Is the relationship between algal diversity and biomass in North American lakes consistent with biodiversity experiments?

Emily K. Zimmerman and Bradley J. Cardinale

E. K. Zimmerman (ezim@umich.edu) and B. J. Cardinale, School of Natural Resources and Environment, Univ. of Michigan, Ann Arbor, MI 48109-1041, USA. BJC also at: Dept of Ecology and Evolutionary Biology, Univ. of Michigan, Ann Arbor, MI 48109-1041, USA.

Over the past few decades, a large body of research has examined how biodiversity loss influences the functioning of ecosystems, as well as the cascading impacts on the goods and services ecosystems provide to humanity. The relationship between biodiversity and ecosystem functions quantified in prior experiments suggests that initial losses of biodiversity have relatively small impacts on properties like community biomass production; however, beyond some threshold, increasing losses lead to accelerating declines in function. Some have questioned whether a saturating relationship between diversity and community biomass production is an artifact of overly simplified experiments that manipulate diversity in homogeneous conditions over short time-scales in which niche differences may not be realized. Others have questioned whether even the modest effects of biodiversity observed in experiments would be discernible in natural systems where they could be overridden by the stronger influence of abiotic factors.

Here, we used a biogeographic dataset to assess how the taxonomic richness of aquatic primary producers relates to community biomass in unmanipulated lake ecosystems in the US, and then compared these findings to prior experiments. We used structural equation modeling to evaluate hypotheses about the effects of algal richness on community biomass while accounting for covariance with environmental parameters measured in the USEPA’s National Lakes Assessment (NLA), which sampled 1157 freshwater lakes. These analyses converged on a single best-fit model ($\chi^2 = 0.31, p = 0.58$) wherein community algal biomass was a function of three explanatory variables – nitrogen, phosphorus, and algal richness. The quantitative magnitude of the algal diversity ($x$) – biomass ($y$) relationship in the NLA dataset is statistically greater than that documented in the average biodiversity experiment. It did, however, lie at approximately the 75th percentile of experimental relationships, indicating the diversity–biomass relationship in unmanipulated lakes is within the range that has been characterized experimentally.
run for many generations do, in fact, find that the impacts of biodiversity on community biomass tend to grow stronger through time (Cardinale et al. 2007, Stachowicz et al. 2008, Reich et al. 2012). Furthermore, some evidence suggests the effects of biodiversity on community biomass grow stronger as the spatial scale of experiments increase (Dimitrakopoulos and Schmid 2004, Cardinale et al. 2011, Griffin et al. 2013).

While many have claimed that biodiversity effects should grow stronger in more realistic environments with greater spatial and temporal heterogeneity, there is an alternative hypothesis that is rarely cited, and which has received far less attention. Some have argued that, when compared to experiments, species diversity should become less important in controlling ecosystem properties in natural systems because other environmental variables (nutrients, temperature, light, herbivores, etc.) will have far stronger impacts on ecosystem functioning (Fridley 2002, Wardle and Zackrisson 2005, Grace et al. 2007). Supporters of this hypothesis argue that, while diversity may have significant impacts that are easily detected in controlled experiments, those impacts may prove to be minor and undetectable in more natural systems compared to the influence of abiotic controls over ecosystem processes. There have been comparatively few direct tests of this hypothesis, and for those that do exist, conclusions are somewhat mixed. Grace et al. (2007), for example, found that the relationship between plant diversity and community biomass production in unmanipulated grassland ecosystems is weak relative to other environmental variables. In contrast, Paquette and Messier (2011) found that the relationship between tree diversity and annual wood production in forests was positive and significant after statistically controlling for other environmental variables that influence tree growth. Clearly, more case studies are needed if we are to assess whether biodiversity–ecosystem functioning relationships in ‘real’ ecosystems are accurately portrayed by the findings of small-scale, simplified biodiversity experiments.

Here, we present results from a study that was designed to evaluate hypotheses about the effects of algal richness on community biomass in natural lakes after accounting for environmental correlates, and then compare the magnitude of the richness–biomass relationship to that measured in past biodiversity experiments. Using a large biogeographic dataset of algal diversity and community biomass in freshwater lakes, we demonstrate that species richness is positively related to algal community biomass, and can be quantified by the same power function that is commonly used to describe experimental results. The scaling coefficient of this power function relating algal richness to community biomass is significantly positive, and remains so after statistically controlling for a suite of other environmental variables that are also known to influence the biomass of primary producers. The magnitude of the mean scaling coefficient quantified for natural lake ecosystems in this study is significantly larger than the mean scaling coefficient previously quantified in experimental manipulations of biodiversity for producers as a whole, and for the more limited set of experiments performed with pelagic microalgae. Even so, the mean scaling coefficient for natural lakes falls well within the range of variation quantified by prior experiments, lying at approximately the 75th percentile of experimental results. Thus, while the average biodiversity experiment does tend to underestimate the relationship between biodiversity and community biomass, the ‘true’ relationship between biodiversity and community biomass in natural systems is captured within the range of experimental results.

