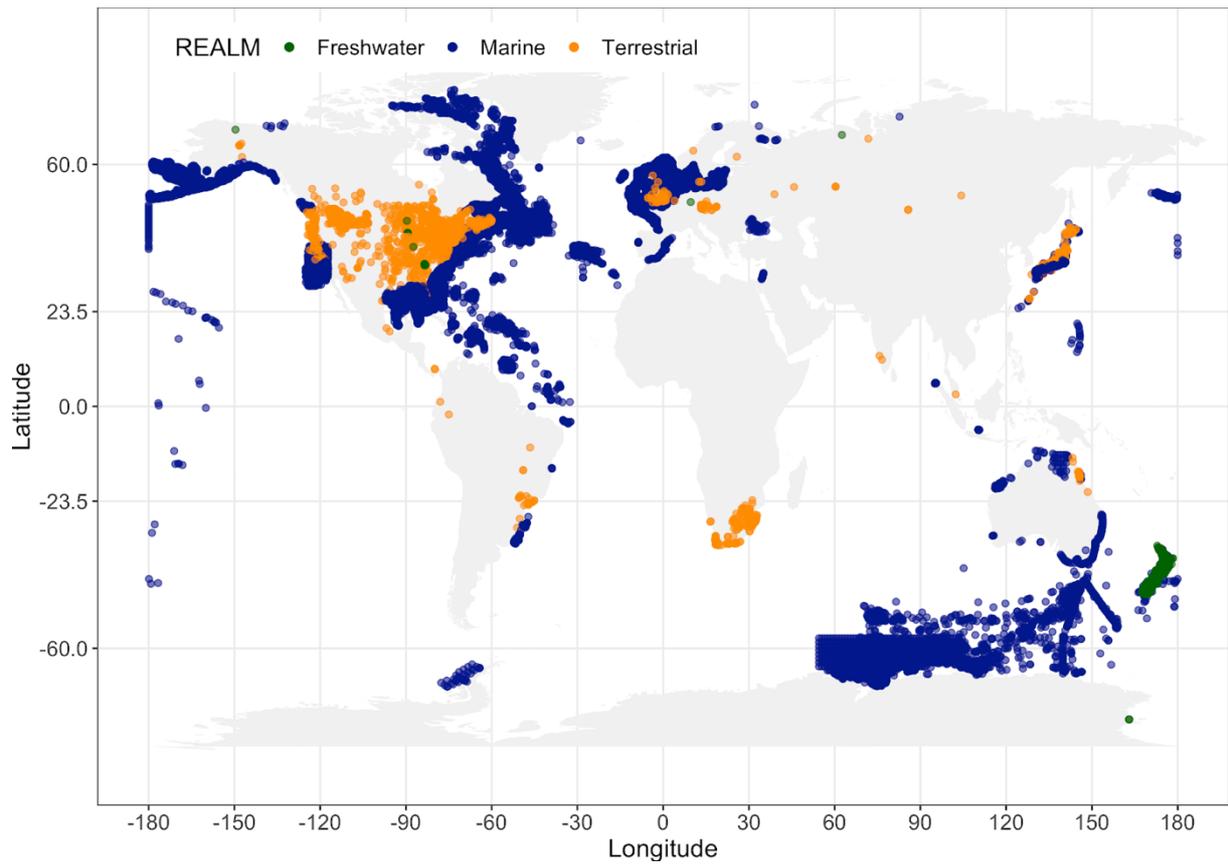


Figure S1.

The geographic distribution of the cell-level time series to which the hierarchical models were fit. Color represents realm as freshwater (green), marine (blue), and terrestrial (orange).

5

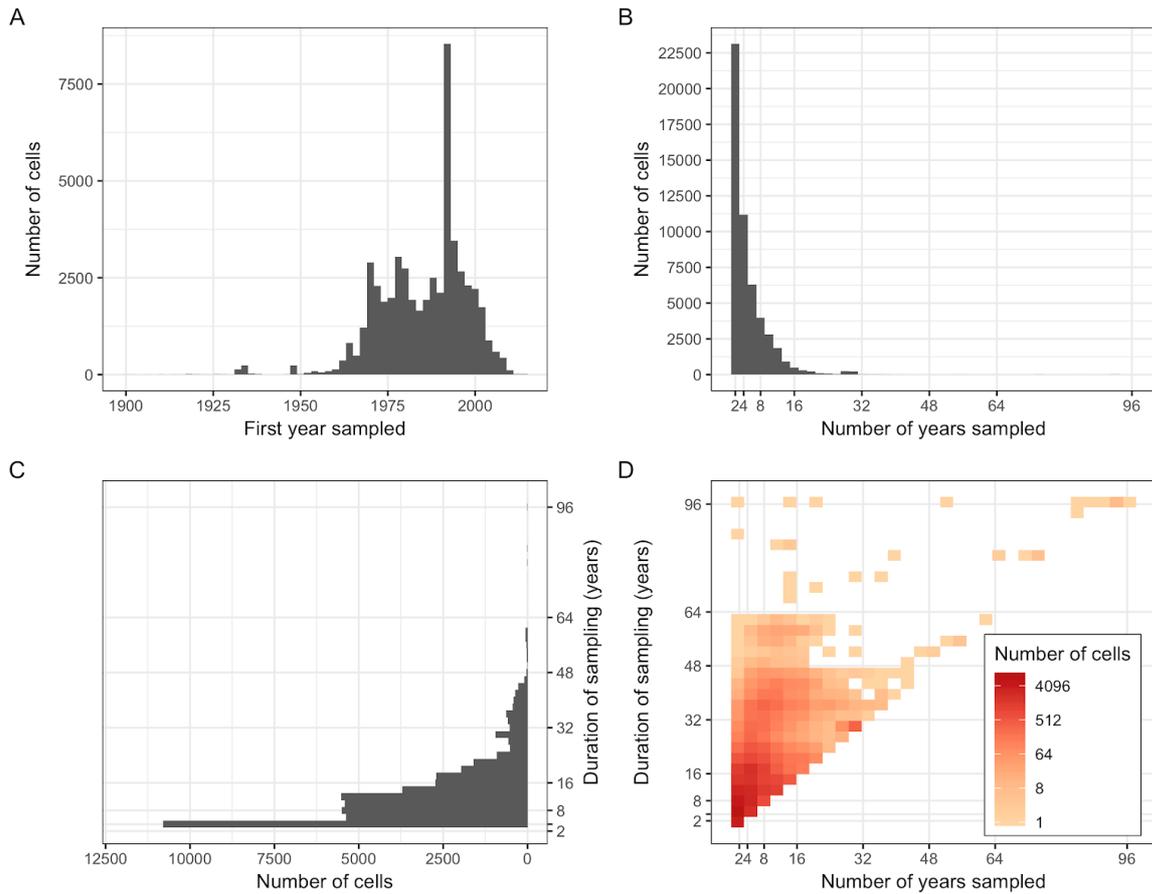


Figure S2.

Variation among the cell-level time series. **A**, Histogram of the starting year of the time series (binned into 2-year periods). **B**, Histogram of the number of years sampled in each time series.

- 5 **C**, Histogram of the duration (time period between first and last sample) of the time series. **D**, Duration of time series as a function of the number of samples, cells are coloured by the density (number of cells) for given combination of duration and number of samples. Note that panel C was rotated to be read in conjunction with B and D: specifically, the y axes of C and D and the x axes of B and D have the same orientation to show the number of cells (D)
- 10 for a given combination of the number of samples (B) and the sampling duration (C).

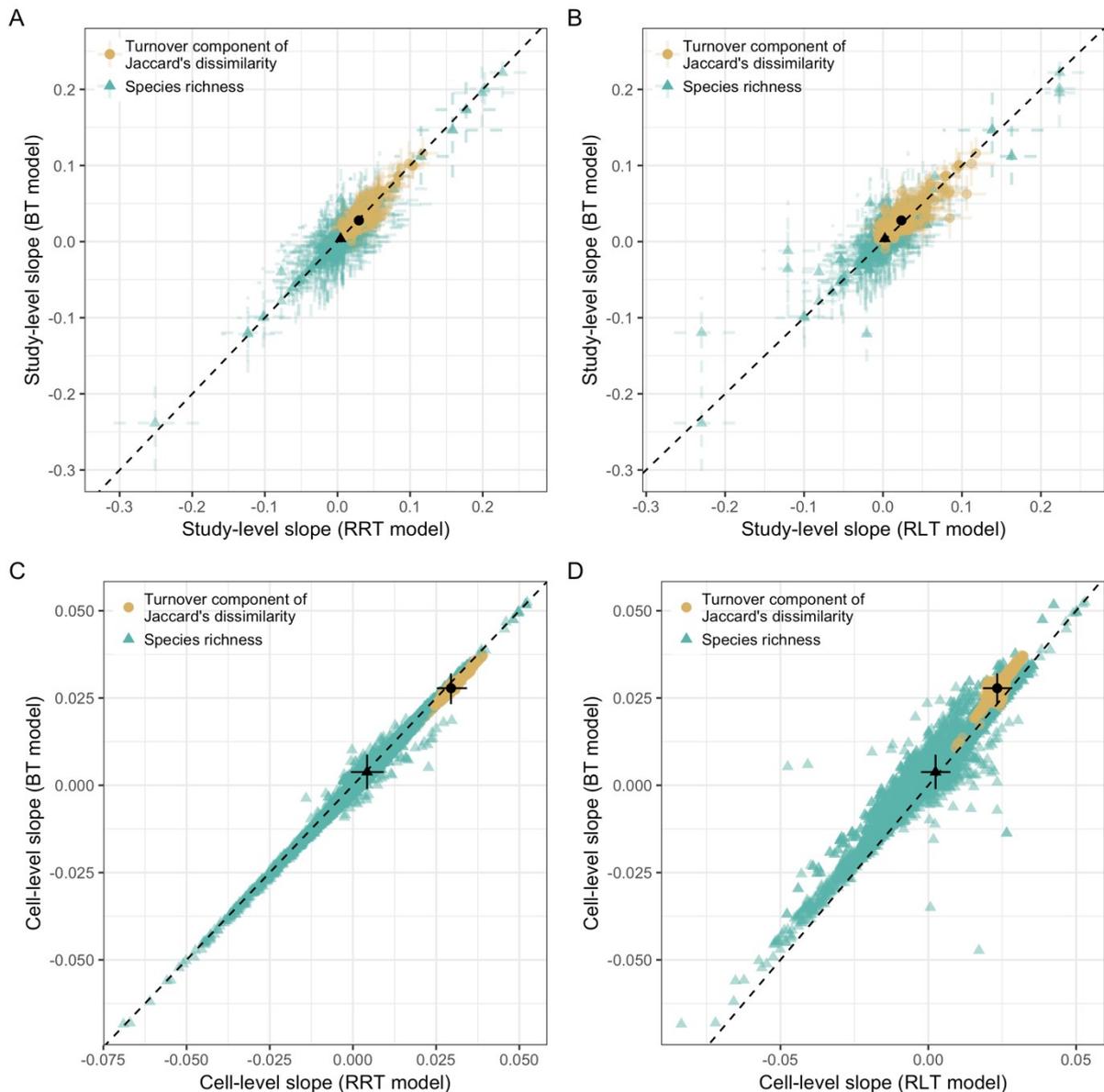


Figure S3.

