Oikos 126: 18–31, 2017
doi: 10.1111/oik.03652

Subject Editor: Dustin Marshall. Editor-in-Chief: Dries Bonte. Accepted 13 June 2016

A general biodiversity–function relationship is mediated by trophic level

Mary I. O'Connor, Andrew Gonzalez, Jarrett E. K. Byrnes, Bradley J. Cardinale, J. Emmett Duffy, Lars Gamfeldt, John N. Griffin, David Hooper, Bruce A. Hungate, Alain Paquette, Patrick L. Thompson, Laura E. Dee and Kristin L. Dolan

M. I. O'Connor, Dept of Zoology and Biodiversity Research Centre, Univ. of British Columbia, Vancouver, BC, V6T 1Z4, Canada. – A. Gonzalez and P. L. Thompson (http://orcid.org/0000-0002-5278-9045), Dept of Biology, McGill Univ., Montreal, QC Canada. – J. E. K. Byrnes (http://orcid.org/0000-0002-9791-9472), Dept of Biology, Univ. of Massachusetts Boston, Boston, MA, USA. – B. J. Cardinale, School of Natural Resources and Environment, Univ. of Michigan, Ann Arbor, MI, USA. – J. E. Duffy, Tennenbaum Marine Observatories Network, Smithsonian Institution, Washington DC, USA. – L. Gamfeldt, Dept of Marine Sciences, Univ. of Gothenburg, Göteborg, Sweden. – J. N. Griffin, Dept of Biosciences, Swansea Univ., Singleton Park, Swansea, UK. – D. Hooper, Dept of Biology, Western Washington Univ., Bellingham, WA, USA. – B. A. Hungate, Center for Ecosystem Science and Society, Dept of Biological Sciences, Northern Arizona Univ., Flagstaff AZ, USA. – A. Paquette, Centre for Forest Research, Univ. du Québec à Montréal, Centre-ville Station, Montréal, QC, Canada. – L. E. Dee, Inst. on the Environment, Univ. of Minnesota, Twin Cities, Saint Paul, MN, USA. – K. L. Dolan, Research Development Office, Univ. of California at San Francisco, San Francisco, CA, USA.

Species diversity affects the functioning of ecosystems, including the efficiency by which communities capture limited resources, produce biomass, recycle and retain biologically essential nutrients. These ecological functions ultimately support the ecosystem services upon which humanity depends. Despite hundreds of experimental tests of the effect of biodiversity on ecosystem function (BEF), it remains unclear whether diversity effects are sufficiently general that we can use a single relationship to quantitatively predict how changes in species richness alter an ecosystem function across trophic levels, ecosystems and ecological conditions. Our objective here is to determine whether a general relationship exists between biodiversity and standing biomass. We used hierarchical mixed effects models, based on a power function between species richness and biomass production ($Y = a \times S^b$), and a database of 374 published experiments to estimate the BEF relationship (the change in biomass with the addition of species), and its associated uncertainty, in the context of environmental factors. We found that the mean relationship ($b = 0.26$, 95% CI: 0.16, 0.37) characterized the vast majority of observations, was robust to differences in experimental design, and was independent of the range of species richness levels considered. However, the richness–biomass relationship varied by trophic level and among ecosystems; in aquatic systems b was nearly twice as large for consumers (herbivores and detritivores) compared to primary producers; in terrestrial ecosystems, b for detritivores was negative but depended on few studies. We estimated changes in biomass expected for a range of changes in species richness, highlighting that species loss has greater implications than species gains, skewing a distribution of biomass change relative to observed species richness change. When biomass provides a good proxy for processes that underpin ecosystem services, this relationship could be used as a step in modeling the production of ecosystem services and their dependence on biodiversity.

A major goal in biodiversity research is to understand the consequences of biodiversity change for ecosystem functioning (Tilman et al. 1997, Petchey 2000, Turnbull et al. 2013). Experiments have shown that species richness positively affects many ecosystem functions, such as standing biomass and resource use (Tilman et al. 2001, Cardinale et al. 2006, Reich et al. 2012). A nonlinear function captures the relationship between species richness and ecosystem functions, and its prevalence among experimental results suggests a common quantitative relationship might characterize the rate of change of function with changing species richness. Generalized empirical relationships in ecology have allowed for comparisons and predictions across complex systems (Peters 1983, Brown and West 2000). Here we tested whether a general empirical relationship adequately describes the relationship between diversity and biomass production, assessing the degree to which this relationship depends on both random experimental factors and a variety of ecological parameters (e.g. ecosystem type and trophic level). We use this variation to estimate changes in biomass expected with changes in species richness of different organism types.

To estimate how biodiversity change will influence changes in ecosystem functioning in contexts beyond controlled experiments, we need 1) a quantitative estimate of how much function is lost with the loss of a species, 2) reliable estimates of variation around the mean estimate of the BEF relationship and, ideally 3) assignment of uncertainty to factors that are known to influence the relationship (e.g. species traits, resource supply, ecosystem type) as well as factors not yet identified. With existing data, an empirical estimate of the relationship between richness and biomass could be used in biodiversity change models to give a first approximation, or testable prediction, for effects of biodiversity change outside experimental settings. The ecosystem function of biomass production, here estimated as standing biomass at a particular time point and referred to throughout as 'standing biomass', has often been described as a positive decelerating function of species richness (Balvanera et al. 2006, Cardinale et al. 2006, Reich et al. 2012). For competitively structured communities, this relationship may follow the Michaelis–Menten function; however in many experiments, the saturation of biomass production with accumulating species is not clear at the levels of species richness tested (Cardinale et al. 2011). An alternative model that captures the strong effects of species richness at low levels of richness, but diminishing effects at higher richness, is a power function

$$biomass = a \times (richness)^b$$

in which b describes the relationship between a change in richness and biomass (when $0 < b < 1$) (Cardinale et al. 2007, 2011, Reich et al. 2012). The power function used here is not meant to imply a particular theoretical BEF mechanism. It is used because it has substantial empirical support from previous syntheses (Cardinale et al. 2011, Gamfeldt et al. 2014, Lefcheck et al. 2015).

To what extent might the response of biomass production to species richness be generalizable across ecosystems and ecological contexts? If the BEF relationship varies in space and time, or with biotic and abiotic conditions, then estimates of b alone may impart minimal information for understanding the role of biodiversity. However, if it is general or varies predictably, the value of b can provide a powerful tool for efforts to generalize the consequences of species loss for ecological function and ecosystem services (Isbell et al. 2015). Previous grassland experiments reported a central tendency toward a value of approximately $b = 0.26$ for effects of species richness on biomass, but with substantial variation in this estimate. Ninety-five percent confidence intervals ranged from 0.15–0.32 (Cardinale et al. 2006), or a standard deviation of 0.27 (Cardinale et al. 2011). Whether that variation reflects systematic and ecologically important differences among BEF relationships in different systems remains an important question.

The evidence is mixed as to whether abiotic and biotic conditions influence the value of b. Individual experimental studies suggest that the change in biomass with accumulating species richness can vary with resource availability (e.g. water, nutrients, CO_2) (Reich et al. 2001, Fridley 2002, Boyer et al. 2009) or presence of a predator (Duffy et al. 2005), and can increase in strength over time (Stachowicz et al. 2008, Reich et al. 2012). Within experiments that share a species pool, experimental design, and other factors, the BEF relationship can also vary among sites (Hector et al. 1999). Such among-site variation could imply that the strength of the richness–function relationship is contingent on species composition and local environmental condition (e.g. soil fertility, climate, etc.) (Hooper et al. 2012). In contrast, meta-analyses of dozens of experiments have demonstrated that, across studies, estimated mean richness-standing biomass relationships (e.g. b) or effect sizes (e.g. log response ratios) are conserved across experiments conducted in different ecosystem types and trophic groups (Cardinale et al. 2006). However, the values of b do vary systematically with attributes of the experimental design, such as additive and substitutive designs or the total number of richness levels (Balvanera et al. 2006), experimental durations (Cardinale et al. 2007), and spatial and temporal scale (Cardinale et al. 2011). In addition, four recent meta-analyses reported differences in BEF effects among trophic levels: marine herbivore richness had a stronger effect on function than richness of algae (Gamfeldt et al. 2015), aquatic herbivore richness had stronger effects on function than plant richness when multiple functions are analyzed (Lefcheck et al. 2015), carnivore richness more strongly affected resource depletion than did richness at lower trophic levels (Griffin et al. 2013), and detritivore richness more strongly increased decomposition rates than did plant litter richness (Srivastava et al. 2009, Hooper et al. 2012). While individual experiments, and broader meta-analyses, have tested the importance of one or a few additional factors (time, resource supply, trophic structure, etc.; Hooper et al 2012), the relative importance of these factors, the uncertainty in their effects, and whether it is necessary to include these parameters in general richness–function models remains unclear.

