Paper:

O'Connor, M., Gonzalez, A., Byrnes, J., Cardinale, B., Duffy, J., Gamfeldt, L., Griffin, J., Hooper, D., Hungate, B., et. al. (2017). A general biodiversity-function relationship is mediated by trophic level. *Oikos, 126*(1), 18-31.
http://dx.doi.org/10.1111/oik.03652

_____

A general biodiversity–function relationship is mediated by trophic level

Mary I. O'Connor[1], Andrew Gonzalez[2], Jarrett E. K. Byrnes[3], Bradley J. Cardinale[4], J. Emmett Duffy[5], Lars Gamfeldt[6], John N Griffin[7], David Hooper[8], Bruce A. Hungate[9], Alain Paquette[10], Patrick L. Thompson[2], Laura E. Dee[12], Kristin Matulich[13]

[1]*Department of Zoology and Biodiversity Research Centre, University of British Columbia, Vancouver, BC Canada*

[2]*Department of Biology, McGill University, Montreal, QC Canada*

[3]*Department of Biology, University of Massachusetts Boston, Boston, MA 02125*

[4]*School of Natural Resources & Environment, University of Michigan, Ann Arbor, MI 48109*

[5] *Tennenbaum Marine Observatories Network, Smithsonian Institution, Washington DC, USA*

[6]*Department of biological and environmental sciences, University of Gothenburg, Box 461, SE-405 30 Göteborg, Sweden*

[7]*Department of Biosciences, Swansea University, Singleton Park, Swansea, SA2 8PP, UK*

[8]*Western Washington University, Department of Biology, Bellingham, WA, U.S.A.*

[9]*Center for Ecosystem Science and Society, Department of Biological Sciences, Northern Arizona University, Flagstaff AZ, USA*

[12]*Bren School of Environmental Science & Management, University of California at Santa Barbara, Santa Barbara, CA, USA 93117*

[10] *Centre for Forest Research, Université du Québec à Montréal, P.O. Box 8888, Centre-ville Station, Montréal, QC H3C 3P8 Canada*

Abstract

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 ecosystem functioning. We used hierarchical mixed effects models, based on a power function between species richness and biomass production ($Y = a*S^b$), and a database of 374 published experiments to estimate the BEF relationship (the change in function 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; b was nearly twice as high for consumers (herbivores and detritivores) compared to primary producers in aquatic systems; in terrestrial ecosystems, b deviated from the overall pattern for terrestrial detritivores. We estimated changes in biomass expected for a range of changes in species richness, highlighting that species loss has greater functional 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

45   of ecosystem services and their dependence on biodiversity.

Introduction

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

50   biomass and resource use (Tilman et al. 2001, Cardinale et al. 2006, Reich et al. 2012). A

nonlinear function commonly captures the relationship between species richness and ecosystem

function, and its prevalence among experimental results in the literature 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

55   predictions across complex systems (Peters 1983, Brown and West 2000, O'Connor et al. 2007).

Here we provide evidence that a general empirical relationship between diversity and function

depends on trophic level, but is otherwise conserved across a large database of BEF experiments.

        To extend the biodiversity-ecosystem functioning paradigm to contexts beyond

controlled experiments, we need (1) a quantitative estimate of how much function is lost with the

60   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 (e.g., species

traits, resource supply, ecosystem) and factors not yet identified. With existing data, an empirical

estimate of the relationship between richness and biomass could be applied to biodiversity

change models to give a first approximation, or testable prediction, for effects of biodiversity

65   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 (e.g., Balvanera et al. 2006, Cardinale et al. 2006, Reich et al. 2012). For competitively structured communities, this relationship has been suggested to follow the Michealis-Menten function;

70   however in many experiments, the saturation of function with accumulating species is not clear at the levels of species richness nested (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*(richness)^b* with scaling parameter *b* (Cardinale et al. 2007, 2011, Reich et al. 2012). The power function used here is not meant to imply a particular

75   theoretical BEF mechanism. It is used because it has substantial empirical support from previous syntheses (Cardinale et al 2011, Gamfeldt et al. 2015, Lefcheck et al 2015).

    Previous grassland diversity experiments reported a central tendency toward a value of approximately *b* = 0.26. These studies have identified variation in this estimate: 95% confidence intervals of 0.15 - 0.32 (Cardinale et al. 2006), or a standard deviation of 0.27 (Cardinale et al.

80   2011). Whether that variation reflects systematic and ecologically important differences among BEF relationships has not been clear. Individual experimental studies suggest this relationship can vary with resource availability (e.g., water, nutrients, CO_2) (Reich et al. 2001, Fridley 2002, Boyer et al. 2009), presence of a predator (Duffy et al. 2005), and can increase in strength over time as experiments are run for more experimental years (Stachowicz et al. 2008, Reich et al.