Comparison of study- and cell-level slopes between the realm-region-taxon (RRT) and realm-latitude-taxon (RLT) with biome-taxon (BT) models. Comparison of study-level estimates for the (A) RRT and BT models, and (B) RLT and BT models; comparisons of the cell-level estimates for the (C) RRT and BT models, and (D) RLT and BT models. Dashed line is the 1:1 line. All point estimates are medians of posterior distributions; uncertainty on A and B are 90% credible intervals.

- 5
- 10 At the study-level, species richness and turnover slopes show slightly stronger concordance between the RRT and BT models, when compared to that between the RLT and BT models. This is expected due to the closer similarity of the model structure (e.g., marine biomes appear as a grouping factor in both models). Similarly, at the cell-level, both species richness and turnover show stronger concordance between the RRT and BT model, when compared to
- 15 that between the RLT and BT models. For the RLT – BT comparison (Fig. S3D), the biome-taxon model has slightly larger slope estimates below zero, but they are closer to the realm-latitude-taxon estimates above zero. Turnover slopes were also qualitatively consistent between the RLT and BT models at the cell-level level, though the biome-taxon model consistently estimated the change to be higher than the realm-latitude taxon model.

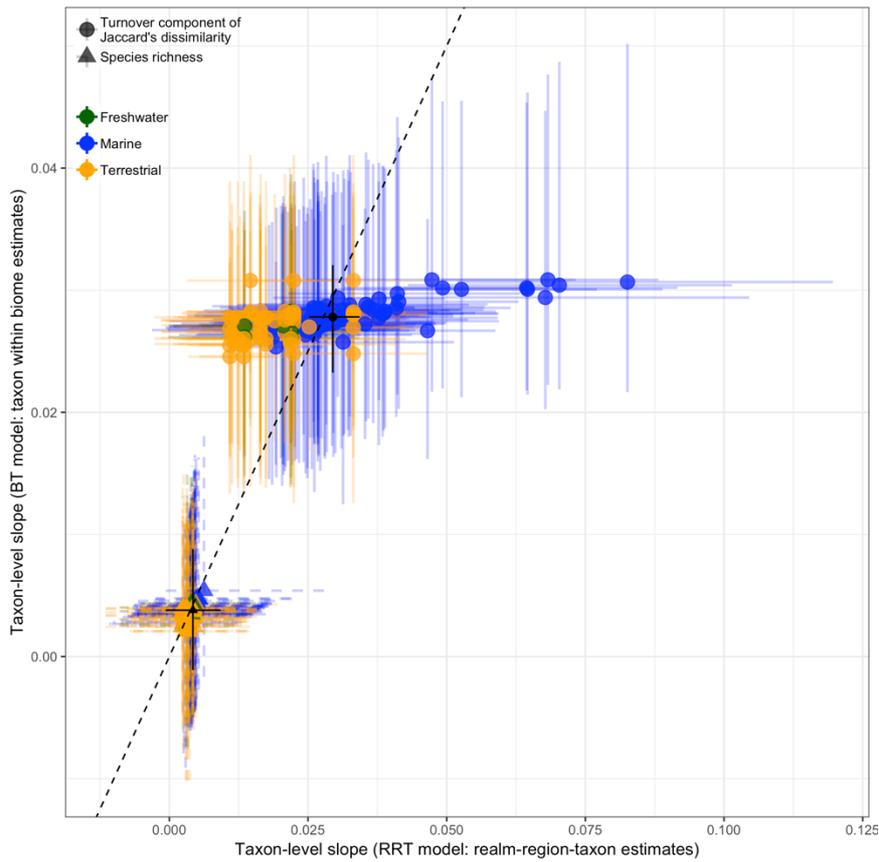


Figure S4

Comparison of taxon-level estimates from the biome-taxon and realm-region-taxon models. Black points and lines show the overall estimates and the 90% credible intervals. Coloured points show taxon level estimates and 90% credible intervals for each realm. Dashed line is 1:1 line. All point estimates are medians of posterior distributions.

5

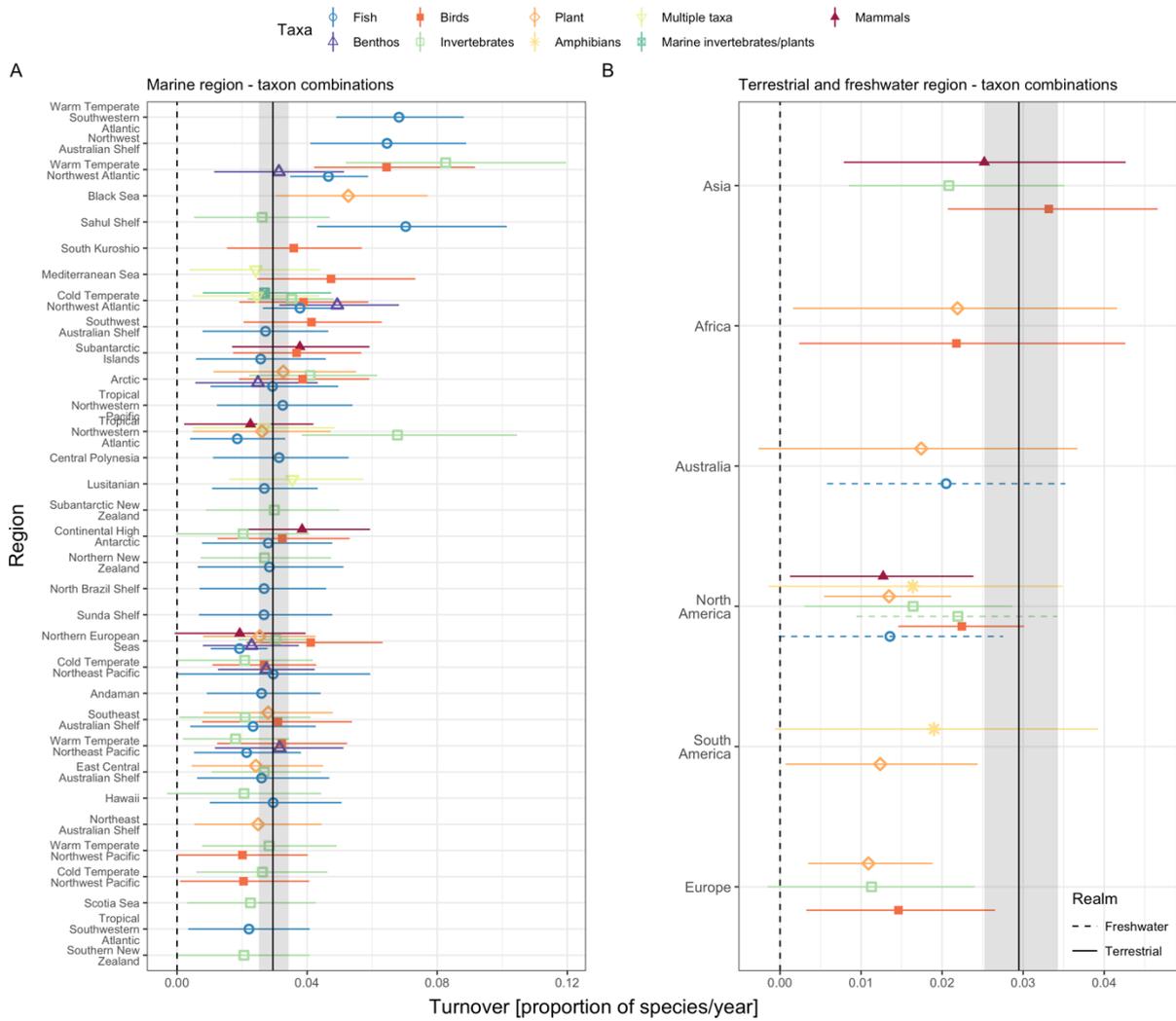


Figure S5

5 Realm-region-taxon slope estimates for turnover in **A** marine and **B** terrestrial and freshwater regions. Each point represents the median of the posterior distribution for a combination of biome-taxon-study and the lines represent the 90% credible interval; point shape and color represent taxon groups; line type differentiates terrestrial and freshwater estimates on **B**. Regions have been ordered by the magnitude of the slope estimates. Solid black vertical line and shading on both panels depict the overall trend (median of the posterior distribution) and 90% credible interval.