Here, we test the hypothesis that a single BEF relationship, expressed as an empirically estimated value of b in a power function, adequately describes the relationship between species richness and standing biomass at fine spatial grains (e.g. m^2 or litres) despite variation across experiments in abiotic conditions, sites, and ecological communities. We then tested biological and experimental conditions, such as different species pools, ecosystem types, trophic levels, resource regimes, and lengths and types of experiments, that might explain variation in this relationship, aiming to identify which factors are essential to understanding the richness–biomass relationship and which might be left out of a general model. We applied our findings to estimate the effects of changes in species richness for changes in standing biomass. Ultimately, our goal is to facilitate integration and quantitative application of the BEF relationship by determining whether experimental evidence supports a general, quantitative relationship (a general b value) between richness and the important ecological function of community biomass production.

Methods

We used a hierarchical mixed effects model to test our hypothesis that there is a constant relationship between species richness and community biomass. We chose standing biomass as the response variable, because theoretical work has centered on this response and hundreds of experimental tests of the relationship between community biomass and richness are published. Standing biomass is closely correlated with net primary production under certain conditions, such as when biomass turnover or size structure is is constant across treatments, or per capita production rates scale isometrically with body size at the individual and population level. When these conditions are not met, biomass may not approximate productivity rates. The data most available – and analyzed here – are for standing stocks only. Modeling standing biomass provides a diverse and large sample to estimate not only the effect of richness on standing biomass but also to test for systematic variation in this relationship among groups.

Our sample of studies included 374 experiments and 558 entries (from 91 studies published between 1985 and 2009, Supplementary material Appendix 1–2) in which richness was manipulated and standing stock of biomass was reported for a species assemblage. The fundamental unit of observation in our analysis was a biomass response variable (e.g. above-ground biomass, density, etc.) reported across a set of species richness (S) treatments varying in the number of species (at least two richness levels), with all other factors controlled, hereafter an 'entry'. For most entries, we lacked data on individual replicates for a treatment in a given experimental unit, and had no choice but to use published means of the richness treatment. Though all entries shared this basic experimental design, they differed in 1) the number of richness levels tested, 2) maximum species richness, 3) the duration of the experiment, 4) whether resources like nutrients or water were added, reduced or unmanipulated, 5) whether experiments were conducted in the lab or in the field, 6) whether the ecosystem studied was aquatic or terrestrial, and 7) in which trophic level diversity was manipulated and biomass reported (Table 1). Many experiments included monocultures (richness = one species), such that 544 of 558 entries included $S = 1$. Across all experiments, the highest richness level tested (S_{max}) increased with the number of richness levels within entries ($r^2 = 0.41$, $p < 0.001$). For each entry, we obtained or estimated a value for each predictor listed above (Table 1). Studies were dropped from the analysis when information on this set of predictors was not available, so there were no unknown values, and the dataset included the same information for all models tested. The number of entries for each level of each predictor was not balanced. The database is dominated by terrestrial plant studies lacking explicit resource manipulations (Table 1). The database also does not include the following combinations: aquatic species richness × resource reduction treatments, terrestrial herbivore richness manipulations, or resource reduction treatments for herbivores or detritivores. There were also insufficient studies reporting the effects of carnivore diversity on carnivore biomass to include in this analysis. Fortunately, hierarchical mixed effects models handle unbalanced designs, and groups with few data points can still contribute some information to the overall analysis (Gelman and Hill 2007).

The search for a single BEF relationship

Our primary objective was to estimate the relationship between richness and biomass. Then, we aimed to test whether a single *b* coefficient described the relationship between richness and biomass given the variation across organisms, ecosystems, and studies performed to date, and if not, to determine what additional information is required to estimate the effect of species richness on biomass. We chose a mixed effects modeling approach that allowed us to characterize the effect of richness on biomass using our structured dataset in which many variables are shared by observations reported from the same experiment or study.

In this dataset, entries within experiments differ in aspects including date sampled or response variable (e.g. above or below ground biomass sampled from the same plot), but share all other attributes such as species richness levels, focal taxa, etc. Experiments within studies differ in treatment

Table 1. Summary of hierarchical dataset on the richness – biomass relationship and predictors analyzed in this study. The most basic unit of observation is an 'entry', which is a single response variable measured at a single time for a set of species richness levels with all other factors controlled. An 'experiment' is the richness manipulation within which all other factors are controlled, but multiple response variables might have been measured at more than one time point, thus there are often multiple entries within each experiment, and several experiments are often published within a single study, and might differ in the level of a factor such as consumer presence, resource supply, etc. Numbers in each column indicate the number of groups (entry, experiment or study) in the dataset for each level of each categorical predictor, and for each continuous predictor the range of values is given for the entire dataset.

Categorical predictors	Levels	Entry (n)	Experiments (n)	Studies (n)
Ecosystem (Sys)	aquatic	134	73	26
	terrestrial	424	301	65
Trophic group (L)	primary producers	501	327	78
	herbivore	26	16	8
	detritivore	31	31	7
Lab/field	lab/greenhouse	178	121	36
	field enclosures or plots	348	221	44
	outdoor mesocosms	46	32	12
Biomass estimator (Units)	biomass	501	339	86
	density	38	30	2
	percent cover	19	3	3
Resource treatment (N)	control	381	241	88
	addition	172	128	22
	reduction	5	5	4

Continuous predictors	Min	Median	Mean	Max
Experimental duration	0.02	1.64	8.48	202.6
Time of measurement (T_G)	0.02	1.05	7.10	202.6
S_{max}	3	6	9.67	43

levels of resources, location or time (e.g. year sampled), but share a publication, research team, and other study-level attributes (Table 1). Mixed effects models allow modeling of variation associated with all unmeasured variables that make parameter estimates from the same group (e.g. a study) similar to each other but distinct from other groups. Hierarchical mixed effects models pool information at the group level, using fewer degrees of freedom and reducing uncertainty in estimated relationships relative to an analysis of each group (e.g. study) independently with regressions (Pinheiro and Bates 2000, Gelman and Hill 2007, O'Connor et al. 2007, Cressie et al. 2009). Hierarchical mixed effects models that account for such structure in datasets are used extensively in social sciences, economics, public health, and other fields where grouped data are the norm (Snijders and Bosker 1999, Gelman and Hill 2007), and provide an information-efficient approach for structured data. Hierarchical modeling is an established ecological research tool well-suited to large datasets comprised of similar, smaller datasets (Myers and Worm 2003, O'Connor et al. 2007, Bolker et al. 2009, Hudson et al. 2013, Lefcheck et al. 2015).

We modeled the biomass–richness relationship at the finest data resolution with the simplest plausible relationship of interest, derived by natural-log-transforming the power function $Y = a \times S^b$. In our case biomass (ln(Y)) at the plot level as predicted by species richness (ln(S)),

$$\ln(\mathrm{Y}_{ijkl}) = B_{0.ijk} + B_{1.ijk} \times \ln(\mathrm{S}_{ijkl}) + \varepsilon_{ijkl} \tag{1a}$$

so that a is estimated by B_0 and the parameter b is estimated by B_1 for plots or mesocosms (l) within each combination of a species richness manipulation and a biomass response (entry, k), entries within experiments (j), and experiments within studies (i). We assumed normally distributed residual error ($\varepsilon_{ijk} \sim (N, \sigma^2)$). Although other formulations have been used to describe this relationship (e.g. Michaelis–Menten, Cardinale et al. 2011), we proceed with the power function, which has also received substantial empirical support and offers greater analytical simplicity, though differs from Michaelis–Menten in that it does not saturate (Cardinale et al. 2007, 2011, Reich et al. 2012, Gamfeldt et al. 2014).

Our hypotheses are centered on the question of how predictable is the value of B_1, the slope of biomass on species richness, or conversely, how variable it is among studies and conditions. Though it is not of primary interest in this study, we also modeled variation in the intercept term, B_0, because predictors of B_1 could influence the intercept (mean biomass), and those influences likely co-vary in some cases with variation in the slope. The intercept term in a BEF regression model represents the absolute value of biomass at standard richness level. Biomass varies among groups of organisms for many reasons – taxonomy of the group involved (algae versus grass versus insects), absolute resource supply rates, etc. Our analysis and dataset are not suited to modeling biomass variation among experimental units (the intercept term). Although we model the intercept, because it may explain some variation in slopes as discussed above, we do not interpret variation in the intercept estimate in terms of the predictors we have included, because we know this set of predictors is insufficient for understanding variation in biomass (the intercept) among entries, experiments and studies.

To test our hypotheses while accounting for variability among experimental conditions and study systems in our dataset, we modeled variation in the slope ($B_{1.ijk}$) and intercept ($B_{0.ijk}$):

$$\begin{aligned} B_{0.ijk} &= \gamma_{00} + \mu_{0.i} + \mu_{0.j} + \mu_{0.k} \\ B_{1.ijk} &= \gamma_{10} + \mu_{1.i} + \mu_{1.j} + \mu_{1.k} \end{aligned} \tag{1b}$$

In Eq. 1b, the slope $B_{1.ijk}$ and intercept $B_{0.ijk}$ for each observation (a set of species richness – biomass observations) are modeled as mean γ_{10} and γ_{00}, respectively. Variation associated with each level of data grouping – entry ($\mu_{1.k}$, $\mu_{0.k}$), experiment ($\mu_{1.j}$, $\mu_{0.j}$) and study ($\mu_{1.i}$, $\mu_{0.i}$) – can be formally considered as random effects normally distributed with variance Σ_0 estimated by the model (Supplementary material Appendix 3).