85   2012, Cardinale et al. 2007). Within experiments that share a species pool, experimental design, and other factors, the BEF relationship varies 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 environmental condition (e.g., soil type, climate, etc). Meta-analyses of dozens of experiments have demonstrated that, across studies, estimated mean richness-standing

90 biomass scaling values (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.

95 2011). Recently, three meta-analyses reported differences in BEF effects among trophic levels

demonstrating that marine herbivore richness has a stronger effect on function than algae

(Gamfeldt et al. 2014), aquatic herbivore richness has stronger effects on function than plants

when multiple functions are analyzed (Lefcheck et al., 2015), and carnivore richness more

strongly affects resource depletion than richness at lower trophic levels (Griffin et al. 2013). If

100 the BEF relationship does vary in space and time, or with biotic and abiotic conditions, then

estimates of the scaling parameter alone may impart minimal information; however, if it is

general or varies predictably it can provide a powerful tool for efforts to generalize the

consequences of species loss for ecological function and ecosystem services (Isbell et al. 2014).

While individual experiments, and broader meta-analyses, have tested the importance of one or a

105 few additional factors (time, resource supply, trophic structure, etc.; (Hooper et al. 2012, Tilman

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

110 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. We then test biological and experimental conditions that

might explain variation in this relationship, aiming to identify which factors are essential to understanding the richness-biomass scaling relationship and which might be left out of a general

115   model. We apply 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 scaling relationship (a specific value) between richness and the important ecological function of community biomass production.

120

2. Methods

We used a hierarchical mixed effects model to test our hypothesis that there is a constant scaling 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

125   experimental tests of the relationship between community biomass and richness are published. Standing biomass is often inferred to be correlated with net primary production, though 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

130   studies included 374 experiments and 558 entries (from 91 studies published between 1985 and 2009, appendix) 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) reported across a set of species richness (S) treatments varying in the number of species (at least 2 richness levels), with all other factors

135   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,

140   5) whether experimental units 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 = 1 species),

such that 544 of 558 entries included $S = 1$, and 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

145   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 were 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

150   does not include the following combinations: aquatic species richness x 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

155   information to the overall analysis (Gelman and Hill 2007).

2.1 The search for a single BEF relationship

Our primary objective was to estimate the scaling relationship between richness and biomass. Then, we aimed to test whether a single scaling coefficient describes the relationship

160 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.

165 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 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

170 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.

175 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.

We began by modeling the BEF relationship at the finest data resolution with the simplest

180   plausible relationship of interest, in our case biomass (ln(Y)) at the plot level as predicted by

species richness (ln(S)),

$$\ln(Y_{ijkl}) = B_{0.ijk} + B_{1.ijk}*\ln(S_{ijkl}) + \varepsilon_{ijkl,} \qquad \textbf{Eqn 1a}$$

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

185   assumed normally distributed residual error ($\varepsilon_{ijk} \sim (N, \sigma^2)$). We modeled the biomass-richness

relationship as a log-log relationship, derived by log-transforming the power function $Y = a*S^b$,

so that *a* is estimated by B_0 and the scaling parameter *b* is estimated by B_1. Although other

formulations (e.g., Michaelis-Menten) have been used to describe this relationship (e.g.,

Cardinale et al., 2011), we proceed with the power function, which has received substantial

190   empirical support (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, 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 covary in some cases with

195   variation in the slope. 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 ($B_{0.ijk}$):

$$B_{0.ijk} = \gamma_{00} + \mu_{0.i} + \mu_{0.j} + \mu_{0.k} \qquad \textbf{Eqn 1b}$$

$$B_{1.ijk} = \gamma_{10} + \mu_{1.i} + \mu_{1.j} + \mu_{1.k}$$

200   In Eqns 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), and 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 (Appendix A).

205   The test of our first hypotheses is whether variable slopes ($\mu_{1.i}$, $\mu_{1.j}$, $\mu_{1.k}$) are required. If so, we would conclude that it is not possible to identify a single scaling parameter ($B_{1.ijk}$) for this dataset. We also tested alternate models (Eqns 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 T_G = duration of experiment in days / generation time of focal organism) (B_2), and the

210   interaction between T_G and ln(S) (B_3):

$$\ln(Y_{ijkl}) = B_{0.ijk} + B_{1.ijk}*\ln(S_{ijkl}) + B_2*\ln(T_{G.ijkl}) + \varepsilon_{ijkl.} \qquad \textbf{Eqn 2}$$

$$\ln(Y_{ijkl}) = B_{0.ijk} + B_{1.ijk}*\ln(S_{ijkl}) + B_2*\ln(T_{G.ijkl}) + B_3*\ln(S_{ijkl})*\ln(T_{G.ijkl}) + \varepsilon_{ijkl.} \qquad \textbf{Eqn 3}$$

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.