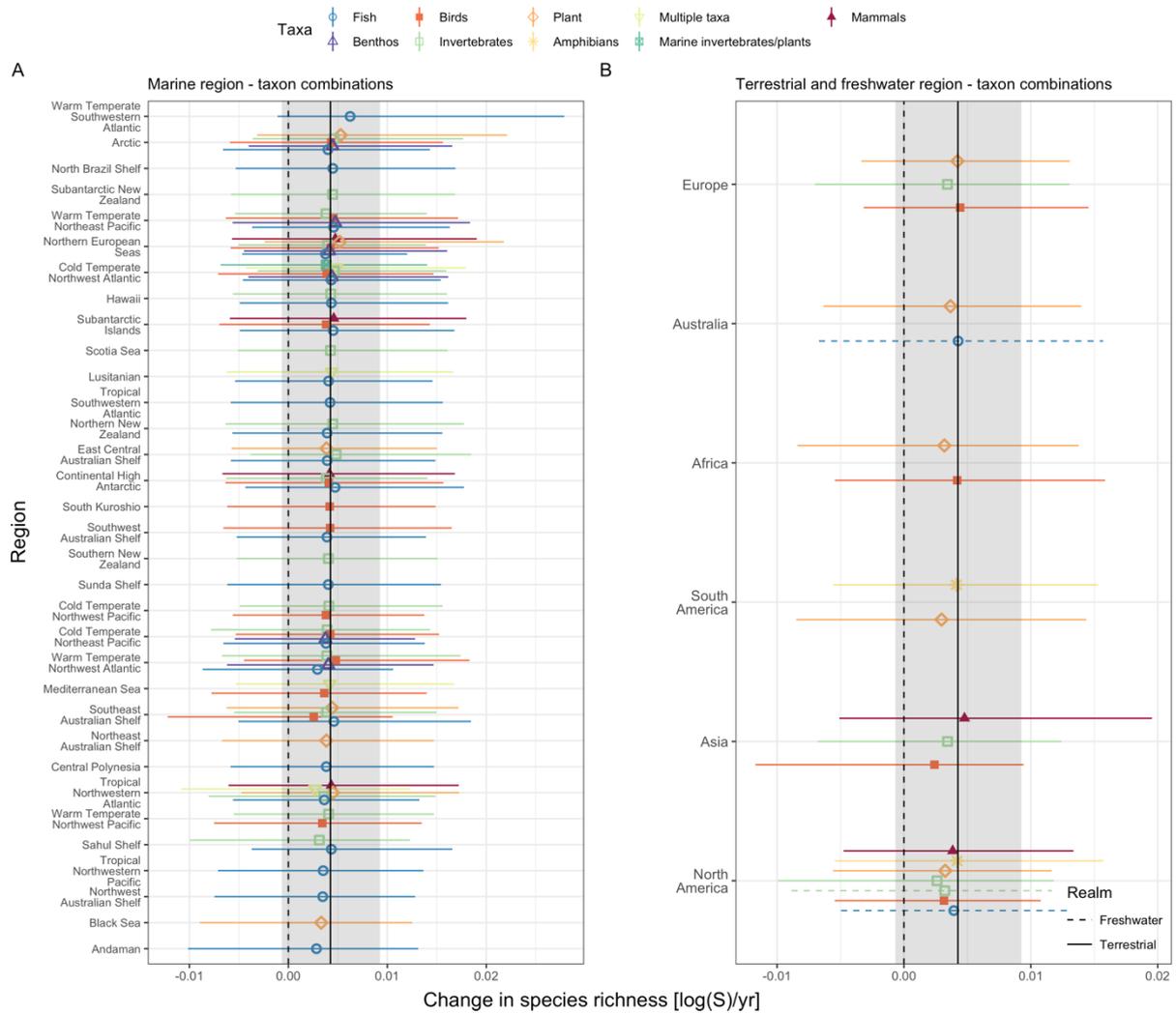


Figure S6

5 Realm-region-taxon slope estimates for species richness in **A** marine and **B** terrestrial and freshwater regions. Each point represents the median of the posterior distribution for a combination of biome-taxon-study and the lines represent the 90% credible interval; point shape and color represent taxon groups; line type differentiates terrestrial and freshwater estimates on **B**. Regions have been ordered by the magnitude of the slope estimates. Solid black vertical line and shading on both panels depict the overall trend (median of posterior distribution) and 90% credible interval.

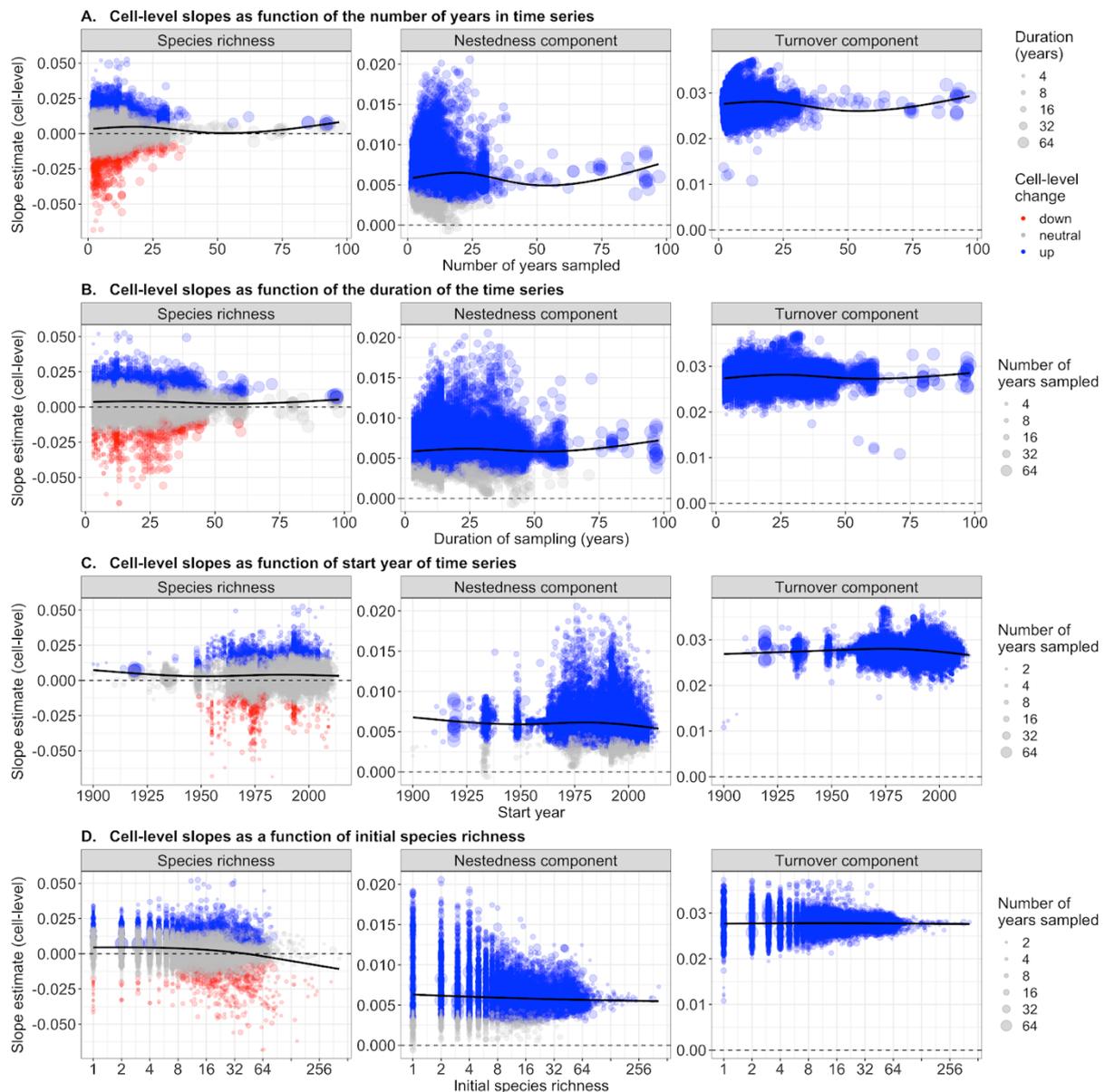


Figure S7.

Sensitivity analyses for the biome-taxon models. We use GAMs for maximum flexibility to detect systematic trends in the data, but no concerning pattern is clear. **A.** Cell-level slope estimates as a function of the number of years (discrete observations) in the time series to which models were fit, size of the points represents the duration of the time series (Duration = $\text{Year}_{\text{end}} - \text{Year}_{\text{start}} + 1$); **B.** cell-level slope estimates as a function of the duration of the time series to which models were fit, size of the points represents the number of time points in the time series; **C.** Cell-level slope estimates as a function of the start year (initial year) in the time series to which models were fit, size of the points represents the number of years sampled in the time series; **D.** Cell-level slope estimates as a function of the initial species richness (observed in start year) in the time series to which models were fit. Colours on all panels show whether the 90% credible interval (or the cell-level slope) overlapped zero (grey = neutral) or not (blue = up, red = down); black line on each panel is a GAM estimated with a penalised cubic regression spline with four knots to prevent overfitting.