### Methods

**United States Environmental Protection Agency’s National Lakes Assessment – overview**

The bulk of data used in this study came from a large observational dataset that was collected by the United States Environmental Protection Agency (hereafter, US EPA) referred to as the National Lakes Assessment (hereafter, NLA). The original purpose of the NLA was to collect base-line measures that detail the condition of the US’s lake ecosystems. To be included in the NLA, lakes had to have been a natural or man-made freshwater lake, pond, or reservoir occurring in the continental United States, be greater than 10 acres (4 ha) in area, be at least one meter in depth, be accessible by field sampling crews, and have a minimum of a quarter acre of open water. The Great Lakes...
and the Great Salt Lake were not included in the survey. Lakes were chosen from the National Hydrography Dataset using two methods; 1033 lakes were chosen randomly using a statistical survey design and 124 lakes were chosen as reference lakes, with the latter selected non-randomly by the US EPA to represent the ‘least-disturbed’ lakes in the US (Fig. 1B).

Collection of data for the NLA took place during the summer of 2007. Field sampling for each of the lakes was completed in one-day increments by 86 teams composed of two to four technicians trained and deployed by the US EPA. At each lake, the field sampling crew followed standardized protocols (publicly available at <http://water.epa.gov/type/lakes/lakessurvey_index.cfm>) to collect samples at the deepest point of the lake and at ten different locations around the perimeter. Data collected on site, such as temperature and pH, were entered by field sampling crews into standardized data forms, whereas other samples were shipped to common laboratories for additional analyses. More than 680,000 measurements of the chemical, physical, biological and recreational characteristics of the lakes were quantified.

The US EPA employed a Quality Assurance Project Plan to ensure quality control at all levels of the study, from data collection by field sampling crews to standardized and central data management. All water chemistry samples were analyzed at the same laboratory under standard operating protocols administered by the US EPA’s Western Ecology Division. Zooplankton and diatom samples were sent to four different laboratories around the country, whereas samples for other phytoplankton were all processed at one laboratory. Laboratories were audited for adherence to the NLA standard operating protocols for benthic processing. All laboratories were subject to internal quality control on sorting and identification using the Integrated Taxonomic Information System and external quality control by independent taxonomists contracted to audit ten percent of each laboratory’s samples. Analyzed data were organized and entered into a series of spreadsheets that are publicly available at <http://water.epa.gov/type/lakes/NLA_data.cfm>.

**National Lakes Assessment – extraction and summary of relevant variables**

We extracted variables from the US EPA NLA that quantified algal diversity, algal biomass, and a suite of environmental variables known to influence algal biomass. Algal richness was taken to be the taxonomic richness in samples taken from a lake (lowest possible taxonomic unit, which was usually genus), and biomass was estimated as the summed biovolume across all algal species in the sample (log (µm³ l⁻¹)). Algal taxon richness and biomass were taken from the ‘Lake Phytoplankton Soft Algae Count Data’ spreadsheet and the ‘Lake Phytoplankton Diatom Count Data’ spreadsheet. We further selected ten environmental variables from the US EPA NLA to include in the analyses, such as measures of nutrient concentrations, zooplankton diversity and abundance, and water chemistry. The rationale for including each of these variables in the analyses, the sources of data for each variable, and the mean, variation and range of each variable measured in NLA lakes are given in Table 1.

It is important to note that the NLA dataset only had measures of standing stock biomass of algae measured at a single time point, and that rates of primary production (neither gross nor net) were not measured in this study. The spatial extent of this dataset is unrivaled (~1200 lakes across the continental US), but the large extent prohibited taking measures of biomass at multiple time points, or measuring gas exchange rates by phytoplankton in sealed bottles to estimate productivity per se. So it should be kept in mind that our analyses focus on how species richness relates to community-level biomass of algae measured at a single point in time across a large biogeographic gradient of many lakes. While the distinction between standing biomass (a stock) versus primary production (a flux) is important, as these do not necessarily relate to biodiversity in the same way (Weis et al. 2007, Power and Cardinale 2009), it should be noted that most studies of biodiversity–productivity relationships actually measure standing stock biomass at a single point in time, not primary production per se (Mittelbach et al. 2001, Cardinale et al. 2011).