The test of our first hypothesis, that there is a constant relationship between species richness and standing biomass, is whether variable slopes (b, as estimated by $B_{1.ijk}$) are required among different studies, experiments and entries ($\mu_{1.i}$, $\mu_{1.j}$, or $\mu_{1.k} \neq 0$, in Eq. 1b) (Table 2). If so, we would conclude that it is not possible to identify a single parameter (b) for this dataset. We also tested alternate models (Eq. 2–3) that include interaction terms for time (T_G) passed since the beginning of the experiment normalized to the generation time of the taxon (such that $\mathrm{T}_G =$ duration of experiment in days/generation time of focal organism) (B_2), and the interaction between T_G and ln(S) (B_3):

$$\begin{aligned} \ln(\mathrm{Y}_{ijkl}) = {} & B_{0.ijk} + B_{1.ijk} \times \ln(\mathrm{S}_{ijkl}) \\ & + B_2 \times \ln(\mathrm{T}_{G.ijkl}) + \varepsilon_{ijkl} \end{aligned} \tag{2}$$

$$\begin{aligned} \ln(\mathrm{Y}_{ijkl}) = {} & B_{0.ijk} + B_{1.ijk} \times \ln(\mathrm{S}_{ijkl}) + B_2 \times \ln(\mathrm{T}_{G.ijkl}) \\ & + B_3 \times \ln(\mathrm{S}_{ijkl}) \times \ln(\mathrm{T}_{G.ijkl}) + \varepsilon_{ijkl} \end{aligned} \tag{3}$$

These models test for effects of plot-scale richness and plot age, and are possible because paired richness, function data were reported for multiple time points in many studies. Generation time was approximated based on body size and knowledge of taxa (Cardinale et al. 2011).

Testing hypotheses about factors that modify the BEF relationship

We tested our second main hypothesis that ecological or experimental parameters that varied across entries, experiments or studies altered the richness–biomass relationship (Table 2). Specifically, we compared mixed effects models with different formulations that represent hypotheses for how various biotic and abiotic factors (listed in Table 1) interact with species richness to affect the relationship.

In addition to the basic hypothesis that biomass changes with increasing species richness and time (model 1–3), we tested the hypothesis that ecosystem (aquatic, terrestrial) and trophic group (primary producer, detritivore, herbivore) influence the richness–biomass relationship (slope $= B_{1.ij}$, model 4, Table 2, 4). The trophic group predictor indicates the group for which species richness was manipulated and biomass was measured. In this hypothesis, we included an interaction between ecosystem and trophic group to allow for the lack of data on terrestrial herbivores. We also tested the hypotheses that in addition to ecosystem and trophic group, increased or reduced resources (water, nutrients,

Table 2. Hypotheses tested for how abiotic and experimental factors could affect the relationship between species richness and biomass. Taken together, this set of hypotheses allowed us to test the overarching hypothesis that a single BEF relationship, expressed as an empirically estimated value of *b* in a power function, adequately describes the relationship between species richness and standing biomass at fine spatial grains (e.g. m^2 or litres) despite variation in abiotic conditions, sites, and ecological community contexts such as different species pools, ecosystems, trophic levels or resource regimes.

Hypothesis	Model
There is a constant relationship between species richness and standing biomass, estimated as B_1 in Eq. 1a. Variation in unmeasured attributes at different levels of organization in the data (entry-, experiment- or study-level variation) did not alter estimates of the richness–biomass relationship	All, with random effects (Eq. 1b)
The richness–biomass relationship increases with time within experiments, estimated as generation times of focal taxa (T_G)	3
The effects of experimental duration (T_G) within experiments varies among ecosystems (Sys) and trophic levels (L)	3a
The effects of experimental duration (T_G) within experiments varies among trophic levels (L)	3b
The richness–biomass relationship varies among ecosystems (Sys) (aquatic, terrestrial) and trophic groups (primary producer, detritivore, herbivore) (L)	4
In addition to variation among ecosystems (Sys) and trophic groups (L), the richness–biomass relationship varies with increased or reduced resources (water, nutrients, CO_2) (N)	5
Experiments that run for a greater number of generations (T_G) show stronger effects of richness when compared across taxa or systems, in the context of resource addition or reduction (N)	6
The effects of maximum experimental duration vary with level of resource addition (N)	7
The richness-biomass relationship varies with attributes of the experimental design – maximum duration, maximum number of species tested (S_{max}), units in which biomass was measured (biomass estimator), and lab versus field (model 8)	8
All biotic, abiotic and experimental factors (model 1–8) modify the BEF relationship (Eq. 3)	9
When all predictors are considered, the interaction between ecosystem and trophic group is not important (Eq. 2)	9.1

CO_2) modified the BEF relationship (model 5). The three categorical levels of the resource treatment predictor (control, addition, reduction) reflect experimental manipulations relative to ambient conditions for any resource explicitly manipulated (water, nitrogen, light, etc). A level of 'control' was assigned to any species richness manipulation that did not specify that resources were added or reduced relative to ambient levels. Some experiments manipulated resource supply to plants and factorially with consumer richness manipulations, and we included these studies. This resource predictor includes no information on whether the resource was a priori shown to be limiting in the system, and not all experiments included factorial resource treatments (Table 1). Consequently, the 'resource' predictor represents a coarse test of whether resource manipulation modifies the richness–function relationship.

In a fourth hypothesis, we tested for an effect of experimental duration among studies standardized to estimated generation time of the manipulated taxa, testing whether experiments that run for a greater number of generations show stronger effects of richness when compared across taxa or systems (model 6, Table 2, 4). We considered time in two ways. First, we modeled the effect of time 'within an experiment', looking at whether the slope parameter changes as an experiment moves from year 1 to year 2 to year 3. We might expect the parameter to change over time based on studies in long-term experiments such as Reich et al. 2012 and Stachowicz et al 2008. This effect of time is captured by the parameter T_G, and models the effect of year (or day) within a multi-year (day) experiment (Table 2, model 2). Second, we examined the effect of time by modeling the effect of total experiment duration on the slope *b*. This model tested whether longer experiments have steeper slopes (the parameter is called 'ln(maxDuration)', which is measured by the number of generations of the focal taxa in the experiment) (Table 2, model 6). We also tested whether the effect of total experiment duration depended on whether resources were added or reduced (Table 2, model 7).

We also tested the hypothesis that the BEF relationship varies with attributes of the experimental design – maximum duration, maximum number of species tested (S_{max}), units in which biomass was measured (biomass estimator), and lab versus field (model 8). Finally, we tested the hypothesis that all factors modify the BEF relationship (model 9, Eq. 4), and that when all are included, the interaction between ecosystem and trophic group is not important (model 9.1). We modeled interactions between intercepts ($B_{0.ijk}$) and slopes ($B_{1.ijk}$) for each group using the following equations, and each hypothesis outlined above was modeled as a nested subset of the full model:

$$\begin{aligned} B_{0.ijk} = {} & \gamma_{00} + \gamma_{01} \times Sys_i + \gamma_{02} \times L_i \\ & + \gamma_{03} \times Sys_i \times L_i + \gamma_{04} \times \mathrm{Units}_j + \gamma_{05} \times LabField_i \\ & + \gamma_{06} \times S_{max.il} + \gamma_{07} \times N_j + \gamma_{08} \times \\ & ln(max(Duration_i)) + \mu_{0.i} + \mu_{0.j} + \mu_{0.l} \end{aligned} \tag{4a}$$

$$\begin{aligned} B_{1.ijk} = {} & \gamma_{10} + \gamma_{11} \times Sys_i + \gamma_{12} \times L_i \\ & + \gamma_{13} \times Sys_i \times L_i + \gamma_{14} \times Units_j + \gamma_{15} \times LabField_i \\ & + \gamma_{16} \times S_{max.il} + \gamma_{17} \times N_j + \gamma_{18} \times \\ & ln(max(Duration_i)) + \mu_{1.i} + \mu_{1.j} + \mu_{1.k} \end{aligned} \tag{4b}$$

with random effects, normally distributed about zero with variance estimated by the model (Supplementary material Appendix 3).

Model selection, analysis and inference

To identify the best model, we first determined the need for variable slopes and intercepts for each candidate model (Eq. 1a, 2, 3) by comparing models with different random effects structures (Supplementary material Appendix 3 Table A1). The test of our first hypothesis is whether the BEF model requires variable slopes at group (entry, experiment and study) levels, implying variation in the BEF relationship among groups. We ranked models with and without variable slopes and intercepts using AIC_c adjusted for degrees of freedom to account for different random effects following Bolker et al. (2009) and Gelman and Hill (2007), and compared them using δAIC_c values (Bolker et al. 2009). If variable slopes were required at the group level, we examined residuals ($\mu_{1.k}$, $\mu_{1.j}$, $\mu_{1.l}$) using caterpillar plots to determine whether only a few studies drove the need for variable slopes at the group level (Verbeke and Molenberghs 2000).