215

2.2 Testing hypotheses about factors that modify the BEF scaling relationship

We tested our second main hypothesis that entry-, experiment- or study-level factors alter estimates of the richness-biomass scaling relationship. Specifically, we compared mixed effects models with different formulations that represent hypotheses for how various biotic and abiotic

220   factors (listed in Table 1) interact with species richness to affect the scaling relationship.

In addition to the basic hypothesis that biomass changes with increasing species richness and time (models 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). The trophic group predictor indicates the group for which species

225 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, CO_2) modified the BEF relationship (model 5).

The three categorical levels of the resource treatment predictor (control, addition, reduction)

230 reflect experimental manipulations relative to ambient conditions. 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. 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

235 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). We considered time in two ways. First, we modeled the effect

240 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 level-1 parameter T_G, and models the effect of year (or day)

within a multi-year (day) experiment (Table 2). Second, we examined the effect of time by

245 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).

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 vs field (model 8). Finally, we tested the hypothesis that all factors modify the BEF relationship (model 9, Eqn 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:

$$B_{0.ijk} = \gamma_{00} + \gamma_{01}*Sys_i + \gamma_{02}*TG_i + \gamma_{03}*Sys_i*TG_i + \gamma_{04}*Units_j + \gamma_{05}*LabField_i$$

$$+ \gamma_{06}*S_{max.il} + \gamma_{07}*N_j + \gamma_{08}*ln(max(Duration_i)) + \mu_{0.i} + \mu_{0.j} + \mu_{0.l} \qquad \textbf{Eqn 4a}$$

$$B_{1.ijk} = \gamma_{10} + \gamma_{11}*Sys_i + \gamma_{12}*TG_i + \gamma_{13}*Sys_i*TG_i + \gamma_{14}*Units_j + \gamma_{15}*LabField_i$$

$$+ \gamma_{16}*S_{max.il} + \gamma_{17}*N_j + \gamma_{18}*ln(max(Duration_i)) + \mu_{1.i} + \mu_{1.j} + \mu_{1.k} \qquad \textbf{Eqn 4b}$$

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

2.3 Model selection, analysis and inference

To identify the best model, we first determined the need for variable slopes and intercepts for each candidate model (Eqn 1a, 2 and 3) by comparing models with different random effects structures (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 AICc adjusted for degrees of freedom to account for different random effects following Bolker et al.

(2009) and Gelman and Hill (2007), and δ_{aic} 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

275   (Verbeke and Molenberghs 2000).

To test our second hypothesis, we compared models with biotic and experimental predictors (models 4-9; Eqns 4a-b). We ranked models using AIC_c, and compared them with δ_{aic} and Akaike weights (w). We defined the best model set as all models with $\delta_{aic} < 2$ (Richards 2005), Burnham and Anderson 2002). If more than one model met our criteria of $\delta_{aic} < 2$, we

280   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 scaling 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.

285   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 Equation 1 revealed 7 experiments from 2 studies that were extreme outliers in the dataset ($< 0.03\%$ of entries), and these were excluded from analysis to meet assumptions of homoscedasticity (Appendix 1).

290   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

13

295   ln(T_G)*ln(S) fixed effect is expected to include any real effects and any possible autocorrelation.

All analyses were done in R (v. 3.2.1). We used lme4 package (v. 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 is available from the authors and at

https://github.com/mioconnor78/OConnor-et-al-BEF-Relationship. We produced caterpillar plots

300   using the package sjPlot (v. 1.8.2).

3. Results

A single, universal scaling relationship ($B_{1.ijk}$ value) was not supported by our analysis of the

species richness-biomass relationship. Variable slopes and intercepts associated with entry,

305   experiment and study were required for each candidate model (Eqns 1-3) (AIC_C > 100 for

comparison of model with variable slopes and intercepts to model with fewer random effects

terms) (Table A1). We did not find strong evidence for a systematic effect of experimental

duration on the BEF relationship across all studies (Table 2). This simple model with random

effects (Model 2; Eqn 2) estimates a BEF relationship of $b = \gamma_{10} = 0.23$ (95% CI: 0.18, 0.28) that

310   applies to most but not all entries (Figure 1A, 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 (Figure 2), suggests this estimate of $b = \gamma_{10}$ adequately describes most

observations (i.e., the confidence intervals for the random effects include 0 in the caterpillar plots

for most ln(S) estimates) (Figure A1). Still, the number of slope residuals deviating from the

315   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 ln(8) to minimize covariation of random effects to 0.07 for Entry, 0.24 for

Experiment, and -0.16 for Study level slopes and intercepts estimated by model 1.

      After concluding that variable slopes at the entry, experiment and study levels confound

320   the identification of a single scaling relationship, we tested our second main hypothesis that

entry-, experiment-, and study-level factors could explain some variation in richness-biomass

scaling relationship, thus eliminating the need for variable slopes ($\mu_{1.i}$, $\mu_{1.j}$ and $\mu_{0.k}$). We found

that the BEF relationship varies systematically between aquatic primary producers and

consumers such that herbivore and detritivore biomass increased with species richness by

325    $b_{aq.herbivores} = \gamma_{10} + \gamma_{12} = 0.47$, 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 3, Figure 1B).