Biotic data available from the NLA (zooplankton abundance and richness, algal taxon richness, total algal community biovolume) required a good deal of post-processing to be made useful for these analyses. Zooplankton were labeled with their functional feeding group by the US EPA. Only herbivorous zooplankton were included in our dataset. Zooplankton abundance was standardized across sites based on volume (abundance ml⁻¹), whereas zooplankton taxon richness was calculated by summing the number of taxa present in each lake. Algal taxon richness was calculated by summing the taxa richness of algae and the taxa richness of diatoms for each site from their respective spreadsheets. Total algal community biovolume for each site was calculated by summing individual taxa biovolume for each site. Any taxa for which biovolume data were not provided by the US EPA was excluded from the analyses. In cases for which biovolume data were not available for a given sample, but were available for the same taxa elsewhere in the dataset, that biovolume was assigned to the given sample.

Distributions of all variables were initially screened for linearity, normality, and correlations. Variables were log-transformed when necessary to improve the normality of the distributions.

**Variables from other sources**

Light was not explicitly measured during the US EPA NLA; yet, its role in photosynthesis and primary production is obvious. Therefore, it was necessary to gather data for light intensities from an alternative source. Data for the monthly average solar radiation incident on the surface of the earth was taken from the National Aeronautics and Space Administration Atmospheric Science Data Center by entering geographical coordinates and month of sampling for each of the 1157 lakes in the US EPA NLA. The geographical coordinates for each lake were provided in the US EPA NLA. Monthly average solar radiation incident (kWh m⁻² day⁻¹) is a direct measure of the light available on the surface of the water (Table 1).
Table 1. Description of environmental variables taken from the US EPA NLA and NASA. The asterisk (*) in the units column indicates that the variable was log transformed before use in the analysis.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Units</th>
<th>Rationale (reference)</th>
<th>Source</th>
<th>Mean ± SD, (min, max)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total nitrogen</td>
<td>(μg N l⁻¹)°</td>
<td>important limiting nutrient to primary production in freshwater ecosystems (Elser et al. 1990)</td>
<td>NLA lake water quality data</td>
<td>1162 ± 2154 (10, 26 100)</td>
</tr>
<tr>
<td>Total phosphorus</td>
<td>(μg P l⁻¹)°</td>
<td>important limiting nutrient to primary production in freshwater ecosystems (Elser et al. 1990)</td>
<td>NLA lake water quality data</td>
<td>107 ± 270 (1, 4679)</td>
</tr>
<tr>
<td>Silica</td>
<td>(μg SiO₂ l⁻¹)°</td>
<td>critical to the growth and ecology of diatoms, and may be a third limiting nutrient in systems lacking nitrogen and phosphorus (Martin-Jezequel et al. 2000, Evans et al. 2011)</td>
<td>NLA lake water quality data</td>
<td>8755 ± 10 630 (25, 91 907)</td>
</tr>
<tr>
<td>Growing days</td>
<td>latitude (decimal degrees)</td>
<td>latitude intended as a proxy for degree days, which are frequently used as a measure of growing period, or in the case of algae, ice-off period (Weyhenmeyer et al. 2013)</td>
<td>NLA information for lakes that were sampled</td>
<td>41 ± 5 (27, 49)</td>
</tr>
<tr>
<td>Dissolved organic carbon</td>
<td>(μg DOC l⁻¹)°</td>
<td>negatively impacts primary production, possibly through shading; of freshwater plankton communities (Carpenter et al. 1998)</td>
<td>NLA lake water quality data</td>
<td>8857 ± 16 846 (340, 290 570)</td>
</tr>
<tr>
<td>Zooplankton taxa richness</td>
<td>lowest taxonomic unit possible</td>
<td>significant influence of consumer diversity on rate of herbivory of algal biomass (Naem and Li 1998)</td>
<td>NLA lake zooplankton count data</td>
<td>7 ± 3 (0, 18)</td>
</tr>
<tr>
<td>Zooplankton abundance</td>
<td>(density ml⁻¹)°</td>
<td>aquatic herbivores remove as much as 79% of algal primary production (Cyr and Pace 1993)</td>
<td>NLA Lake zooplankton count data</td>
<td>70 ± 203 (0, 2602)</td>
</tr>
<tr>
<td>Water temperature</td>
<td>°C</td>
<td>influences photosynthesis (Davinson 1991) as well as lake turnover and stratification (Kalff 2002)</td>
<td>NLA lake profile data</td>
<td>21 ± 5 (7, 34)</td>
</tr>
<tr>
<td>Conductivity</td>
<td>(μS cm⁻¹ at 25°C)°</td>
<td>frequently noted environmental variable in freshwater algal studies; ions in solution may affect algal growth rates and uptake of phosphorus (Tilman et al. 1982)</td>
<td>NLA lake water quality data</td>
<td>662 ± 2455 (4, 50 590)</td>
</tr>
<tr>
<td>Water pH</td>
<td>pH units</td>
<td>frequently noted environmental variable in freshwater algal studies; hypothesized to play a role in nutrient availability and uptake of algal species (Tilman et al. 1982)</td>
<td>NLA lake water quality data</td>
<td>8 ± 1 (4, 10)</td>
</tr>
<tr>
<td>Light</td>
<td>kWh m⁻² day⁻¹</td>
<td>important environmental variable in primary production; measured as monthly average solar radiation incident (kWh m⁻² day⁻¹), which is a direct measure of the light available on the surface of the water</td>
<td>NASA</td>
<td>5 ± 1 (3, 8)</td>
</tr>
</tbody>
</table>