To test our second hypothesis, we compared models with biotic and experimental predictors (model 4–9; Eq. 4a–b). We ranked models using AIC_c, and compared them with δAICc and Akaike weights (*w*). We defined the best model set as all models with $\delta\text{AIC}_c < 2$, (Burnham and Anderson 2002, Richards 2005). If more than one model met our criteria of $\delta_{aic} < 2$, we averaged these models to produce coefficient estimates (Burnham and Anderson 2002). Model averaging produces estimates for all coefficients in the best model set, weighted by the importance (*w*) of each model in the set. To estimate the parameter *b* for each study, we summed coefficients for each richness manipulation ($b = B_{1.ijk} + \mu_{1.i} + \mu_{1.j} + \mu_{1.k}$) (Gelman and Hill 2007) from the best model set.

We proceeded with a linear mixed effects model, although in our dataset, $\ln(Y_{ijkl})$ values have a fat-tailed distribution and are not strictly normally distributed (Shapiro-test, $p < 0.001$), differing from normal but without significant skew. Analysis of residuals of Eq. 1 revealed seven experiments from two studies that were extreme outliers in the dataset (< 3% of entries), and these were excluded from analysis to meet assumptions of homoscedasticity. Although we tested for an effect of time and there is a risk that observations are temporally autocorrelated, we could not include a temporal autocorrelation term in the model because time and richness are modeled at the finest resolution of our hierarchical data. Thus, there are multiple observations (biomass at multiple richness levels) for each level of T_G within each entry, and we cannot isolate potential autocorrelation in time from among richness levels. Therefore, the $\ln(T_G) \times \ln(S)$ fixed effect is expected to include any real effects and any possible autocorrelation. All analyses were done in R (ver. 3.2.1). We used lme4 package (ver. 1.1-8) for mixed effects analyses, comparing models fit with REML = FALSE but used REML = TRUE for estimation of coefficients. Data and analytical code are available from the authors at <https://github.com/mioconnor78/OConnor-et-al-BEF-Relationship>. We produced caterpillar plots using the package sjPlot (ver. 1.8.2).

To consider the relationship between a change in richness and a change in function, we simply generated a normal distribution of numbers between 0 and 2, centered on 1, to represent a distribution of proportional species richness changes centered on no net change. We then used the empirical estimates of *b* produced from the previous analyses to translate a distribution in species richness change to an expected distribution of change in biomass. We present these as proportional changes in species richness and proportional changes in function to compare the effect of species loss or gain across systems with different total species richness.

Results

A single, universal relationship ($B_{1.ijk}$ value) was not supported by our analysis of the species richness–biomass relationship. Variable slopes and intercepts associated with entry, experiment and study were required for each candidate model (Eq. 1–3) ($\text{AIC}_C > 100$ for comparison of model with variable slopes and intercepts to model with fewer random effects terms) (Supplementary material Appendix 3 Table A1). We did not find strong evidence for a systematic effect of experimental duration on the BEF relationship across all studies (Table 3). This simple model with random effects (model 2; Eq. 2) estimates a BEF relationship of $b = \gamma_{10} = 0.23$ (95% CI: 0.18, 0.28) that applies to most (> 94.00%) but not all entries (Supplementary material Appendix 3 Fig. A1, Table A1). Examination of the variation in slopes ($\mu_{1.i}$, $\mu_{1.j}$ and $\mu_{0.k}$), plotted as the deviation of each slope's estimated random effect from the mean slope fixed effect (Fig. 2), suggests this estimate of $b = \gamma_{10}$ adequately described most observations (i.e. the confidence intervals for the random effects include 0 in the caterpillar plots for most ln(S) estimates) (Supplementary material Appendix 3 Fig. A1). Still, the number of slope residuals deviating from the central estimate (γ_{10}) is sufficient that removing those observations neither eliminates the need for variable slopes, nor is justified based on the dataset. In all models, richness values were centered on the value that minimized co-variances of random effects for slopes and intercepts estimated by model 1 (ln(8)). These covariances were 0.07 for Entry, 0.24 for Experiment and –0.16 for Study. Examination of standardized residual plots suggests models were not overfit (Supplementary material Appendix 3 Fig. A3).

After concluding that variable slopes at the entry, experiment and study levels confound the identification of a single relationship, we tested our second main set of hypotheses that ecological and experimental factors could explain some variation in richness–biomass relationship, thus eliminating

Table 3. Results of model selection for basic species richness–biomass model. Models relate total estimated biomass (ln(Y)) to species richness (ln(S)) and experimental duration, estimated in terms of number of generations of focal taxa ($\ln(T_G)$). Models are ranked by AICc, and compared using AIC weight (*w*), δAIC values and likelihood ratio tests. Likelihood ratio tests (p-values) compare each model with the top-ranked (lowest AICc value) model (first row) and facilitate interpretation of the significance of differences in similar AICc values. All models include variable slope and intercept coefficients at the entry, experiment and study level (Table A1).

Model		AICc	*w*	DF	modLik	δ	p
2	$\ln(Y) = \ln(S) + \ln(T_G)$	1451.5	0.49	13	–712.67	0.00	–
3	$\ln(Y) = \ln(S) \times \ln(T_G)$	1452.4	0.30	14	–712.13	0.95	0.30
1	$\ln(Y) = \ln(S)$	1453.1	0.21	12	–714.50	1.65	0.09

the need for variable slopes ($\mu_{1.i}$, $\mu_{1.j}$ and $\mu_{0.k}$) based on diagnosis using caterpillar plots (Supplementary material Appendix 3 Fig. A1–2). We found that the BEF relationship varied systematically between aquatic primary producers and consumers (model 4, Table 4) such that herbivore biomass increased with species richness by $b_{aq.herbivores} = \gamma_{10} + \gamma_{12} = 0.47$, whereas detritivore biomass increased with species richness by $b_{aq.\ detritivores} = \gamma_{10} + \gamma_{12} = 0.55$, both stronger than the relationship between primary producer (plants, algae) biomass and primary producer species richness ($b_{prim.prod} = \gamma_{10} = 0.26$) (Table 4, Fig. 1B). Though there was no difference between aquatic and terrestrial primary producers, terrestrial detritivores had a much weaker relationship between richness and biomass than all other groups (aquatic primary producers and herbivores, and terrestrial plants), with a negative value for *b* (Table 5, Fig. 1B–C). Estimates for the intercept term varied among trophic and ecosystem groups, as expected by their very different biomasses (Table 5, Supplementary material Appendix 3 Fig. A2). The top-ranked model of our set was model 4 (Table 4), which included the interaction between trophic group and ecosystem (Fig. 1B). None of our other hypotheses about variation in the BEF relationship were comparable to this 'best' model (AIC weight = 0.784).

The best model indicates that variable slopes and intercepts are still required, even with fixed effects for trophic group and ecosystem (model 4, Supplementary material Appendix 3 Table A2, Fig. 2). Thus, systematic variation remains among entries, experiments, and studies that prohibits a single estimate of a BEF relationship between ln(Y) and ln(S) (Fig. 2). The larger variance components associated with study and entry compared to experiment suggests that most of the unexplained systematic variation is at those levels. Slope estimates did not differ systematically for experiments with or without monocultures (Supplementary material Appendix 3 Fig. A4).

Our model comparison results allowed us to reject some of our alternate hypotheses (Table 4). We rejected the hypothesis that the basic model (Eq. 2) is sufficient to explain the relationship between richness and biomass. We also rejected the hypotheses (model 8, 9 and 9.1) that differences in experimental designs (number of species tested, lab versus field experiment, and the method of estimating biomass) explain variation in the observed BEF relationship, and that our coarse grouping of resource supply condition (addition, reduction or control) provided needed information for understanding the richness–biomass relationship. Results of the test of the importance of number of generations (T_G) within experiments did not support the hypothesis that including the interaction between ln(S) and ln(T_G) led to a significantly better fit (Table 3). AICc values and the likelihood ratio test suggest that the richness–biomass

Table 4. Comparison of alternative models for how richness (ln(S)) affects biomass. Model terms are as shown in Eq. 4a and b (intercept term not shown here), ranked from left to right by their quality (high to low) as a description of this dataset. Models differed in fixed effects, indicated by ✓, but all included variable intercepts (μ_{0i},μ_{0j},μ_{0k}) and slopes (μ_{1i},μ_{1j},μ_{1k}) at the level of the study (*i*), experiment (*j*) and entry (*k*). Models were ranked based on AICc, and differences assessed using δAIC and Akaike weights (***w***) and likelihood ratio tests. We used likelihood ratio test results (– $p > 0.05$, $^*p < 0.05$, $^{**}p < 0.01$) to compare models with the top-ranked (lowest AICc) model only for comparisons in which one model can be derived from the other by constraining parameter values. When this was not possible, likelihood ratio tests were not performed. A significant p-value indicates that the model with the lower AICc value is a better description of the data. When the likelihood ratio test indicates no differences, the model with fewer parameters is preferred.