Though there was no difference between aquatic and terrestrial primary producers, terrestrial

detritivores had a much weaker and negative relationship between richness and biomass than all

330   other groups (aquatic primary producers and herbivores, and terrestrial plants) (Figure 1B, C).

The top-ranked model of our set was model 4 (Table 3), which included the interaction between

trophic group and ecosystem (Figure 1B). None of our other hypotheses about variation in the

BEF relationship could be considered as equivalent to this 'best' model.

      The best model indicates that variable slopes and intercepts are required, even with fixed

335   effects for trophic group and ecosystem (Model 4, Table A2, Figure 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) (Figure 2). Still, the value estimated by model 4 of $b =$

$\gamma_{10} = 0.26$ (95% CI: 0.16, 0.37) describes the majority of the dataset (Figure 2). The larger

15

variance components associated with study and entry compared to experiment suggests that most
of the unexplained systematic variation is at those levels.

Our model comparison results allowed us to reject some of our alternate hypotheses
(Table 3). We rejected the hypothesis that the basic model (Eqn 2) is sufficient to explain the
relationship between richness and biomass. We also rejected the hypotheses that differences in
experimental designs (number of species tested, lab vs field experiment, and the method of
estimating biomass) explain variation in the observed BEF relationship. Results of the test of the
importance of (no. generations, T_G) within experiments did not support an interaction between
$\ln(S)$ and $\ln(T_G)$ (Table 2). AICc values and the likelihood ratio test suggest that the model with
and without the interaction term are equivalent, and we therefore proceeded with the simpler
model without the interaction between time and richness (model 2) inferring that the richness-
biomass relationship did not depend on the duration of an experiment relative to the generation
time of the organisms being studied (e.g., no. generations, T_G) across the 374 experiments in our
dataset (Table 2). We conducted two additional tests of the hypothesis that experimental duration
might affect the strength of the relationship. In the first of these, we tested the effect of
maximum duration on the scaling relationship for only final observations of each experiment.
Model comparisons for this dataset were consistent with the full dataset, and suggested no effect
of maximum duration on the scaling parameter (Table A3). Second, we expanded model 3 (Table
2) 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). The inclusion of the
interaction term for time in this model suggests an effect of time could be informative, yet
coefficients for the T_G interactions do not differ from 0 except for herbivores, which suggests a

16

very weak negative effect of time on the BEF relationship. We did not find evidence for this effect of time on plants or detritivores across all studies (Tables S4 and S5).

4. Discussion

365   We used 558 experimental diversity manipulations to quantify the richness-biomass relationship, and found that most (but not all) primary producer assemblages conform to a single scaling coefficient ($b = 0.26$) for how biomass increases with increasing species richness, with evidence for a stronger relationship among aquatic consumers (herbivores, $b = 0.47$; detritivores, $b = 0.54$) compared to plants and algae. We conclude that in addition to a change in species richness,

370   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 provides 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

375   hundreds of experimental tests. Our failure to reject the need for variable slopes among groups (entries, experiments, and studies) indicates that systematic variation in the BEF relationship exists among studies and experiments, not captured by our hypotheses.

One application of a parameterized BEF relationship is to estimate the change in function (biomass) associated with a change in species richness. The identification of empirically

380   supported scaling parameters for the power function allow the expectation that not only does species loss or gain *per se* affect standing biomass, but also to estimate how the proportional loss of species translates to a specific expected loss in standing biomass. Our results suggest that a proportional change in species richness has greater effects on standing biomass for aquatic

herbivores and detritivores than for primary producers, and possibly a negative effect for

385 terrestrial detritivores (Figure 1C). The nonlinearity of the BEF relationship informs

interpretation of this relationship. For a distribution of possible changes in species richness

(Figure 3), Jensen's inequality implies that for positive values of *b*, the mean change in species

richness will always exceed the associated change in biomass expected from this change in

species number (not considering other dimensions of biodiversity, non-random species loss, or

390 co-varying environmental conditions). This is due to the decelerating nature of the BEF curve;

gains accrue less change in function than is lost when an equivalent number of species are lost

(Figure A3). Thus, even for a distribution in observed species richness changes centered on 0,

we'd expect an associated distribution in biomass change with a mean value less than 0 (Figure

3). This highlights the fact that species losses have greater effects on function than gains, given

395 values of *b* less than one, and large proportional losses are expected to lead to disproportionately

large losses in biomass. This is particularly true for systems initially low in species richness for

which the loss of even a few species amounts to a large proportional loss.