Analyses

**Structural equation modeling**

In the first step of the analyses, we used structural equation modeling (SEM) to quantify the relationship between algal taxa richness and algal community biomass while statistically controlling for the potential influence of other environmental variables. The initial model included 12 explanatory variables (Table 1, Fig. 2). While the NLA measured many more variables than this (e.g. land use surrounding sampling sites, maximum depth of sampled lake, perimeter of lake, lake trophic condition, etc.), many of the variables have only indirect effects on algal biomass, and represent proxies for variables that have a more direct influence on algal biomass. For example, land use does not have a direct effect on algae, but instead, leads to changes in nutrient loading, conductivity, and other metrics of water quality that are of direct consequence. We identified 10 explanatory variables from the NLA dataset that are known to have direct influence on algal richness or biomass (Table 1), and we specified five covariance paths in the initial model (Fig. 2) based on previously documented correlations in the literature. To confirm that the covariance paths predicted from the literature were actually covarying in the NLA dataset, we ran a correlation matrix in R ver. 2.15.0 and verified that all correlations specified based on the literature had p-values < 0.1 (Supplementary material Appendix A3 Table A3). These were included in the initial model (Fig. 2). For example, prior studies have shown that nitrogen and phosphorus loading are highly correlated in lakes, and that the two inorganic nutrients have synergistic effects on primary production that leads to more biomass in tandem than individually (Harpole et al. 2011). In the NLA dataset, the correlation between TN and TP was, in fact, strong and statistically significant ($r = 0.81$, p < 0.01). Thus, we included a covariance term between TN and TP in the initial structural equations model.

Beginning with the initial structural equation model (Fig. 2), we used the Lavaan package in R ver. 2.15.0 (Rosseel 2012) to parameterize the relationship between
quantitative magnitude of the algal diversity ($x$) – biomass ($y$) relationship in unmanipulated lakes to the magnitude of the relationship that has been quantified in prior experimental manipulations of diversity. We made the comparison using the scaling coefficient $b$ from the power function, $y = ax^b$ where $y$ is algal community biomass, and $x$ is algal species richness (note that richness and biomass are on a log scale in the SEM; thus the partial regression coefficient is equal to $b$ in the power equation). The power function has previously been used to summarize this same diversity–function relationship in prior BEF experiments (Cardinale et al. 2011). The power function is considerably more flexible than the Michaelis–Menten, and can fit a wider variety of functional relationships. In addition to having greater flexibility, experimental data fit to power functions had $R^2$-values that were very similar to the Michaelis–Menten function ($R^2_{\text{power}} = 0.71$, $R^2_{\text{MM}} = 0.73$) (Cardinale et al. 2011). Thus, while the power function is not

algal taxa richness and algal community biomass while considering the specified covariance terms and the effects of the selected environmental variables. We removed variables that improved the overall goodness of fit of the structural equation model by increasing parsimony. Specifically, we iteratively removed pathways with high p-values from the full model that improved fit (decrease in AIC, RMSEA). We used the $\chi^2$-statistic to assess the overall significance of model fits to data ($p > 0.05$) and Akaike information criteria (AIC) and root mean square error of approximation (RMSEA) to compare models of differing complexity (Akaike 1974, Grace 2006). In addition, we calculated the Akaike weights to estimate the likelihood that each model was the best fit to the observed data, given the suite of models hypothesized (Johnson and Omland 2004).