		Model									
Predictor	Term	4	5	6	4.2	9	2	9.1	3	8	7
ln(S)	γ_{10}	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Time (ln(T_G))	B_2	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Ecosystem (Sys)	γ_{01}	✓	✓	✓	✓	✓	✓	✓		✓	
Trophic group (L)	γ_{02}	✓	✓	✓	✓	✓	✓	✓		✓	
Sys*L	γ_{03}	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Resource treatment (N)	γ_{07}		✓	✓		✓		✓		✓	✓
ln(max(Duration))	γ_{18}			✓		✓		✓		✓	✓
Lab vs field experiment	γ_{05}					✓		✓		✓	✓
Biomass estimator (Units)	γ_{04}					✓		✓		✓	✓
ln(S_{max})	γ_{06}					✓		✓		✓	✓
ln(S) × ln(T_G)	B_3							✓	✓	✓	✓
ln(S) × Sys	γ_{11}	✓	✓	✓	✓	✓		✓		✓	✓
ln(S) × Sys × ln(T_G)	γ_{13}	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
ln(S) × L	γ_{12}	✓	✓	✓	✓	✓		✓		✓	✓
ln(S) × N	γ_{17}		✓	✓		✓		✓		✓	✓
ln(S) × ln(max(Duration))	γ_{18}			✓		✓		✓		✓	✓
ln(S) × (lab vs field)	γ_{15}					✓		✓		✓	✓
ln(S) × Units	γ_{14}					✓		✓		✓	✓
ln(S) × S_{max}	γ_{16}					✓		✓		✓	
	DF	21	25	27	19	35	13	33	14	23	19
	AIC_c	1437.8	1442.9	1443.1	1443.1	1443.5	1451.5	1452.3	1452.4	1455.4	1459.1
	δ	0	5.13	5.33	5.34	5.70	13.72	14.52	14.67	17.68	21.30
	w	0.784	0.060	0.054	0.054	0.045	0.001	0.001	0.001	0	0
	logLik	–697.7	–696.2	–694.3	–702.4	–686.2	–712.7	–692.7	–712.1	–704.5	–710.4
	p		–	–	**	–	**		**		

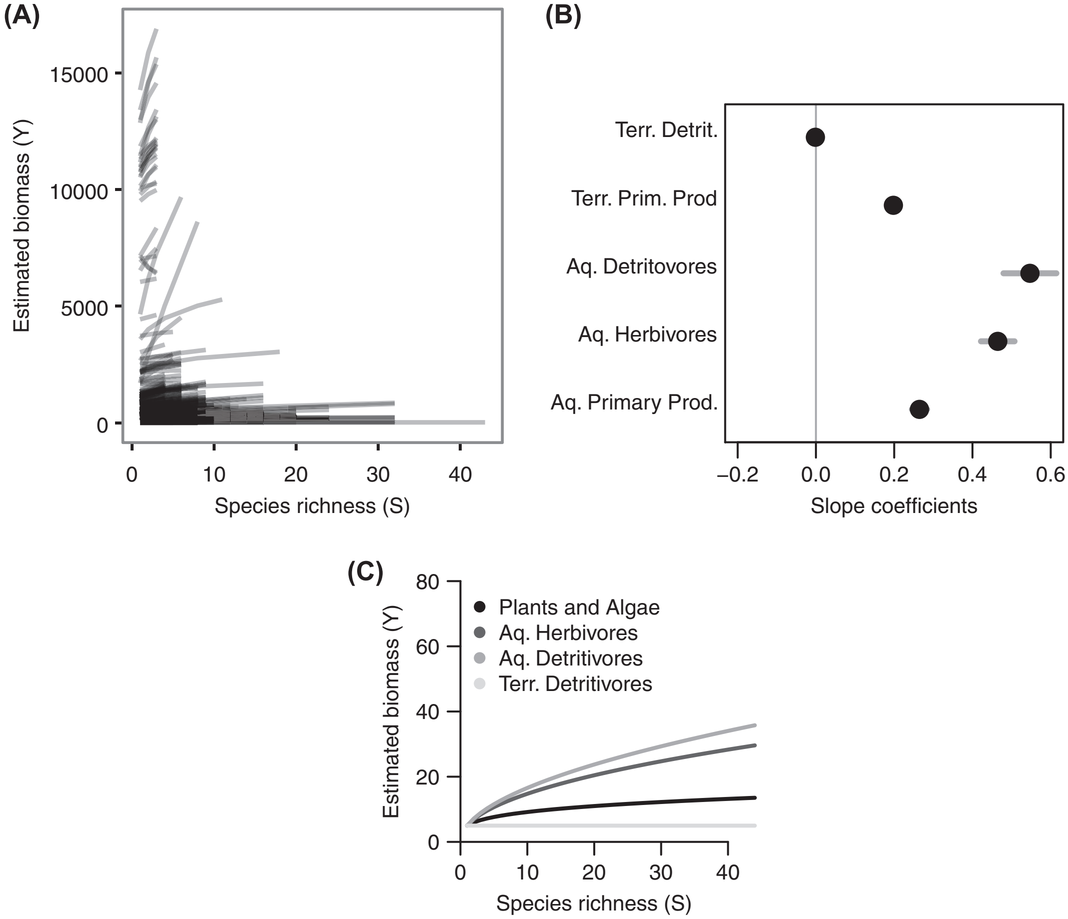

Figure 1. (A) Standing stock (biomass) plotted against species richness as a power function ($Y = a \times S^b$) relating standing biomass (Y) to species richness (S) via an intercept (*a*) and scaling parameter (*b*) for each entry in our database ($n = 558$). Each entry is plotted in gray, dark lines indicate overlapping lines. Each entry was analyzed in a hierarchical mixed effects model using a linearized power function (Eq. 1, 2). (B) Empirically estimated scaling parameters for BEF relationships vary among trophic groups and between aquatic and terrestrial systems. Estimates are based on model coefficients for the slope term ($B_{1.ijk}$) from the best model (model 4, Table 3; black points). Standard errors shown in this figure are errors of the mean estimate from the distribution of fitted slopes for this dataset. Confidence intervals estimated from the model output are shown in Table 4. (C) Power functions plotted with *b* values shown in panel B for primary producers, aquatic herbivores and aquatic detritivores.

relationship did not depend on the duration of an experiment expressed as the generation time of the organisms being studied (e.g. no. generations, *L*) across the 374 experiments in our dataset (no support for the interaction between time and richness, model 2 in Table 4). We conducted two additional tests of the hypothesis that experimental duration might affect the strength of the relationship. In the first of these, we expanded model 3 (Table 4) to test alternate hypotheses that there is an interaction between ecosystem, time and richness (model 3a) or between trophic group, time and richness (model 3b) (Supplementary material Appendix 3 Table A4). The inclusion of the interaction term for time in these models suggests an effect of time could be informative, yet coefficients for the T_G interactions did not differ from 0 except for herbivores, which suggests a weak negative effect of time on the BEF relationship. Second, we tested the effect of maximum duration on the relationship for only final observations of each experiment (model 6, 7). Model comparisons for this dataset were consistent with the full dataset, and suggested no effect of maximum duration on the parameter *b* (Supplementary material Appendix 3 Table A3). We did not find evidence for this effect of time on plants or detritivores across all studies (Supplementary material Appendix 3 Tables A4–A5).

When applied to a distribution of scenarios of species richness change over time, the estimated values of *b* (Table 5) produced distributions of expected biomass change that reveal net negative effects of species richness change on biomass. A distribution of species richness changes centered on no change (or proportion of richness before:after change = 1), produces a distribution of expected biomass change with a mean proportional change < 1 (Fig. 3), and greater extreme values for loss of function than for gain. The larger *b*-value for aquatic herbivores suggests much greater losses or gains in biomass expected for a given change in species richness relative to primary producers. For example, a 20% loss of species richness for herbivores leads to a 10% loss in herbivore biomass, while the same loss of plant species richness leads to a 6% loss of plant biomass.

Table 5. Coefficients for modeled effect of richness on standing stock. Mean (± 95% CI) estimate from the best model with fixed effects (trophic level, duration, lab versus field tests and ecosystem) and variable slopes and intercepts (model 4). Estimates give effect sizes relative to plant biomass in a terrestrial ecosystem under nutrient control conditions. Values in bold indicate parameter estimates contributing to the slope term that differ significantly from zero and thus modify the relationship between richness and biomass.