The larger effects of species richness change for aquatic herbivore and detritivore

biomass than primary producer biomass lead to the hypothesis that changes in diversity could

400 create positive feedbacks in aquatic systems if as species richness declines, the magnitude of the

consequences increase. If diversity declines occur more among consumers than resources

(Byrnes et al 2007, Duffy 2003), then the BEF relationship could also shift the importance of top

down control as grazer species loss disproportionately reduces grazer biomass, potentially

releasing plant assemblages from grazing pressure, and indirectly increasing plant productivity.

405 The observed stronger BEF scaling relationship in primary consumers relative to primary

producers has been predicted conceptually (Duffy 2002, 2003). Although early syntheses did not

detect this difference, that could be explained by smaller numbers of studies and relatively

simplistic 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

410   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

415   that herbivores did not differ from plants can now be revised.

We did not find a systematic relationship between the BEF relationship and experimental

duration (T_G) across this dataset (Table 2). 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

420   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 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 scaling

coefficient were required, and one of the main differences between entries within an experiment

425   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

430   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.

Among-study variation explained the majority of the variation in the random effects in

our model (Table 4). Random, study-level variation is distinguishable from residual variation

435   (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. Therefore, our model has captured sufficient variation to provide an estimate of

440   how biomass scales with species richness and to apply the model to similar systems.

Strengths and limitations of the empirical scaling 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

445   leads to a change in biomass for primary producers, and distinct values for aquatic herbivores

and detritivores. The strength of this approach is a large database of experimental observations

using the same experimental design and a range of taxa (Table 1). Based on this, we have

confidence that the scaling value is not highly dependent on taxonomic differences, or on

differences associated with researchers, studies, or experimental settings.

450   The implications of this finding for other trophic groups or for 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 scaling 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.

The interpretation of the trophic level effect is complicated by biases in the available data. For example, this dataset includes no terrestrial herbivore manipulations for which herbivore biomass was reported, and we therefore have no basis for inference about scaling relationships for terrestrial herbivores. Similarly, we lacked sufficient estimates of carnivore biodiversity manipulations that reported effects on carnivore biomass to include them in this analysis.

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

480   relationship. To determine whether evidence of such a relationship exists would merit further

theoretical development requires first determining 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

485   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

490   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.

Furthermore, when biomass provides a good proxy for the processes and functions that underpin

ecosystem services, this scaling relationship could be used as a step in modeling the production

495   of ecosystem services and their dependency on biodiversity. For instance, this BEF relationship

can be part of an ecosystem service production function (Isbell et al. 2014), 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

500   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 scaling 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 services, standing biomass is not a direct proxy of many ecosystem services

505   that directly contributes to human well-being (for example, secondary productivity or nutrient

cycling). Therefore, there is a need to determine whether this scaling relationship holds more

generally to other response variables that are also closely linked to human well-being (e.g., food

production, water quality). If so, integrating such a scaling relationship into production functions

could represent an important step towards the development of new tools to forecast the

510   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.

515

23

Authorship statement, following ICMJE guidelines:

MO and AG designed this work and drafted the manuscript, MO conducted the analysis, BC and

KM acquired the data, all authors contributed to interpretation of the analysis and implications,

525   revision of intellectually important content, and approved of its final publication and have agreed

to be accountable for the work.

530 **References**:

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

Balvanera, P. et al. 2014. Linking biodiversity and ecosystem services: assessing uncertainties and finding a way forward. - Front. Ecol. Environ. 64: 49–57.

535 Barbier, Edward B. "Valuing ecosystem services as productive inputs. "*Economic Policy* 22.49 (2007): 178-229.

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

540 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 G. B. West (eds). 2000. Scaling in Biology. Oxford University Press.

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

545 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. U. S. A. 104: 18123–18128.

550 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:
555 201–219.

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.

560 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.

565 Gamfeldt, Lars, J. S. Lefcheck, J.EK Byrnes, B. J. Cardinale, J. E. Duffy, and J. N. Griffin. "Marine biodiversity and ecosystem functioning: what's known and what's next?." *Oikos* 124, no. 3 (2015): 252-265.

Gelman, A. and Hill, J. 2007. Data analysis using regression and multilevel/hierarchical models.
570   - Cambridge University Press.

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.

575 Hooper, D. U. et al. 2012. A global synthesis reveals biodiversity loss as a major driver of ecosystem change. - Nature 486: 105–U129.

Isbell, F., Tilman D., S. Polasky, and M. Loreau. 2015. The biodiversity-dependent ecosystem service debt. Ecology Letters 18(2): 119-134.

Lefcheck, Jonathan S., Jarrett EK Byrnes, Forest Isbell, Lars Gamfeldt, John N. Griffin, Nico
580   Eisenhauer, Marc JS Hensel, Andy Hector, Bradley J. Cardinale, and J. Emmett Duffy. "Biodiversity enhances ecosystem multifunctionality across trophic levels and habitats." *Nature communications* 6 (2015).

Nelson, E. et al. 2009. Modeling multiple ecosystem services, biodiversity conservation,
585   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. 104: 1266–1271.