**Parameter comparisons**

After finding the most parsimonious, best-fit structural equation model, we used the resulting parameter estimates (partial regression coefficients) from that model to compare the

---

**Figure 2.** Full structural equation model including all variables and specified covariance terms. Solid lines with arrow head represent causal pathways. Dashed lines with dual arrow heads represent covarying variables. Standardized partial regression coefficients are noted. Significance is indicated by asterisks. The model’s chi-square was 2929.46, and the AIC was 37,244.75 ($p = 0.000$). Variance explained by the model was 0.209. Analyses were completed in Lavaan package in R. Growing season is measured by latitude (decimal degrees); though growing season is usually measured by degree days, latitude is highly correlated with degree days and could be log transformed to achieve normality.
the single best fit to prior experimental data, its greater flexibility to fit a variety of datasets coupled with its comparable explanatory power makes it useful for our purposes (see the same argument used to justify analyses by Reich et al. 2012).

For BEF experiments, values for the scaling parameter \( b \) were taken from the dataset of Cardinale et al. (2011) who summarized the form and strength of several hundred producer diversity–biomass relationships previously documented in experimental manipulations of species richness. From this dataset, we tallied 31 observations that have quantified how the richness of pelagic microalgae influences producer community biomass in aquatic ecosystems (freshwater or marine). The average value of \( b \) in these observations was \( 0.22 \pm SD 0.48 \). Unfortunately, the 31 observations that produce this estimate only come from five independent experiments (Gamfeldt et al. 2005, Zhang and Zhang 2006, Weis et al. 2007, Power and Cardinale 2009, Striebel et al. 2009), one of which (Weis et al. 2007) contributed 18 of the observations with a time-series that showed highly variable scaling relationships over time (thus, it could not be summarized into a single number). Because of the limited experimental data available for microalgae, we simultaneously expanded our comparison to include 288 observations that examined the effects of primary producer diversity on community biomass for systems composed of marine macroalgae (\( n = 29 \)), emergent wetland macrophytes (\( n = 13 \)), and systems of terrestrial herbaceous vegetation – mostly grassland studies (\( n = 214 \)). Clearly, these systems of primary producers are fundamentally different in the traits of the focal organisms, and in the ecological controls over their diversity and composition. In spite of obvious differences, explicit comparisons of diversity–biomass relationships have failed to detect statistically significant differences in this relationship among most ecosystems (Cardinale et al. 2006, 2011). Therefore, we felt it was potentially informative to compare results from the NLA dataset to this broader set of producer systems.

**Variables and relationships not considered**

The single greatest challenge in quantifying diversity–function relationships in natural systems is statistically controlling for the myriad of potentially confounding environmental variables. The second greatest challenge is limiting one’s search for confounding variables to a list that is manageably sized, and which only includes variables that have a direct causal influence on the focal relationship of interest (in this case, richness and biomass). We extracted variables from the NLA that have been shown to influence the biomass of aquatic primary producers via well-known biological mechanisms (Table 1). At the same time, we excluded from consideration certain relationships or variables for which we could not envision a direct causal pathway. One variable that is prominently missing from our analyses is lake area. While many studies have demonstrated a positive correlation between area and both biodiversity and biomass production (Conner and McCoy 1979, McGuinness 1984, Lomolino 2001), it is generally understood that area per se is not the direct causal factor. Rather, area is associated with other factors such as habitat heterogeneity, resource loading, or species population sizes that do directly influence biodiversity or ecosystem function (Williams 1964, Harner and Harper 1976). We recognize that other researchers might have a differing opinion on the utility of surrogates like area, so we have included supplemental analyses (Supplementary material Appendix A1 Fig. A1A–B) that suggest lake area has no bearing on the conclusions that follow from our analyses of this dataset.