	Factor	Term	Model 4
Fixed effects	Intercept	γ_{00}	4.33 [3.45, 5.21]
	ln(S)	γ_{10}	**0.26 [0.16, 0.37]**
	ln(T_G)	B_2	0.16 [0.03, 0.28]
	Ecosystem - Terrestrial	γ_{01}	1.50 [0.48, 2.51]
	L – Herbivore	γ_{02}	–0.33 [–1.38, 0.66]
	L – Detritivore	γ_{02}	1.22 [–0.93, 3.39]
	Terrestrial × Detritivore	γ_{03}	–1.93 [–5.11, 1.26]
	ln(S) × ecosystem –Terrest.	γ_{11}	–0.07 [–0.18, 0.05]
	ln(S) × L – Herbivore	γ_{12}	**0.21 [0.03, 0.38]**
	ln(S) × L – Detritivore	γ_{12}	**0.29 [0.01, 0.56]**
	ln(S) × Terrest. × Detrit.	γ_{13}	**–0.58 [–0.98, –0.17]**
Random effects	Entry – intercept	Σ_{0k}	0.38
	Entry – ln(S)	Σ_{1k}	0.03
	Experiment – intercept	Σ_{0j}	0.72
	Experiment – ln(S)	Σ_{1j}	< 0.01
	Study – intercept	Σ_{0i}	3.31
	Study – ln(S)	Σ_{1i}	0.03
	Residual	σ^2	0.02

Discussion

We synthesized over 500 experimental tests of the effect of species richness on biomass, and found strong support for a relationship that varies among trophic levels (Fig. 1B–C) and potentially across ecosystem types. We found that most (but not all) primary producer assemblages in both terrestrial and aquatic environments conform to a single coefficient ($b = 0.26$) quantifying how biomass increases with increasing species richness. However, aquatic consumers had much stronger effects of diversity on biomass (herbivores, $b = 0.47$; detritivores, $b = 0.54$), and terrestrial detritivores had no clear relationship ($b = –0.001$), compared to plants and algae. We conclude that information about trophic group and ecosystem can inform estimates of the consequences of species loss or gain for one ecosystem function, standing biomass. Our hierarchical mixed effects modeling approach provided one of the more comprehensive analyses of the richness–biomass relationship to date, simultaneously considering the potential dependence of the richness–biomass relationship on 8 abiotic and biotic factors and additional systematic variation across hundreds of experimental tests. Finally, our results indicate that, on average, species losses will result in greater losses of biomass than will species gains result in increased biomass. We discuss the potentially important implications of this result later in the Discussion.

The observed stronger BEF relationship in primary consumers relative to primary producers has been predicted conceptually (Duffy 2002, 2003). Although early data syntheses did not detect this difference, that could be explained by smaller numbers of studies and relatively simple statistical methods of data synthesis (Cardinale et al. 2006). Recently, using variants of the dataset we used here, Gamfeldt et al. (2014) found that in marine studies, herbivore biomass increases more strongly with richness than does primary producer biomass, and Lefcheck et al. (2015) reported stronger effects of aquatic herbivore than primary producer diversity on multiple functions. Similarly, Griffin et al. (2013) found stronger effects of species richness on resource depletion rates for higher trophic groups. We confirm this result for a larger dataset that includes terrestrial studies, suggesting that as more data has become available, previous findings that herbivores did not differ from plants can now be revised.

The larger effects of species richness change for aquatic herbivore and detritivore biomass than primary producer biomass leads to the hypothesis that changes in diversity could create positive feedbacks in aquatic systems. Because the magnitudes of the consequences increase nonlinearly as species richness declines, greater diversity declines among consumers than resources (Byrnes et al. 2007, Duffy 2003) could shift the importance of top–down control: as grazer species loss disproportionately reduces grazer biomass, plant assemblages released from grazing pressure could increase plant productivity. Similar mechanisms could influence the relationship between species richness and biomass for decomposers, and also for decomposition rates. A meta-analyses based on the same data set found that changes in detritivore diversity had greater effects on changes in decomposition rates than did changes in plant litter diversity, particularly among aquatic detritivores (Hooper et al. 2012).

We found that patterns in the richness–biomass relationship across trophic levels differed in terrestrial systems. In this dataset, detritivores were represented by primary decomposers: unicellular fungi (Replanksy and Bell 2009), multicellular fungi (Setälä and McLean 2004), and bacteria (three studies, 24 Experiments with distinct species compositions at each richness level, 24 Entries). While the former two studies found evidence for positive effects of fungal diversity on fungal biomass without disturbance, disturbance by drought reversed this effect (Setälä and McLean 2004), and bacterial diversity had no effect on a variety of soil processes. The pattern of weaker richness–biomass relationships among terrestrial detritivores relative to primary producers in this small set may reflect the taxonomic

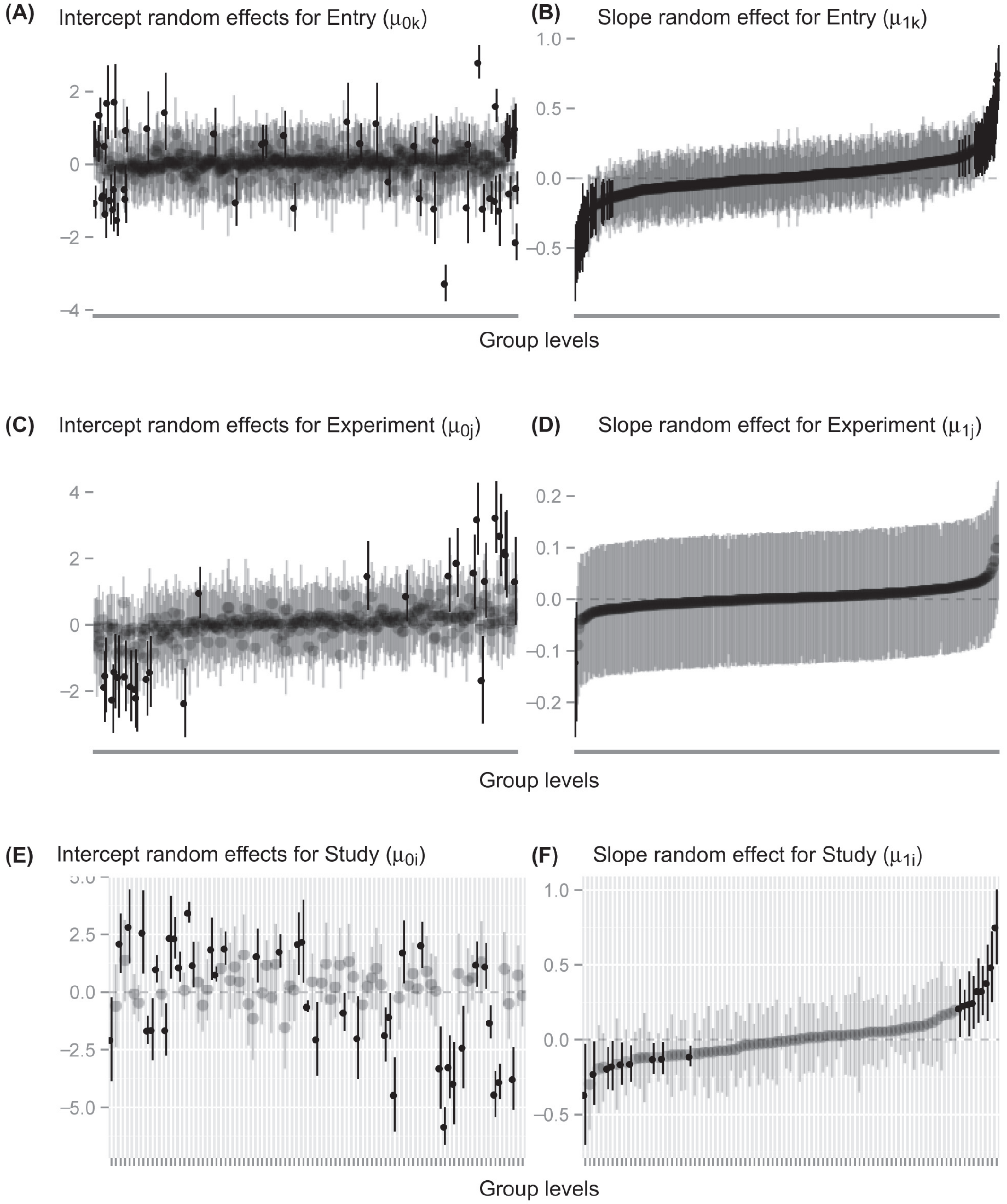

Figure 2. Random effects (± CI) estimated by the best model (model 4, Table 2) associated with entry$_k$ (plots A–B), experiment$_j$ (panels C–D) and study$_i$ (panels E–F) for intercepts (μ_0) and slopes (μ_1), ranked by slope random effects (μ_1). Gray CI's include 0, indicating that the estimated random effect cannot be distinguished from the fixed effect for slope or intercept. Random effects different from zero imply that the coefficient for that study can be estimated as the fixed effect plus the random effect.

bias toward microbial consumers in terrestrial systems relative to larger-bodied detritivores (e.g. macroinvertebrates) in the aquatic studies included here. However, the interpretation of a general, cross-system trophic level effect in terrestrial systems is hampered by lack of available data. The few studies using terrestrial detritivores are insufficient to understand the generality or causes of variation in that relationship. Furthermore, our dataset included no terrestrial herbivore manipulations for which herbivore biomass was reported, so we have no basis for inference about that relationship. Similarly, we lacked sufficient estimates of carnivore biodiversity manipulations that reported effects on carnivore biomass to include them in this analysis. These issues highlight the need for further exploration of biodiversity–ecosystem functioning relationships in these understudied groups.