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

590 Peters, R. H. 1983. The ecological implications of body size. - Cambridge University Press.

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

Reich, P. B. et al. 2001. Plant diversity enhances ecosystem responses to elevated CO2 and nitrogen deposition. - Nature 410: 809–812.

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

26

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

Snijders, T. A. B. and Bosker, R. J. 1999. Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling. - Sage Publishers.

600   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. - In: The Year in Ecology and Conservation Biology 2009. in press.

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

Tilman, D. et al. 2001. Diversity and Productivity in a Long-term Grassland Experiment. - Science 294: 843–845.

Tilman, D. et al. 2012. Biodiversity impacts ecosystem productivity as much as resources,
610   disturbance, or herbivory. - Proc. Natl. Acad. Sci. U. S. A. 109: 10394–10397.

Turnbull, L. A. et al. 2013. Coexistence, niches and biodiversity effects on ecosystem functioning (JHR Lambers, Ed.). - Ecol. Lett. 16: 116–127.

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

615

620

625

630

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 (TG)	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

Table 2. Results of model selection for Level-1 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(Tg)). Models are ranked by AICc, and compared using AIC weight (*w*) and 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(Tg)**	**1451.5**	**0.49**	**13**	**-712.67**	**0.00**	**--**
3	ln(Y) = ln(S)*ln(Tg)	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

Table 3. Comparison of alternative level-2 models for how richness (ln(S)) affects biomass. Model terms are

655   Equations 4a and b (intercept term not shown here), ranked from left to right by their quality (high to low) as a d

dataset. Models differed in fixed effects, indicated by ✓, but all included variable intercepts ($\mu_{0i}, \mu_{0j}, \mu_{0k}$) and slo

level of the study (i), experiment (j) and entry (k). Models were ranked based on AICc, and differences assessed

Akaike weights (w) and likelihood ratio tests. We used likelihood ratio test results (- $P > 0.05$, * $P < 0.05$, ** P

models with the top-ranked (lowest AICc) model only for comparisons in which one model can be derived from

660   constraining parameter values. When this was not possible, likelihood ratio tests were not performed. A significa

that the model with the lower AICc value is a better description of the data. When the likelihood ratio test indica

model with fewer parameters is preferred.

665

Predictor	Term	Model							
		4	**5**	**6**	**4.2**	**9**	**2**	**9.1**	**3**
ln(S)	γ_{10}	✓	✓	✓	✓	✓	✓	✓	✓
Time (ln(T_G))	B_2	✓	✓	✓	✓	✓	✓	✓	✓
Ecosystem (Sys)	γ_{01}	✓	✓	✓	✓	✓	□	✓	
Trophic group (TG)	γ_{02}	✓	✓	✓	✓	✓	□	✓	
Sys*TG	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)*TG	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
	AIC$_c$	1437.8	1442.9	1443.1	1443.1	1443.5	1451.5	1452.3	1452.4
	δ	0	5.13	5.33	5.34	5.70	13.72	14.52	14.67
	w	0.784	0.060	0.054	0.054	0.045	0.001	0.001	0.001
	logLik	-697.7	-696.2	-694.3	-702.4	-686.2	-712.7	-692.7	-712.1
	P		-	-	**	-	**		**

Table 4. Coefficients for modeled effect of richness on standing stock. Mean ($\pm$95% CI) estimate from the best model with fixed effects (trophic level, duration, lab vs 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

670 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]
	TG - Herbivore	γ_{02}	-0.33 [-1.38, 0.66]
	TG - 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)*TG – Herbivore	$\boldsymbol{\gamma_{12}}$	**0.21 [0.03, 0.38]**
	ln(S)*TG – 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

675

Figure 1. A) Standing stock (biomass) plotted against species richness as a power function (Y = a*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 (Eqn 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.