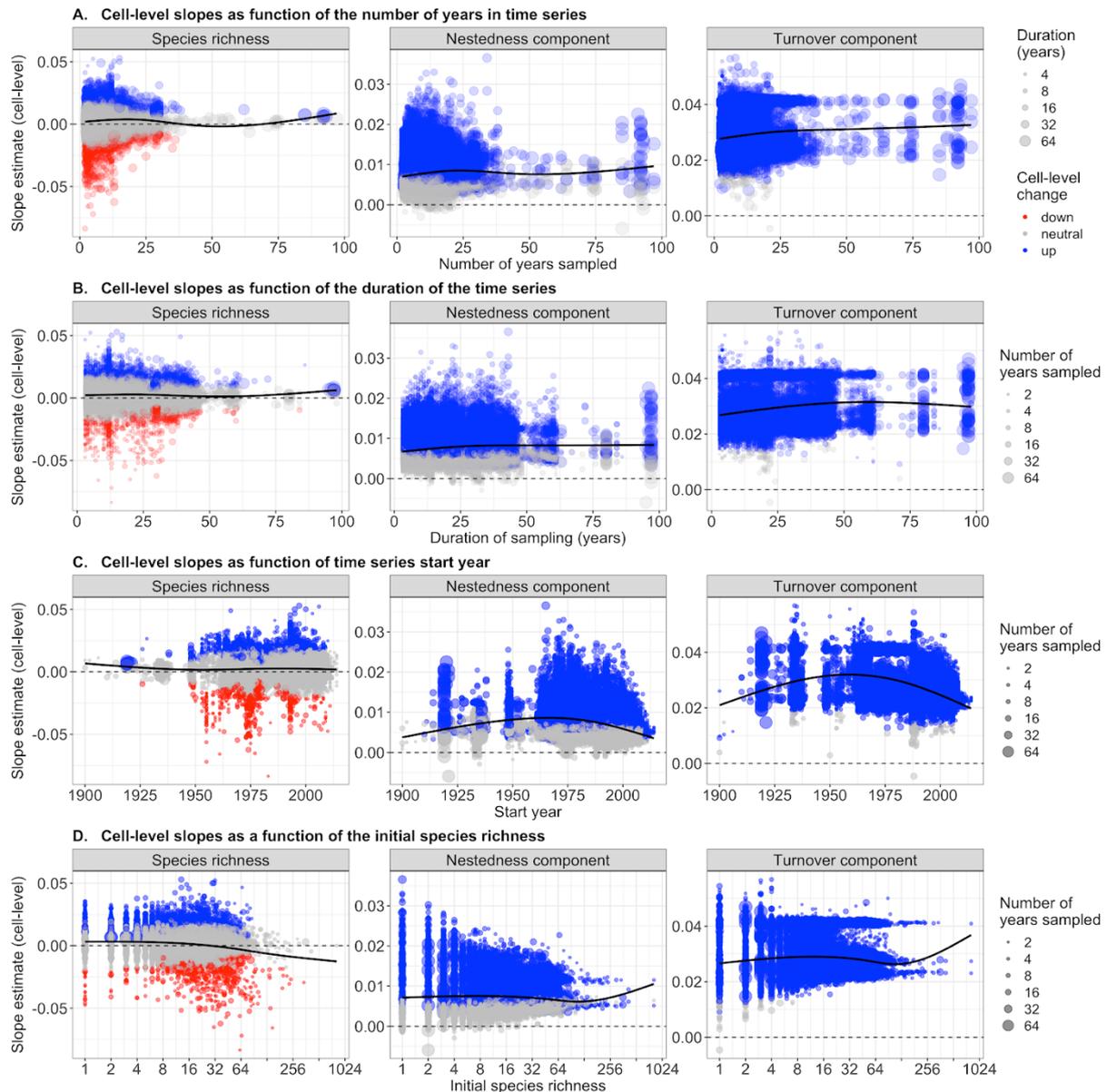


Figure S8.

Sensitivity analyses for the realm-latitude-taxon models. We use GAM fits for maximum flexibility to detect systematic trends in the data, but no concerning pattern is clear. **A.** Cell-level slope estimates as a function of the number of years (discrete observations) in the time series to which models were fit, size of the points represents the duration of the time series ($\text{Duration} = \text{Year}_{\text{end}} - \text{Year}_{\text{start}} + 1$); **B.** Cell-level slope estimates as a function of the duration of the time series, size of points represents the number of years sampled; **C.** Cell-level slope estimates as a function of the start year (initial year) in the time series to which models were fit, size of the points represents the number of years sampled in the time series; **D.** Cell-level slope estimates as a function of the initial species richness (observed in start year) in the time series to which models were fit. Colours on all panels show whether the 90% credible interval (or the cell-level slope) overlapped zero (grey = neutral) or not (blue = up, red = down); black line on each panel is a GAM estimated with a penalised cubic regression spline with four knots to prevent overfitting.

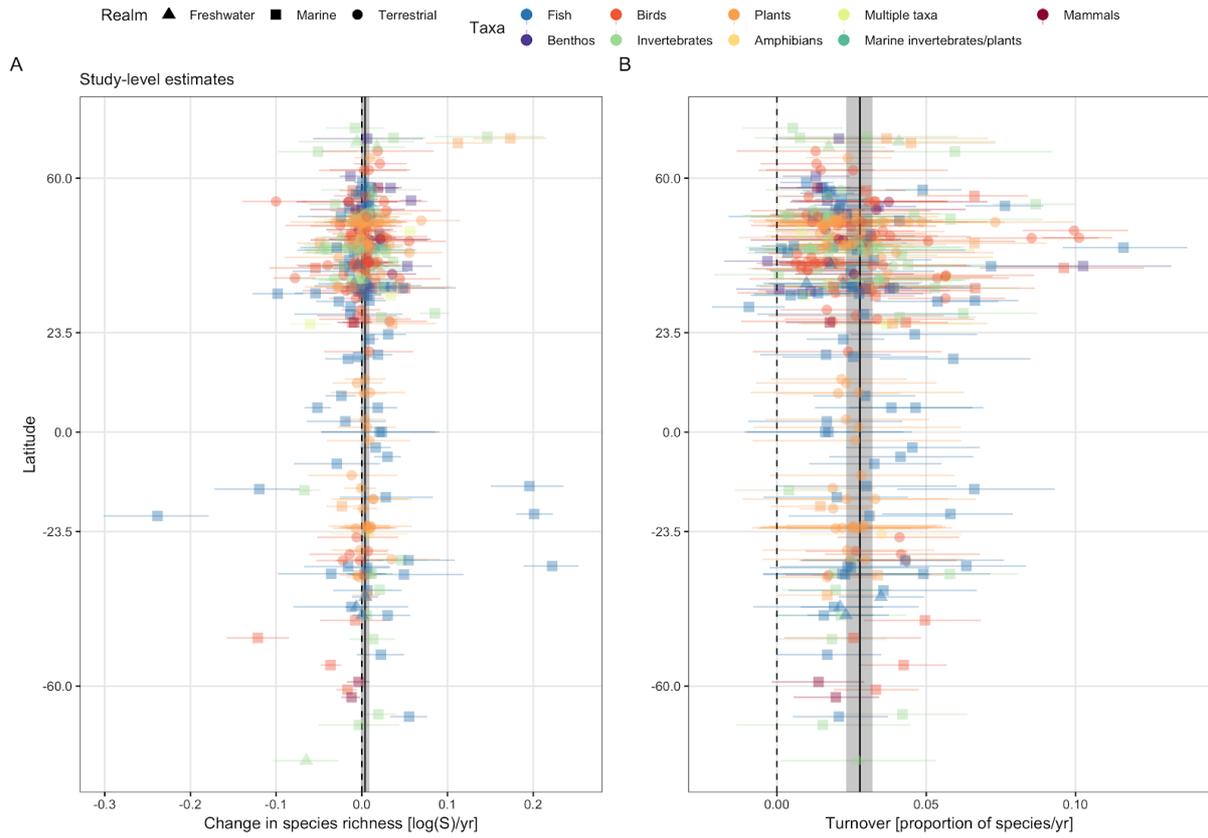


Figure S9.

Study-level slope estimates of the biome-taxon model for **A**, species richness and, **B**, turnover. Each point represents the median of the posterior distribution of a combination of biome-taxon-study and the lines represents the 90% credible interval; point shape represents realm, and color represents taxon.

5

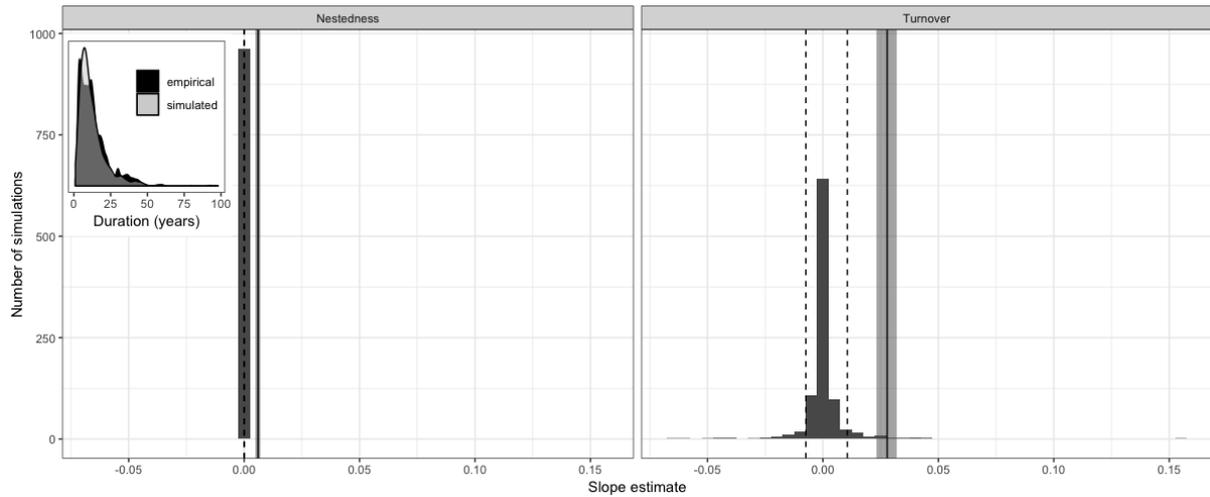


Figure S10.

5 Distributions of the estimated rates of change in nestedness and turnover when time series are random samples from regional pools of a constant size ($n_{\text{sim}} = 1000$). Solid vertical line and shading show the estimates from the biome-taxon models for nestedness and turnover and the 90% credible interval. Dashed vertical lines show the 90% quantiles of the simulated rates of change (NB: these are both equal to zero on the nestedness panel and not visible). Inset shows the probability densities of the duration of time series in the empirical and simulated data.