We did not find a systematic relationship between the BEF relationship and experimental duration (T_G) across this dataset (Table 3). Still, for several reasons, we cannot reject the hypothesis that the BEF relationship changes though time within a community. First, there is strong evidence in the literature, including one meta-analysis, that have reported that the richness–biomass relationship strengthens through time (Cardinale et al. 2007, Stachowicz et al. 2008, Reich et al. 2012). Further, in some of the longest-running

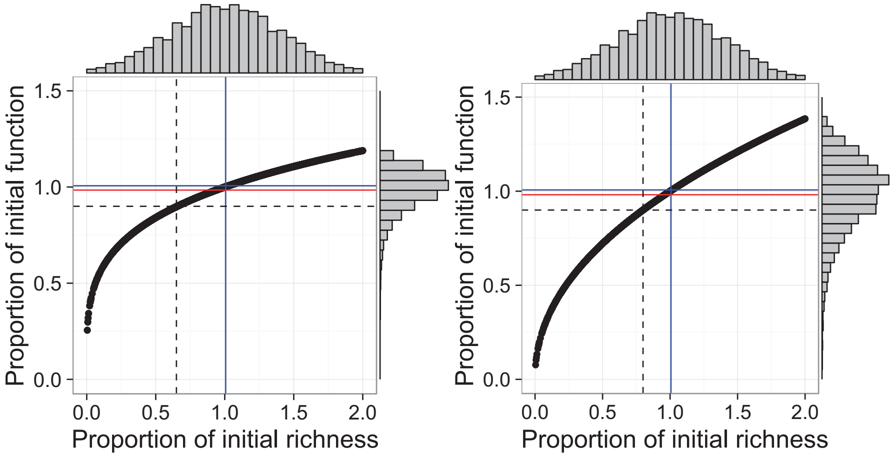

Figure 3. Expected change in biomass associated with changes in species richness. Distribution of species richness changes (top histogram), expressed as a response ratio ($\ln(S_{\mathrm{Time1}}/S_{\mathrm{Time2}})$), and the distribution of associated change in biomass (vertical histogram), expressed as $\ln(Y_{\mathrm{Time1}}/Y_{\mathrm{Time0}})$ expected for (A) primary producers and (B) herbivores. The distribution of expected function was produced using $Y = a \times S^b$ (the plotted curve) for values of $b = 0.26$ for plants, and $b = 0.47$ for herbivores (Table 4, Fig. 1B), Solid blue lines indicate response ratios of 1 = no change in richness; and the red lines indicate the mean expected function. Dashed lines identify a 10% decline in standing biomass, and the intersection with the BEF curve identifies the change in richness expected to cause a 10% change in function: a 35% reduction in plant richness, and a 20% reduction in herbivore richness.

BEF experiments, an effect of duration is clear after several years (Stachowicz et al. 2008, Reich et al. 2012). Second, time may have had variable effects among studies. We found that entry-level random effects for the coefficient *b* were required, and one of the main differences between entries within an experiment is the time of measurement. The persistence of the entry-level random effect for the BEF relationship could reflect temporal variation, that variation may not be linear through time in all studies, or that co-varying factors such as climate conditions explain effects associated with time. A similar argument could be made for the importance of total experimental duration (maximum duration, Table 1), which varies among studies and could therefore also be accounted for in the study-level variance component. At the study-level, variation in study duration is typically confounded with variation in spatial scale and body size of the focal taxa (Cardinale et al. 2011), such that time effects cannot be clearly distinguished.

Our failure to reject the need for variable slopes among certain groupings (entries, experiments, and studies) indicates that systematic variation in the BEF relationship exists among studies and experiments, not captured by our hypotheses. This suggests that additional research and synthesis is needed to determine whether there is a single BEF relationship, or whether attributions such as climate or higher resolution treatment of predictors, such as resource supply, could explain the remaining variation. Among-study variation explained the majority of the variation in the random effects in our model (Table 5). Random, study-level variation is distinguishable from residual variation (error) and implies that in addition to the fixed effects that we modeled, there is still systematic variation in how richness affects function among studies. This variation could result from climate, site environmental parameters (e.g. soil pH), taxonomic groups studied, species or functional trait composition within those groups, or other ecological or scientific particularities of the research studies.

Our results also help shed light on potential changes in function resulting from changes in local species richness. Our results show that, on average, species losses lead to greater losses of biomass than species gains lead to increased biomass (Fig. 3). The disproportionate effects of species losses compared to gains follows from Jensen's inequality theorem, which shows that the change in *y* per increase in *x* is different than the change in *y* per the same unit decrease in *x* for any convex or concave function (Supplementary material Appendix 3 Fig. A4). Our analyses confirmed that, despite some systematic variation, the relationship between species richness and biomass is almost always concave down with positive values of *b* for the power function $0 < b < 1$. Thus, on average, species gains accrue less change in biomass than is lost when an equivalent number of species are lost.

There are several implications of Jensen's inequality that warrant consideration in future studies. First, Jensen's inequality suggests that efforts to conserve species could have disproportionately large effects on ecological functions compared to efforts to restore a similar number of species that have already been lost from a system. Indeed, restoration of ecosystem functions may require the addition of a larger number of native species than were originally lost. Jensen's inequality may also be important for assessing arguments in an ongoing debate about how local changes in species richness impact ecosystem functioning (Vellend et al. 2013, Gonzalez et al. 2016). Vellend et al. (2013) recently summarized time-series from 346 studies that had monitored plant species richness in 'local-scale' vegetation plots. These authors found that species richness has increased through time in roughly half the plots, but decreased through time in the other half. When averaged together, Vellend et al. concluded there has been no 'net' change in terrestrial plant species richness, and went on to argue that, if there has been no net change in species richness, then there cannot be changes in ecological function driven by local species loss.

Gonzalez et al. (2016) have criticized the study of Vellend on multiple grounds: 1) their dataset was not representative of global patterns of plant species richness nor the primary drivers of diversity change, 2) their dataset was unduly influenced by studies performed in ecosystems where biodiversity is likely recovering from historical destruction (e.g. forests recovering from logging), and 3) the unjustified logic of overextending conclusions of monitoring programs of biodiversity to ecological functions that were never measured in the original studies. Our results add a fourth issue to consider when interpreting the functional implications of changes in biodiversity in such syntheses as Vellend et al. 2013. Even if Vellend et al.'s primary conclusion that species gains and losses have averaged out to no 'net' loss of species richness is correct, it is still incorrect to suggest that no net loss of richness means there has been no net loss of ecological function. Assuming that species being gained and lost are, on average, functionally similar, our results suggest that losses could still have disproportionately large impacts on productivity compared to additions.

Strengths and limitations of the empirical relationship

The main insight supported by our analysis is that there is empirical evidence to support the use of a single value of *b* (in a power function) to describe how a change in species richness leads to a change in biomass for primary producers, but distinct values for aquatic herbivores and detritivores. The strength of this approach is a large database of experimental observations using similar experimental designs and a range of taxa (Table 1). Our analysis does not provide information about the other parameter in the power function, *a*, which can be thought of as the intercept of a linearized power function. Conceptually, *a* is the biomass of the average monoculture for a given community, and therefore will vary substantially among communities depending on their traits and environment. For example, *a* will be very low for phytoplankton and quite high for shrub assemblages, and with these community types will vary with resource availability, climate, etc. We did not have estimates for predictor variables that would be suitable to model variation in *a* among the very different study systems and environments included in this synthesis. For any specific application of the empirically estimated scaling exponent we provide here, the value of *a* will need to be estimated for the ecosystem under consideration.

The implications of our finding that a single value of *b* applies to most observations for application to other trophic groups or to support inferences about theoretical mechanisms remain limited for two reasons. First, some predictors in our analysis should be interpreted with caution. For example, studies differed widely in whether and how resources were controlled or manipulated. Thus, our predictor of 'resource level' is coarse and does not represent resource limitation in these systems. Resource manipulation (addition, control or reduction) was included in a plausible (but unlikely) model (Table 3). Based on the model ranking and the coarseness of the biological meaning of the resource predictor, we do not reject the hypothesis that resource supply can change the BEF relationship. Our analysis was limited by sufficient resource limitation data to conclusively test this hypothesis. Previous studies have shown mixed results, with some individual studies finding that increased nitrogen availability led to greater diversity effects on aboveground production (Reich et al. 2001, Fridley 2003) and a meta-analysis, using some of the same data as our study finding the opposite (Hooper et al. 2012). Some of this variability, and that found in our current study, could result from different effects on aboveground versus belowground versus total production, effects of different resources (e.g. CO_2 versus nutrients), different levels of resource addition, and compositional variation among communities (Reich et al. 2001, Fridley 2002, Hooper et al. 2012). More work is needed to fully test the dependence of the BEF relationship on resource supply, ideally with studies that factorially manipulate both diversity and resource supply within expected ranges of environmental change. Even with the heterogeneity in the data and the coarseness of some predictor levels, the patterns we observe are consistent with previous findings from analyses that tested a subset of these predictors on smaller datasets.

Finally, while not the objective of our study, the existence of a power law scaling relationship in other areas of ecology has been inferred to imply self-similar systems and a certain class of mechanistic driver (Brown et al. 2002). We caution that our empirical study, fitting a power function to BEF data, does not imply such a mechanistic driver to a BEF relationship. Determining whether such a relationship exists would merit further theoretical development, including assessing whether a power law is indeed the best descriptor of the BEF relationship. Instead, we aimed to test for a general empirical pattern. It remains to be determined whether there is a single best functional form to describe the BEF relationship, and whether this relationship is predicted or explained by any single theoretical framework.