A)
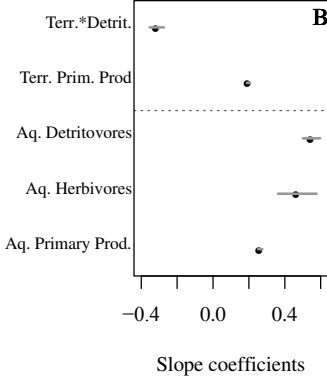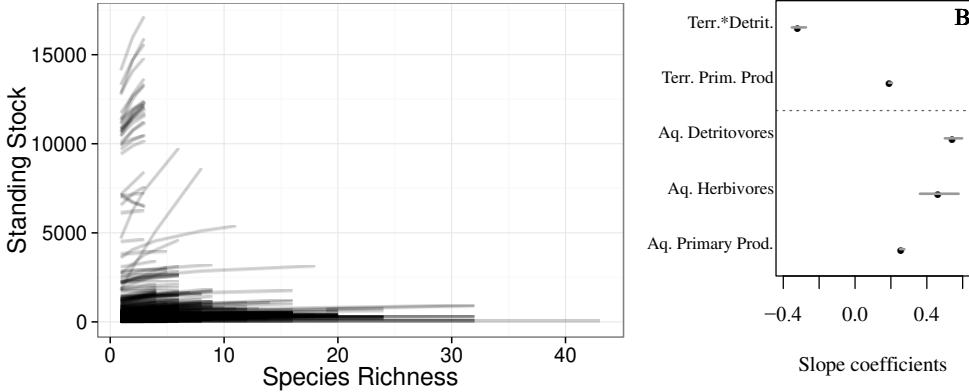
690
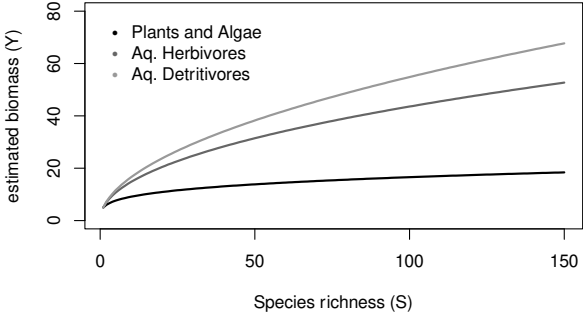
Figure 2. Random effects ($\pm$ CI) estimated by the best level-1 model (model 4, Table 2) associated with entry$_k$ (plots A-B), experiment$_j$ (panels C-D) and study$_i$ (panels E-F) for intercepts (μ_0) and slopes (μ_1), ranked by slope random effects (μ_1). Gray CI's include 0, indicating that the 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.