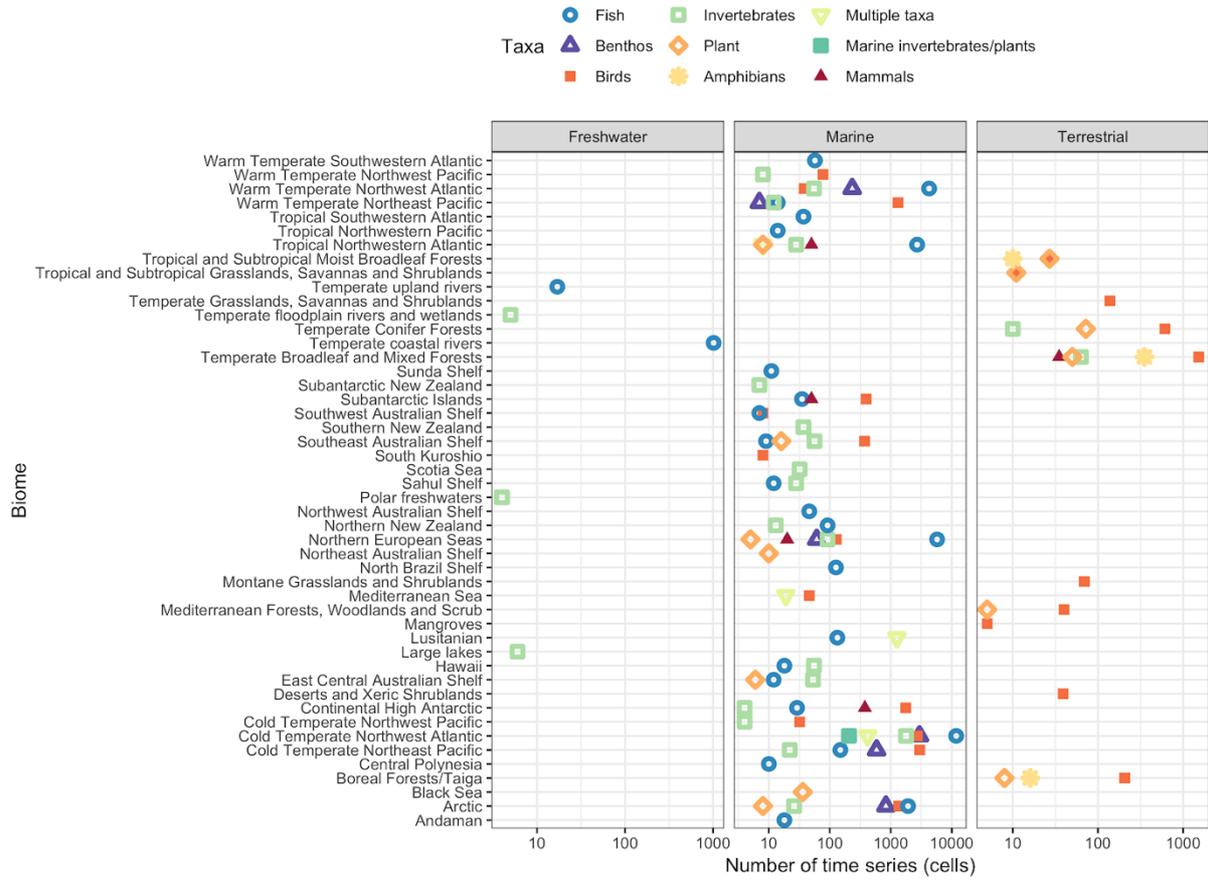
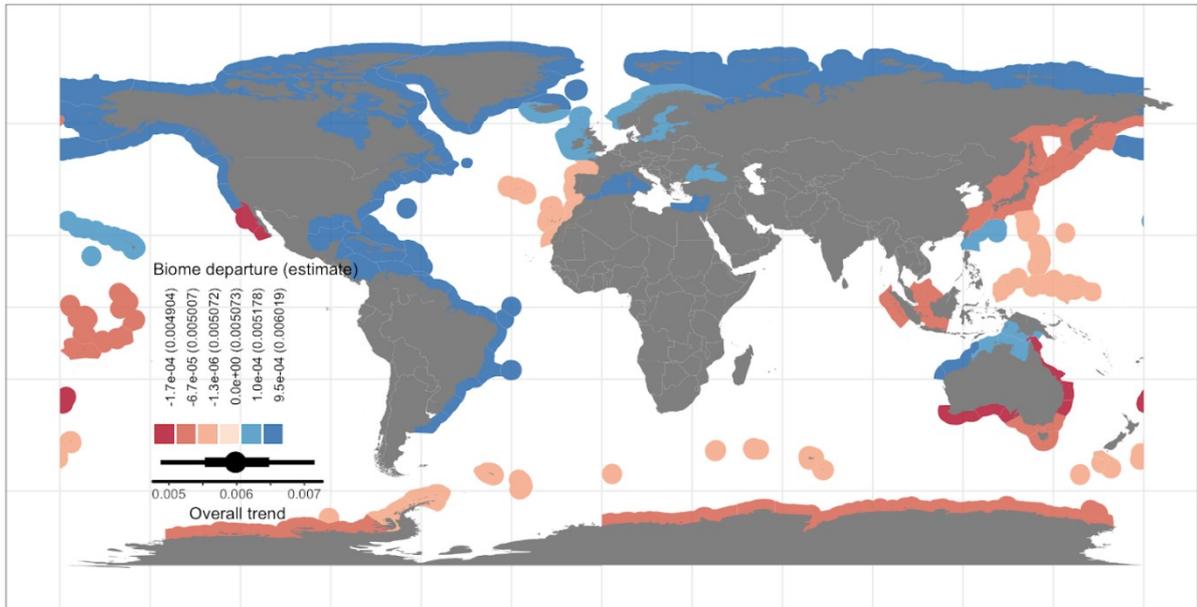


Figure S11.

The number of cell-level time series for each biome-taxon combination in the biome-taxon model; color and shape of each point represents the taxon group sampled within a given biome.

5

A Marine biomes



B Terrestrial and freshwater biomes

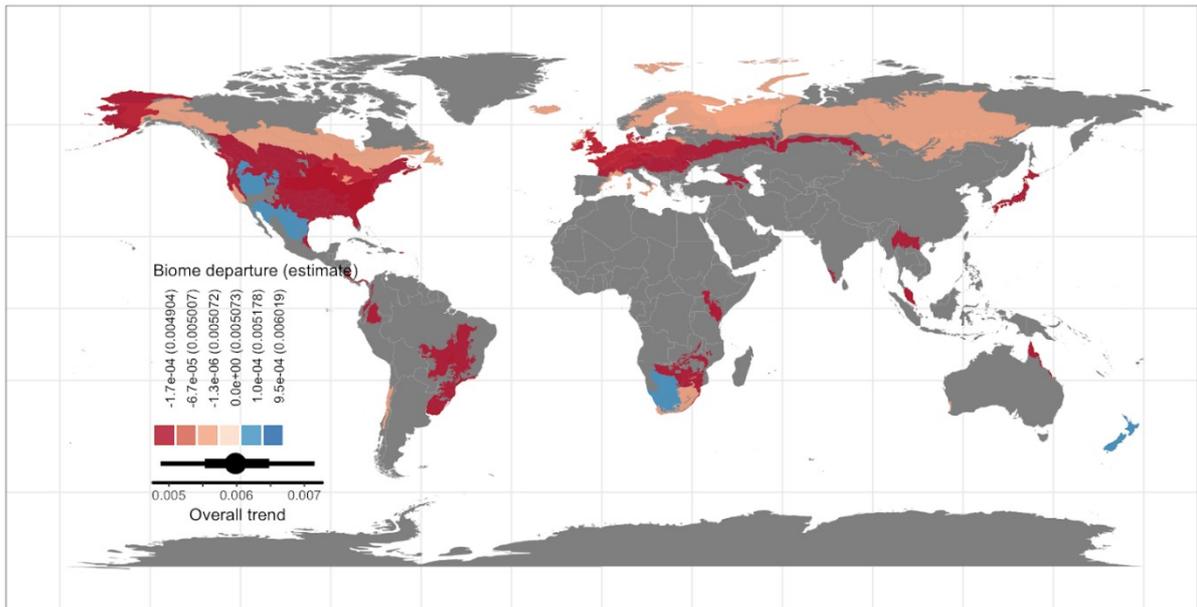


Figure S12.

Change in the nestedness component of Jaccard's dissimilarity as estimated by the biome-

taxon model. Inset shows the overall trend (median of posterior distribution), and the

5

depicts 50% (thick) and 90% (thin) credible intervals. **A**, Departures from the overall trend in

marine biomes are mixed, but more than half of marine biomes exceed the overall trend (blue

shading). **B**, Terrestrial and freshwater biomes. have both positive and negative departures

from the overall trend, but are strongly skewed towards negative departures.