It is also worth noting that we specifically chose to examine how biodiversity influences biomass, rather than the converse relationship. We recognize that a large body of historical research has examined how the ‘productivity’ of ecosystems influences species diversity (Connell and Orias 1964, MacArthur 1965). This assumed direction of causality has recently been criticized because mathematical models of competition make it clear that resource supply rates – not biomass production – are what influences biodiversity by regulating mechanisms of competition and coexistence (Loreau et al. 2001, Schmid 2002, Cardinale and Gross 2007, Cardinale et al. 2009b, c). Even so, empiricists routinely quantify standing biomass (which is easy to measure) as a proxy for resource supply rates (which are hard to measure), and then plot richness as a function of biomass as if this is a causal relationship. Because this historical perspective of causality remains prominent in ecology, we ran additional analyses that modeled algal species richness as a function of algal biomass (Supplementary material Appendix A2 Table A2). We began with the model in Fig. 2 except with richness modeled as the response variable and biomass modeled as an upstream causal variable (Supplementary material Appendix A2 Fig. A2). We again removed variables that improved the goodness of fit and achieved the lowest AIC. All 14 models considered had p-values < 0.05 for the \( \chi^2 \)-statistic, indicating that the predicted and observed covariance matrices were significantly different; thus, models with biomass as a ‘cause’ of species richness were unable to reproduce the observed dataset.

**Results**

The US EPA NLA dataset was characterized by a large amount of variation in algal taxa diversity, primary production, and the different environmental parameters, many of which ranged by several orders of magnitude. A total of 1006 taxa of phytoplankton were identified across the 1157 lakes, which ranged from a minimum taxon richness of 22 to a maximum of 85. Taxa representing all major taxonomic groups of freshwater algae were found, including Bacillariophyceae, Chlorophyta, Cryptophyta, Cyanophyta, Euglenophyta and Pyrrophyta. The most commonly occurring taxonomic group in the dataset was Bacillariophyceae. Total community biovolume in each lake spanned 7 orders of magnitude (22 to 1 310 659 599 \( \mu m^3 \) biovolume l\(^{-1} \)). The range of the various environmental parameters is given in Table 1. The extreme levels of variation are ideal for purposes of this study, as modeling approaches like SEM become increasingly powerful and more reliably quantify relationships when variables span a large range.

Results of the SEM analysis indicate that the initial model (Fig. 2), which included all variables and specified covariance terms, was not the best-fit model. To find the best-fit model, we iteratively removed variables to improve the goodness-of-fit between the predicted and observed covariance matrices.
Table 2. Summary of 16 SEMs run to determine influences of algal diversity and environmental variables on total algal community biovolume. Models arranged using model identification number. Best-fit model is the first listed. Models were selected based on significant $\chi^2$-values ($p > 0.05$) and lowest AIC value. The difference between the model $i$ and the best-fit model ($\Delta i$) was calculated for each model. Using these values, the likelihood of a model, $m_i$, given the data, $y$, ($L(m_i | y)$) was calculated as $L(m_i | y) = \exp(-1/2 \Delta_i)$. The Akaike weight ($w_i$) was calculated using the ($L(m_i | y)$) calculations. Akaike weight provides a relative weight of evidence for each model, and can be interpreted as the probability of model $i$ being the best-fit model of the candidate models, given the data. Akaike weight was calculated by normalizing model likelihood values ($L(m_i | y)$) across all models.