695

700   A) Intercept random effects for Entry (μ_{0k})   B) Slope random effect for Entry (μ_{1k})
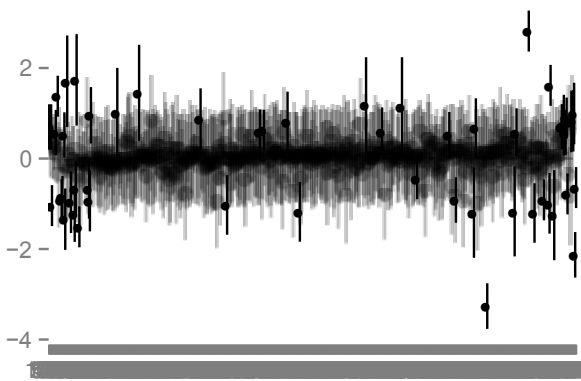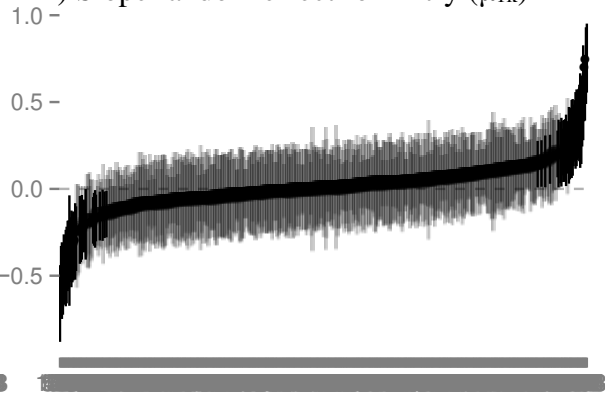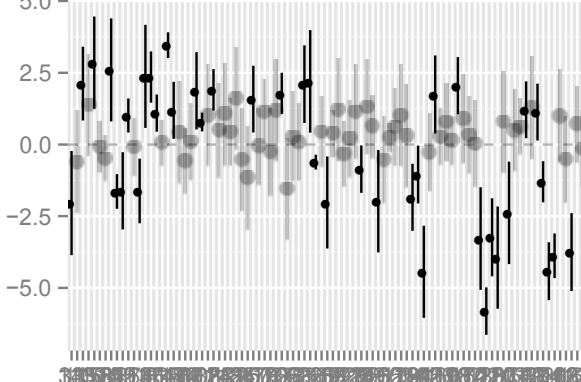
C) Intercept random effects for Experiment (μ_{0j})   D) Slope random effect for Experiment (μ_{1j})
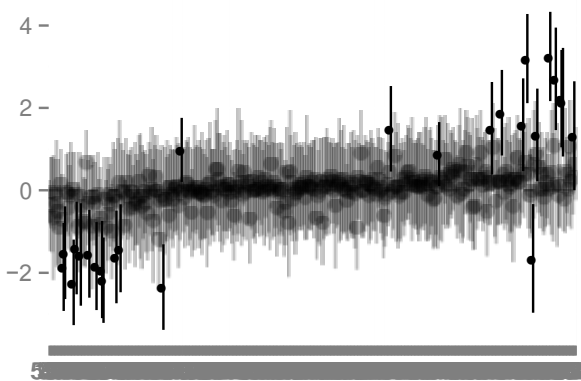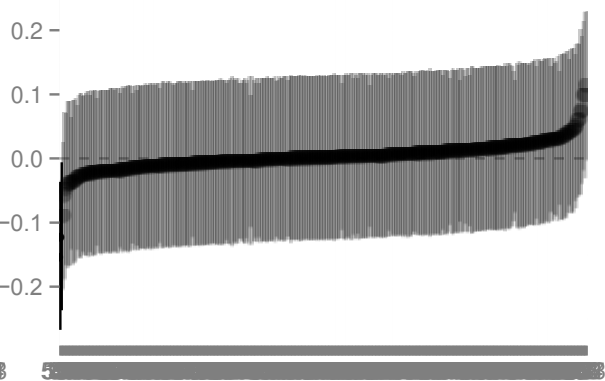
705   E) Intercept random effects for Study (μ_{0i})   F) Slope random effect for Study (μ_{1i})
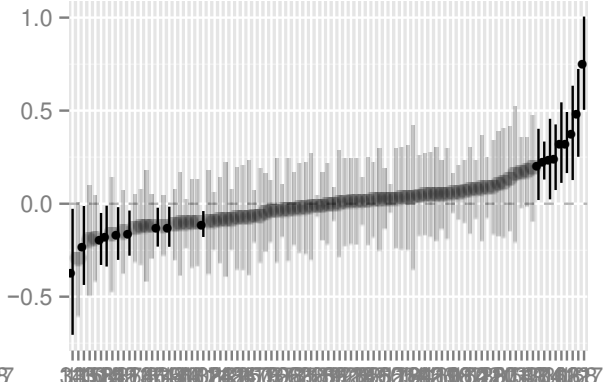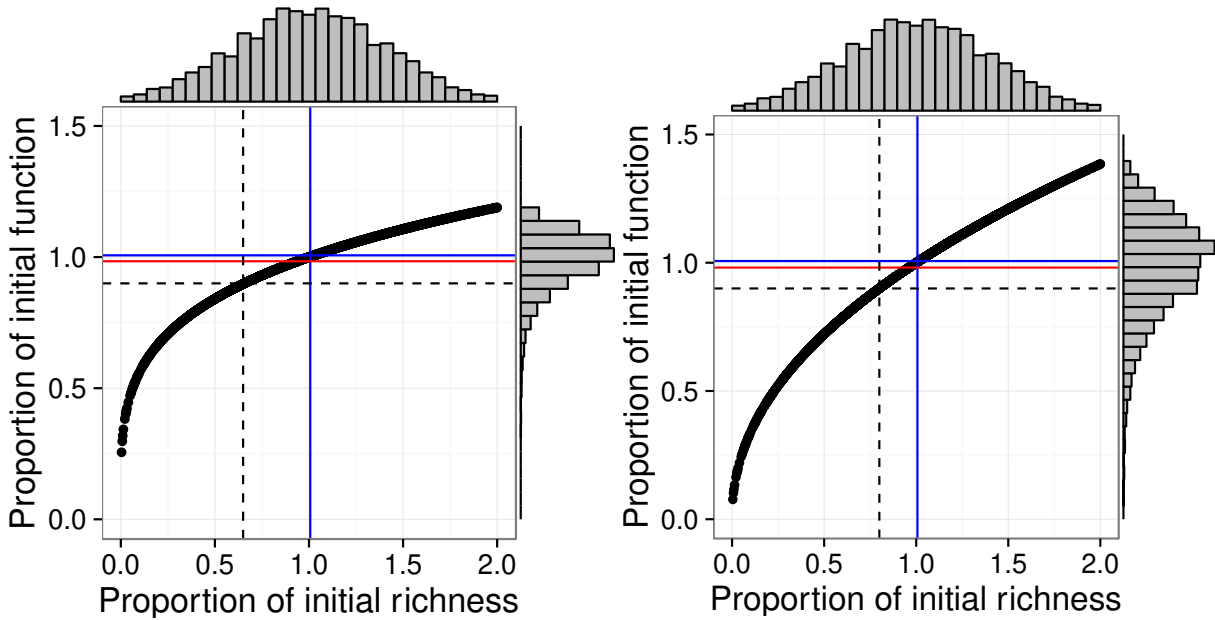
Group levels

36

Figure 3: Expected change in biomass associated with changes in species richness. Distribution of species richness changes (top histogram), expressed as a response ratio ($\ln(S_{Time1}/S_{Time2})$), and the distribution of associated change in biomass (vertical histogram), expressed as

710 $\ln(Y_{Time1}/Y_{Time0})$ expected for A) primary producers and B) herbivores. The distribution of expected function was produced using $Y = aS^b$ (the plotted curve) for values of $b = 0.26$ (Table 4). 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%

715 change in function: a 35% reduction in plant richness, and a 20% reduction in herbivore richness.

720