10

15

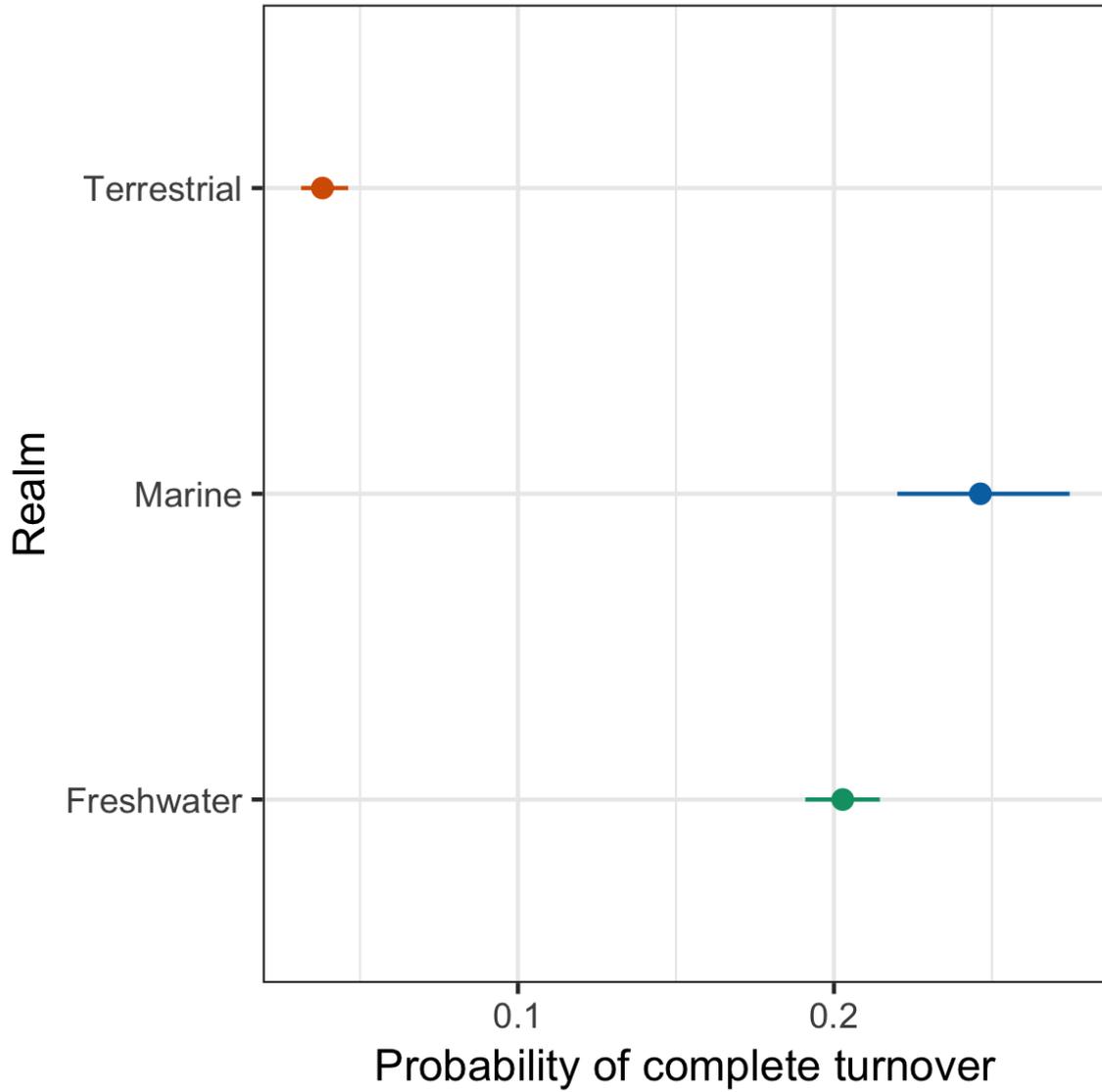


Figure S13.

The conditional probability (i.e., the probability of turnover being equal to one, given it was equal to zero or one) of complete turnover was highest in the marine realm. Points show the median of the posterior distribution and the lines represent the 90% credible interval.

5

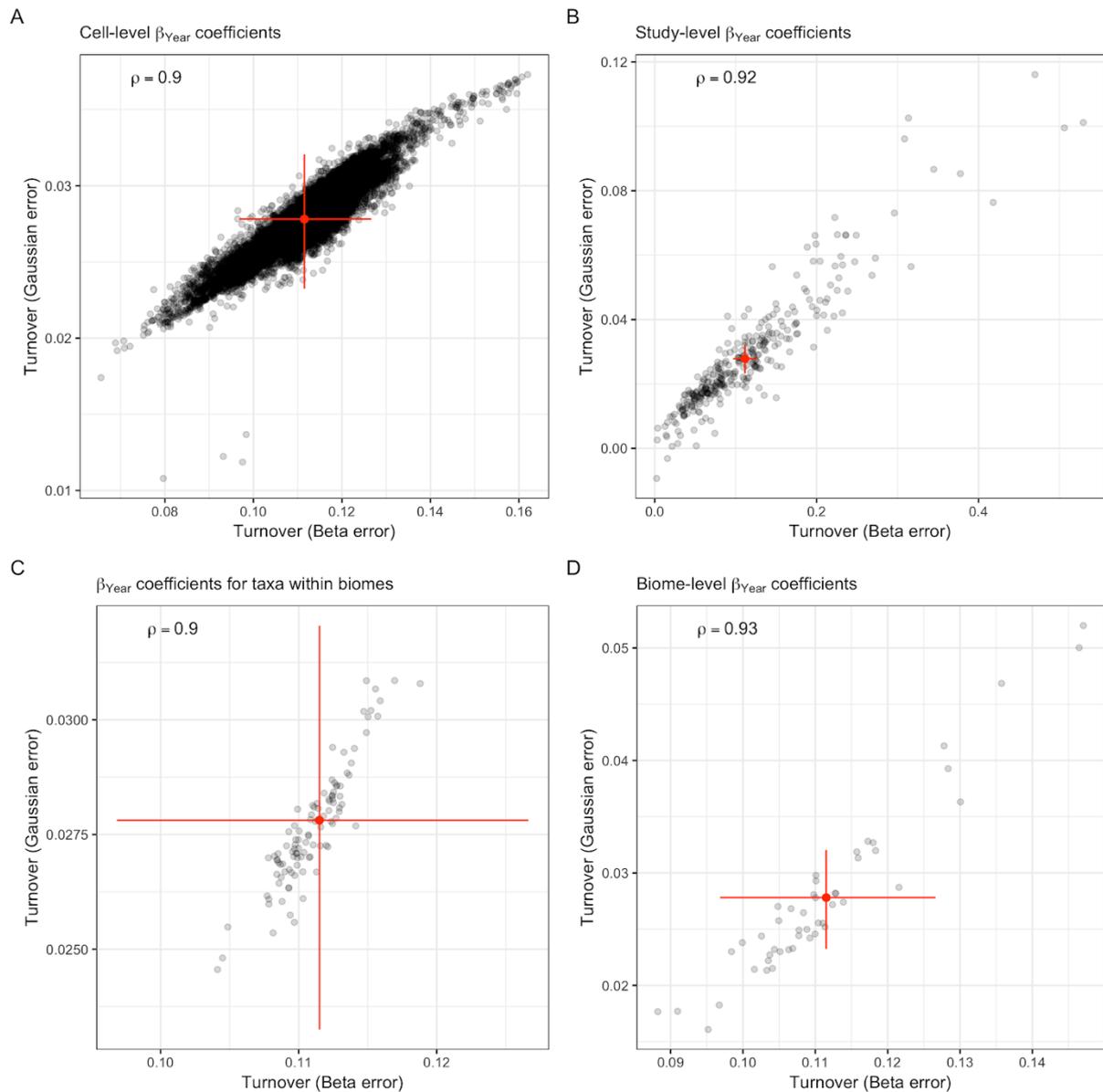


Figure S14.

Slope coefficients for the turnover component of Jaccard's dissimilarity. Each panel show the estimate from the biome-taxon model estimated with Gaussian error as a function of the model estimated with Beta error, where zeros (+1e-6) and ones (-1e-3) were shifted slightly so as they were retained in the regression. **A**, cell-level, **B**, study-level, **C**, taxon-level, and **D**, biome-level slope estimates. Red point and whiskers show the global estimates and 90% credible interval for both models; ρ on each panel reports Spearman's rank correlation.