| Model (m) | Model description | DF | $\chi^2$ | p | AIC | $\Delta_i$ | $L(m_i | y)$ | wi |
|-----------|-------------------|----|----------|---|-----|--------|-------------|----|
| 15        | community biomass of algae = f(μg nitrogen l$^{-1}$, μg phosphorus l$^{-1}$, algal taxa richness) covariance: μg nitrogen l$^{-1}$ – μg phosphorus l$^{-1}$, μg nitrogen l$^{-1}$ – algal taxa richness | 1.00 | 0.31 | 0.58 | 5952.45 | 0.00 | 1.00 | 0.74 |
| 16        | community biomass of algae = f(μg nitrogen l$^{-1}$, μg phosphorus l$^{-1}$, algal taxa richness) covariance: μg nitrogen l$^{-1}$ – μg phosphorus l$^{-1}$ | 2.00 | 4.35 | 0.11 | 5954.49 | 2.04 | 0.36 | 0.26 |
| 11        | community biomass of algae = f(μg nitrogen l$^{-1}$, μg phosphorus l$^{-1}$, algal taxa richness, growing period) covariance: μg nitrogen l$^{-1}$ – μg phosphorus l$^{-1}$ | 4.00 | 42.98 | 0.00 | 12932.43 | 6979.99 | 0.00 | 0.00 |
| 12        | community biomass of algae = f(μg nitrogen l$^{-1}$, μg phosphorus l$^{-1}$, algal taxa richness, temperature) covariance: μg nitrogen l$^{-1}$ – μg phosphorus l$^{-1}$ | 4.00 | 222.80 | 0.00 | 12953.72 | 7001.27 | 0.00 | 0.00 |
| 10        | community biomass of algae = f(μg nitrogen l$^{-1}$, temperature, algal taxa richness, growing period) covariance: growing period – temperature, μg nitrogen l$^{-1}$ – algal taxa richness | 4.00 | 225.5 | 0.00 | 18501.55 | 12549.11 | 0.00 | 0.00 |
| 14        | community biomass of algae = f(μg nitrogen l$^{-1}$, μg phosphorus l$^{-1}$, growing period, temperature) covariance: μg nitrogen l$^{-1}$ – μg phosphorus l$^{-1}$ | 4.00 | 283.44 | 0.00 | 18845.74 | 12893.3 | 0.00 | 0.00 |
| 13        | community biomass of algae = f(μg phosphorus l$^{-1}$, temperature, algal taxa richness, growing period) covariance: growing period – temperature | 5.00 | 252.8 | 0.00 | 19369.95 | 13417.5 | 0.00 | 0.00 |
| 9         | community biomass of algae = f(μg nitrogen l$^{-1}$, μg phosphorus l$^{-1}$, growing period, temperature, algal taxa richness) covariance: μg nitrogen l$^{-1}$ – μg phosphorus l$^{-1}$ | 7.00 | 297.91 | 0.00 | 19582.13 | 13629.69 | 0.00 | 0.00 |
| 8         | community biomass of algae = f(μg nitrogen l$^{-1}$, μg phosphorus l$^{-1}$, light, growing period, temperature, algal taxa richness) covariance: μg nitrogen l$^{-1}$ – μg phosphorus l$^{-1}$ | 12.00 | 367.21 | 0.00 | 22349.97 | 16397.53 | 0.00 | 0.00 |
| 7         | community biomass of algae = f(μg nitrogen l$^{-1}$, μg phosphorus l$^{-1}$, light, growing period, temperature, zooplankton abundance/ml, algal taxa richness) covariance: μg nitrogen l$^{-1}$ – μg phosphorus l$^{-1}$ | 17.00 | 481.13 | 0.00 | 24854.3 | 18901.86 | 0.00 | 0.00 |
| 6         | community biomass of algae = f(μg nitrogen l$^{-1}$, μg phosphorus l$^{-1}$, light, growing period, temperature, zooplankton abundance ml$^{-1}$, zooplankton taxa richness, algal taxa richness) covariance: μg nitrogen l$^{-1}$ – μg phosphorus l$^{-1}$ | 24.00 | 732.33 | 0.00 | 30634.57 | 24682.12 | 0.00 | 0.00 |
| 5         | community biomass of algae = f(μg nitrogen l$^{-1}$, μg phosphorus l$^{-1}$, conductivity, light, growing period, temperature, zooplankton abundance ml$^{-1}$, zooplankton taxa richness, algal taxa richness) covariance: μg nitrogen l$^{-1}$ – μg phosphorus l$^{-1}$ | 32.00 | 1420.71 | 0.00 | 32612.12 | 26659.67 | 0.00 | 0.00 |
| 4         | community biomass of algae = f(μg nitrogen l$^{-1}$, μg phosphorus l$^{-1}$, pH, light, growing period, temperature, zooplankton abundance/ml, zooplankton taxa richness, algal taxa richness) covariance: μg nitrogen l$^{-1}$ – μg phosphorus l$^{-1}$ | 32.00 | 1172.98 | 0.00 | 33235.18 | 27822.73 | 0.00 | 0.00 |

(Continued)
This led to 15 additional models (Table 2), all of which represented alternative hypotheses to explain patterns of covariance in the dataset. In all models, including the initial model presented in Fig. 2, the pathway between algal taxa richness and total algal community biomass production was positive and significant. The best-fitting model to explain total algal community biomass was far more parsimonious than the original model, and included just three explanatory variables – total nitrogen, total phosphorus and algal taxa richness (Fig. 3). The strongest predictor of total algal community biomass in this final model was total nitrogen with a standardized partial regression coefficient of 0.26 (p < 0.05). The standardized partial regression coefficient from total phosphorus to total algal community biomass was 0.15 (p < 0.05), and the standardized partial regression coefficient from algal taxa richness to total algal community biomass was 0.23 (p < 0.05). Additional positive covariance terms existed between algal taxa richness and total nitrogen and between total nitrogen and total phosphorus. Only the latter of these was statistically significant (p < 0.05). The best-fit model with total nitrogen, total phosphorus, and algal taxa richness (Fig. 3) explained ~21% of the total variation in algal community biomass across all 1157 lakes with a χ² of 0.31 (p = 0.58) and a root mean square error of approximation (RMSEA) of 0.07 (p = 0.88). For the χ²-statistic, a p-value greater than 0.05 indicates that the expected and observed covariance matrices are not significantly different; thus, the hypothesized model cannot be rejected as a viable explanation of the data.