Conclusion

Our analysis of the richness–biomass relationship allows practitioners to apply an empirically-derived, a priori prediction for the BEF relationship as a quantitative estimate for the expected importance of a change in biomass with a change in species richness. This estimate provides a starting hypothesis that investigators can use to determine whether additional factors modify the diversity–biomass relationship, or that they can attempt to falsify or improve upon. Furthermore, when biomass provides a good proxy for the processes and functions that underpin ecosystem services, this estimate of *b* could be used as a step in modeling the production of ecosystem services and their dependency on biodiversity. For instance, this BEF relationship can be part of an ecosystem service production function (Barbier 2007, Isbell et al. 2015), where production functions describe the relationship between various inputs (e.g. ecosystem properties, harvesting effort, etc.) and the level of a service that is produced (Barbier 2007). These production functions can support management decisions targeting provisioning of ecosystem services, such as by evaluating ecosystem service provisioning under different scenarios (Barbier 2007, Nelson et al. 2009, Tallis and Polasky 2009). Such an approach can also determine how different estimates of this parameter influence estimates of ecosystem service supply, and provide insight into the marginal value of maintaining diversity in terms of the value of an ecosystem service. However, for many ecosystem services, standing biomass is not a direct proxy for many ecosystem services that directly contribute to human well-being (for example, secondary productivity or nutrient cycling). Therefore, there is a need to determine whether this relationship holds more generally for other response variables that are also closely linked to human well-being (e.g. food production, water quality), and to what extent these findings extend to cases of non-random species loss. If so, integrating such a relationship into production functions could represent an important step towards the development of new tools to forecast the magnitude of change in important ecosystem services due to biodiversity loss, for a broader array of services. In the meantime, there is sufficient evidence to support the application of this parameterized power function to efforts such as integrated ecosystem function models or the generation of production functions linking biodiversity change to ecosystem functions and services directly related to biomass.

Acknowledgements – Funding – This work was conducted as a part of the Biodiversity and the Functioning of Ecosystems: Translating Model Experiments into Functional Reality Working Group supported by the National Center for Ecological Analysis and Synthesis, a Center funded by NSF (Grant no. EF-0553768), the University of California, Santa Barbara, and the State of California. MO is supported by an NSERC discovery grant and the Alfred P. Sloan Foundation. AG is supported by an NSERC discovery grant, the Canada Research Chair Program and the Quebec Centre for Biodiversity Science.
Statement of authorship (following ICMJE guidelines): MO and AG designed this work and drafted the manuscript, MO conducted the analysis, BC and KD acquired the data, all authors contributed to interpretation of the analysis and implications, revision of intellectually important content, and approved of its final publication and have agreed to be accountable for the work.

References

Balvanera, P. et al. 2006. Quantifying the evidence for biodiversity effects on ecosystem functioning and services. – Ecol. Lett. 9: 1–11.

Barbier, E. B. 2007. Valuing ecosystem services as productive inputs. – Econ. Policy 22: 177–229.

Bolker, B. M. et al. 2009. Generalized linear mixed models: a practical guide for ecology and evolution. – Trends Ecol. Evol. 24: 127–135.

Boyer, K. E. et al. 2009. Biodiversity effects on productivity and stability of marine macroalgal communities: the role of environmental context. – Oikos 118: 1062–1072.

Brown, J. H. and West, G. 2000. Scaling in biology (Brown, J. H. and West, G. B. eds.). – Oxford Univ. Press.

Burnham, K. P. and Anderson, D. R. 2002. Model selection and multimodel inference. – Springer.

Byrnes, J. E. et al. 2007. Invasions and extinctions reshape coastal marine food webs. – PLoS ONE 3: e295.

Cardinale, B. J. et al. 2006. Effects of biodiversity on the functioning of trophic groups and ecosystems. – Nature 443: 989–992.

Cardinale, B. J. et al. 2007. Impacts of plant diversity on biomass production increase through time because of species complementarity. – Proc. Natl Acad. Sci. USA 104: 18123–18128.

Cardinale, B. J. et al. 2011. The functional role of producer diversity in ecosystems. – Am. J. Bot. 98: 572–592.

Cressie, N. et al. 2009. Accounting for uncertainty in ecological analysis: the strengths and limitations of hierarchical statistical modeling. – Ecol. Appl. 19: 553–570.

Duffy, J. E. 2002. Biodiversity and ecosystem function: the consumer connection. – Oikos 99: 201–219.

Duffy, J. E. 2003. Biodiversity loss, trophic skew and ecosystem functioning. – Ecol. Lett. 6: 680–687.

Duffy, J. E. et al. 2005. Ecosystem consequences of diversity depend on food chain length in estuarine vegetation. – Ecol. Lett. 8: 301–309.

Fridley, J. D. 2002. Resource availability dominates and alters the relationship between species diversity and ecosystem productivity in experimental plant communities. – Oecologia 132: 271–277.

Fridley, J. D. 2003. Diversity effects on production in different light and fertility environments: an experiment with communities of annual plants. – J. Ecol. 91: 396–406.

Gamfeldt, L. et al. 2015. Marine biodiversity and ecosystem functioning: what's known and what's next? – Oikos 124: 252–265.

Gelman, A. and Hill, J. 2007. Data analysis using regression and multilevel/hierarchical models. – Cambridge Univ. Press.

Gonzalez, A. et al. 2016. Estimating local biodiversity change: a critique of papers claiming no net loss in local diversity. – Ecology 97: 1949–1960.

Griffin, J. N. et al. 2013. Effects of predator richness on prey suppression: a meta-analysis. – Ecology 94: 2180–2187.

Hector, A. et al. 1999. Plant diversity and productivity experiments in European grasslands. – Science 286: 1123–1127.

Hooper, D. U. et al. 2012. A global synthesis reveals biodiversity loss as a major driver of ecosystem change. – Nature 486: 105–108.

Hudson, L. N. et al. 2013. The relationship between body mass and field metabolic rate among individual birds and mammals. – J. Anim. Ecol. 82: 1009–1020.

Isbell, F. et al. 2015. The biodiversity-dependent ecosystem service debt. – Ecol. Lett. 18: 119–134.

Lefcheck, J. S. et al. 2015. Biodiversity enhances ecosystem multifunctionality across trophic levels and habitats. – Nat. Commun. 6: 6936.

Myers, R. A. and Worm, B. 2003. Rapid worldwide depletion of predatory fish communities. – Nature 423: 280

Nelson, E. et al. 2009. Modeling multiple ecosystem services, biodiversity conservation, commodity production, and tradeoffs at landscape scales. – Front. Ecol. Environ. 7: 4–11.

O'Connor, M. I. et al. 2007. Temperature control of larval dispersal and implications for marine ecology, evolution and conservation. – Proc. Natl Acad. Sci. USA 104: 1266–1271.

Petchey, O. L. 2000. Species diversity, species extinction and ecosystem function. – Am. Nat. 155: 696–702.

Peters, R. H. 1983. The ecological implications of body size. – Cambridge Univ. Press.

Pinheiro, J. and Bates, D. 2000. Mixed-effects models in S and S-PLUS. – Springer.

Reich, P. B. et al. 2001. Plant diversity enhances ecosystem responses to elevated CO_2 and nitrogen deposition. – Nature 410: 809–812.

Reich, P. B. et al. 2012. Impacts of biodiversity loss escalate through time as redundancy fades. – Science 336: 589–592.

Replansky, T. and Bell, G. 2009. The relationship between environmental complexity, species diversity and productivity in a natural reconstructed yeast community. – Oikos 118: 233–239.

Richards, S. A. 2005. Testing ecological theory using the information-theoretic approach: examples and cautionary results. – Ecology 86: 2805–2814.

Setälä H. and McLean, M. A. 2004. Decomposition rate of organic substrates in relation to the species diversity of soil saprophytic fungi. – Oecologia 139: 98–107.

Snijders, T. A. B. and Bosker, R. J. 1999. Multilevel analysis: an introduction to basic and advanced multilevel modeling. – Sage Publishers.

Srivastava, D. S. et al. 2009. Diversity has stronger top–down than bottom–up effects on decomposition. – Ecology 90: 1073–1083.

Stachowicz, J. J. et al. 2008. Diversity enhances cover and stability of seaweed assemblages: the role of heterogeneity and time. – Ecology 89: 3008–3019.

Tallis, H. and Polasky, S. 2009. Mapping and valuing ecosystem services as an approach for conservation and natural-resource management. – Ann. N. Y. Acad. Sci. 1162: 265–283.

Tilman, D. et al. 1997. Plant diversity and ecosystem productivity: theoretical considerations. – Proc. Natl Acad. Sci. USA 94: 1857–1861.

Tilman, D. et al. 2001. Diversity and productivity in a long-term grassland experiment. – Science 294: 843–845.

Turnbull, L. A. et al. 2013. Coexistence, niches and biodiversity effects on ecosystem functioning. – Ecol. Lett. 16: 116–127.

Vellend, M. et al. 2013. Global meta-analysis reveals no net change in local-scale plant biodiversity over time. – Proc. Natl Acad. Sci. USA 110: 19456–19459.

Verbeke, G. and Molenberghs, G. 2000. Linear mixed models for longitudinal data. – Springer.

Supplementary material (available online as Appendix oik-03652 at <www.oikosjournal.org/appendix/oik-03652>). Appendix 1–3.