5

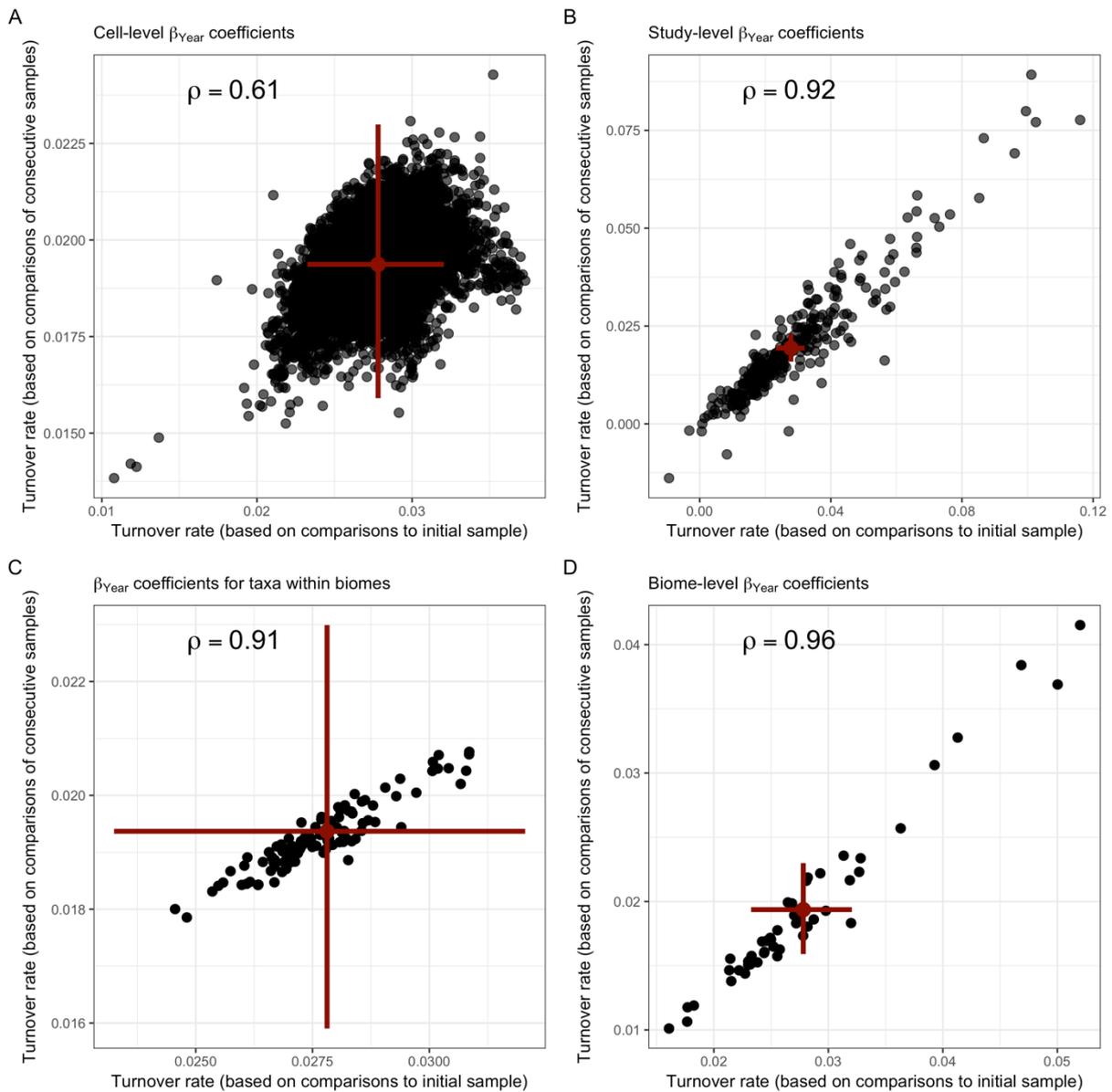


Figure S15.

Slope coefficients for the turnover component of Jaccard's dissimilarity. Each panel show the estimates from the biome-taxon model estimated with Gaussian error for turnover calculated between consecutive assemblages as a function of turnover compared to the initial assemblage. **A**, cell-level, **B**, study-level, **C**, taxon-level, and **D**, biome-level slope estimates. Red point and whiskers show the global estimates and 90% credible interval for both models; ρ on each panel reports Spearman's rank correlation.

5

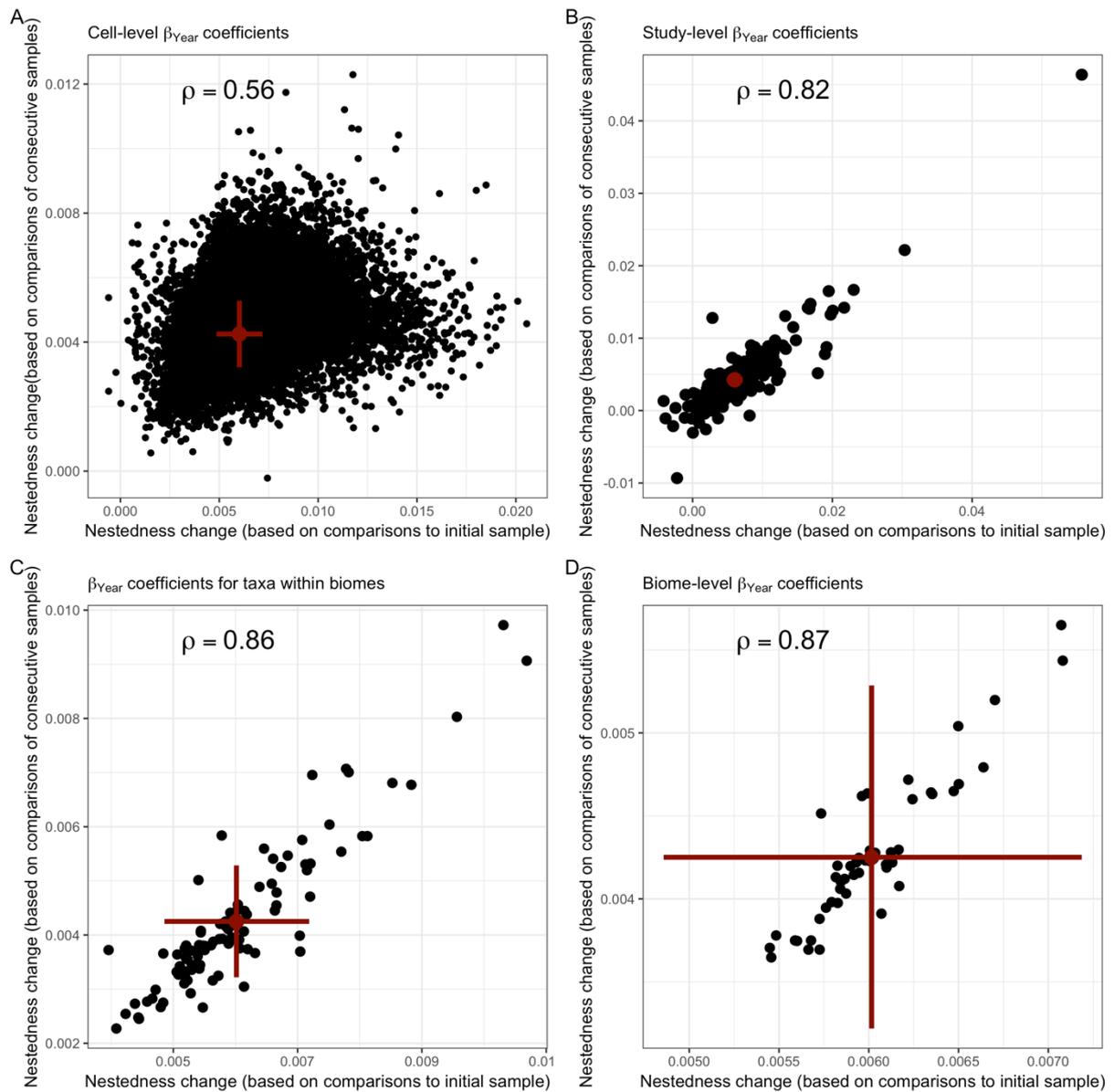


Figure S16.

Slope coefficients for the nestedness component of Jaccard's dissimilarity. Each panel show the estimates from the biome-taxon model estimated with Gaussian error for turnover calculated between consecutive assemblages as a function of turnover compared to the initial assemblage. **A**, cell-level, **B**, study-level, **C**, taxon-level, and **D**, biome-level slope estimates. Red point and whiskers show the global estimates and 90% credible interval for both models; ρ on each panel reports Spearman's rank correlation.

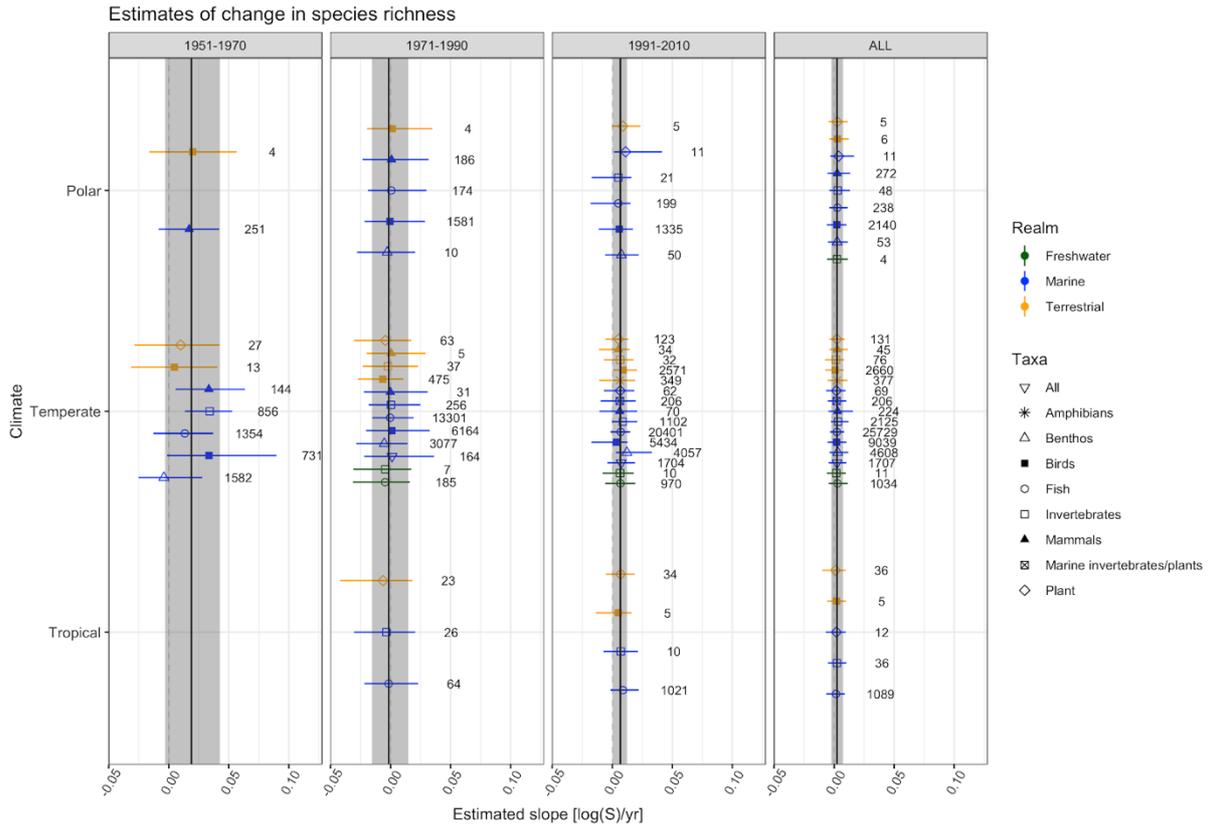


Figure S17.

The estimated slope coefficients for the hierarchical realm-latitude-taxon (RLT) model of species richness change fit to different time periods, or all of the data. Points show the median of the posterior distribution for each combination of realm, latitude and taxon and the line shows the 90% credible interval; solid black vertical line is the global slope estimate (median of posterior distribution) and shaded grey area is its 90% credible. Panels depict either models fit to a specific time period: 1951-1970 (global slope = 0.02; 90% CI = -0.003 - 0.05), 1971-1990 (global slope = -0.001; 90% CI = -0.016 - 0.15) or 1991-2010 (global slope = 0.007; 90% CI = 0.0005 - 0.01), or the model fit to all the data (global slope = 0.002; 90% CI = -0.002 - 0.007); numbers denote the number of cells for each of the groups to which the model was fit. The global slope estimates for the period 1991-2010 differs from zero at the 95% level, compared to 90% or more for the other periods. This suggests that our finding of a no global trend in species richness change is not overly sensitive to time series starting date. All models assumed poisson error and a log link function.

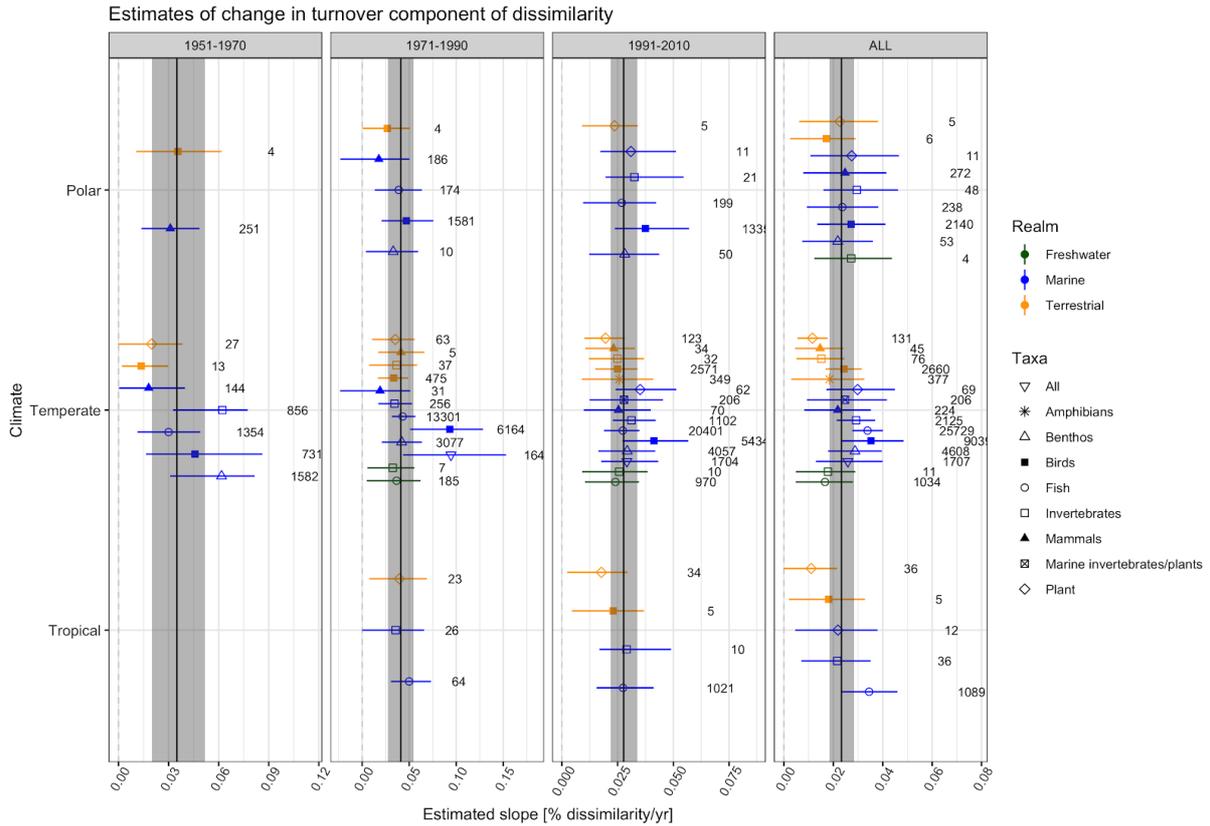


Figure S18.

The estimated slope coefficients for hierarchical realm-latitude-taxon (RLT) model of change in the turnover (replacement) component of dissimilarity fit to different time periods, or all of the data. Points show the median of the posterior distribution for each combination of realm, latitude and taxon and the 90% credible interval; solid black is the global slope estimate (median of posterior distribution) and shaded grey area is 90% credible interval. Panels depict either models fit to a specific time period: 1951-1970 (global slope = 0.035; 90% CI = 0.02 - 0.05), 1971-1990 (global slope = 0.04; 90% CI = 0.027 - 0.055), or 1991-2010 (global slope = 0.03; 90% CI = 0.022 - 0.034), or the model fit to all the data (global slope = 0.02; 90% CI = 0.019 - 0.029); numbers denote the number of cells for each of the groups to which the model was fit. Note that none of the global slope estimates for each of the three time periods falls outside the 90% credible interval for the other time periods, indicating that the results are not strongly sensitive to time series starting date. Model assumed Gaussian error and an identity link function.

Table S1: Supplementary references for data. Study IDs marked with an asterisk are not available in BioTIME (Dornelas et al. 2018 *GEB*), but are available elsewhere (see references).

Study ID	Reference(s)
10	69
18	70
33	(71, 72)
39	(73-76)
41	77
*42	(78-84)
*44	(85-87)
45	(88, 89)
46	90
47	91
51	(92, 93)
52	(94, 95)
53	(96, 97)
54	98
56	99
57	100
58	(101, 102)
59	103
60	(104-109)
63	(110, 111)
67	112
68	113
69	114
70	115
71	116
72	117
73	(118, 119)
74	(120, 121)
75	(122-124)
76	(125, 126)
77	(127, 128)
78	129
81	(130, 131)
84	132
85	133
86	134
87	(135-137)
90	138
91	139
92	140
96	(141, 142)
97	143
98	(144, 145)
99	146
*100-101	(147-149)
108	150
110	151
112	152

113	153
117	154
119	155
120	156
121	157
122	158
123	159
124	(160, 161)
125	162
126	163
127	164
128	165
129	166
133	(167, 168)
135	169
142	(170-172)
143	(173, 174)
147	175
148	176
150	(177-180)
152	181
162	182
163	183
164	184
166	(185-190)
169	(191-194)
171	195
172	(196-199)
173	200
176	201
178	202
180	203
182	204
183	205
184	146
186	206
187	207
189	208
190	209
191	210
192	211
*193	(147-149)
194	212
195	213
196	214
*197	215
*198	216
199	217
200	210
201	218
202	219
204	220
*205	221

*206	222
*207	223
*208	224
*209	225
*210	226
211	227
212	228
213	229
214	230
*215	231
*216	232
217	233
*218	234
*219	235
*220	235
221	236
224	(237, 238)
225	239
226	240
227-228	241
229	242
230	243
231	244
232	245
233	246
234	(247-253)
235	(254-257)
236	258
237	259
238	260
239	261
240	262
241	(108, 263)
242	264
243	(265, 266)
*244	267
245	268
246	269
247	270
248	271
249	272
252	273
253	274
254	275
255	276
*256	225
257	259
270	277
271	278
272	279
273	280
274	281
275	282

276	283
*277 & 279	(85-87)
278	176
280	(144, 145)
281	284
282-283	(123, 124)
284	169
285	169
286-287	159
288	155
289	155
290	285
291	285
292	285
*293	286
294	(287, 288)
295-296	289
297	290
298	291
299	292
300	293
301	(294, 295)
302	296
303	297
304	(298-300)
305	301
306	(302, 303)
307	304
308	305
324-326	319
*309	306
310	307
311	308
312	309
313	310
314	310
315	316
316	372
317	313
*318	314
319	315
321	311
322	317
323	318
327	320
328	321
329	321
330	322
*331	(323-326)
332	327
333	328
334	(329-331)
335	332

336	103
337	333
338	334
339	335
340	336
341	337
342	338
343	339
344	340
345	341
346	342
347	343
348	344
349	(345, 346)
350	(347, 348)
351	(349, 350)
352	351
353	352
354	353
355	354
356	355
357	356
358	357
359	358
360	(359-365)
361	366
362	(367, 368)
363	369
364	370
365	371
366	312
367	373
368	374
369	375
370	376
371	377
372	378
373	379
374	380
375	381
376	382
377	381
378	383
379	383
380	(303, 384)
381	385
382	(303, 386)
383-401	387
402	388
403	389
404	390
405	391
406	392

407-411	<i>(393, 394)</i>
412	395
413	396
414	396
415	396
416	396
418	397
419	398
420	399
421	400
422-423	401
424	402
425	403
426	<i>(404, 405)</i>
427	<i>(404, 405)</i>
428	<i>(406-409)</i>
429	433
430	411
431	411
432	411
433	<i>(412, 413)</i>
434	414
435	415
436	416
437	417
438	418
439-441	419
442-443	420
444	421
445	422
446	423
447	424
448	425
449	426
450-451	427
452	428
453	429
454	430
455	431
456	432
457	410
458-465	<i>(434-436)</i>
466	437
467	<i>(438-441)</i>
468	442
469	443
471	444
473	445
475	446
476	447
477	<i>(448, 449)</i>
478	450
479	<i>(451-456)</i>

480-483	<i>(453-456)</i>
484	<i>(453-457)</i>
485-498	<i>(453-456)</i>
499	<i>458</i>
500	<i>459</i>
501	<i>460</i>
502	<i>461</i>
503	<i>462</i>
504	<i>463</i>
505	<i>464</i>
507	<i>465</i>
508-509	<i>(453-456)</i>
510	<i>466</i>
511	<i>467</i>
512	<i>(468, 469)</i>
513	<i>(470, 471)</i>
514	<i>472</i>
515	<i>(473-476)</i>
516	<i>(477-481)</i>

* These studies are not published in BioTIME but are publicly available and can be accessed via the primary references.