We calculated the Akaike weights of all 16 SEM models, which use the likelihood values to compare how probable each model i as an explanation for the observed data, given the suite of models hypothesized. The Akaike weight for the best-fit model (Fig. 3) was 0.74, indicating that the best-fit model in Fig. 3 had a 74% chance of being the best fit model for the observed data, given the suite of models hypothesized (Table 2). The second most likely model, which only differed in that the covariance term between algal richness and total nitrogen was deleted, had a 26% chance of being the best fit model for the observed data (Table 2). All 14 remaining models had Akaike weights of 0.00, indicating that they were highly unlikely to be viable explanations of the observed dataset (Table 2).

![Figure 3. Best-fit structural equation model. Solid lines with arrow head represent causal pathways. Dashed lines with dual arrow heads represent covarying variables. Standardized partial regression coefficients are noted first. Unstandardized partial regression coefficients are noted second in parentheses. Significance is indicated by asterisks. The model’s χ² was 0.307, and the AIC was 5,952.45 (p = 0.580). Variance explained by the model was 0.209. Analyses were completed in Lavaan package in R.](image-url)
One surprising result of the model comparison is that several models that included variables routinely thought to control algal biomass proved to be inferior explanations of the observed covariance matrix, leading us to reject them as viable explanations of the data. For example, zooplankton taxa richness and zooplankton abundance – both of which can influence the magnitude of herbivory on algae, were not retained in our best-fitting models. This was surprising given the wealth of studies demonstrating the role of herbivory on biomass production in aquatic ecosystems (Cyr and Pace 1993). In addition to herbivory, light was also not retained in our best-fitting models, which was surprising given the obvious role of light in maintaining primary production.

After identifying the single best-fitting model, we used the parameter values from this model to compared the strength of the algal taxa richness \( (x) \) – biomass \( (y) \) relationship in the NLA dataset (Fig. 3) to the strength of relationships documented in prior experiments using the power function, \( y = ax^b \). The unstandardized scaling coefficient \( b \) relating biovolume to taxa richness from the best-fitting SEM was 0.52 ± SEM 0.07 (note that richness and biomass are both on a log scale in the SEM; thus the estimated slope is equal to \( b \) from the power function). Figure 4 compares this scaling coefficient to that documented in past biodiversity experiments. For experiments that have manipulated the richness of pelagic microalgae, \( b \) has averaged 0.22 with a 95% confidence interval of 0.05 to 0.40 (n = 31 observations). This is fairly similar to the values that have been estimated for other systems of primary producers (Fig. 4); in fact, across all experimental manipulations of producer diversity, \( b \) has averaged 0.23 with a confidence interval of 0.20 to 0.26. Thus, the value of \( b \) documented in this study for the NLA lakes dataset is significantly higher than the upper bound of the confidence interval for experiments. However, it is noteworthy that estimate of \( b \) from the NLA dataset falls well within the range of estimates provided from prior experiments (past estimates), and lies at roughly the 75th percentile for experiments performed with pelagic microalgae (Fig. 4).

**Discussion**

Over the past two decades, accelerating rates of biodiversity loss, and the growing concern about how these losses impact ecosystem functions, have led to an increase in the number of experiments focused on understanding how biodiversity impacts ecosystem-level processes like community biomass production (Loreau 2010). These experiments have provided strong evidence that community biomass increases as a function of biodiversity; yet, many of the experiments have been completed at relatively small spatial and temporal scales (Cardinale et al. 2011). While many have called on researchers to increase the spatial and temporal scales used in BEF experiments in order to provide a more ‘realistic’ comparison to natural ecosystems, the practical limitations (funding, personnel, experimental control, etc.) to performing large-scale, long-term manipulative experiments make it highly unlikely that we will routinely mimic realistic scenarios of extinction any time soon. As an alternative, some have suggested that we complement the expanding number of BEF experiments with analyses of large, observational datasets that are representative of unmanipulated ecosystems (Loreau et al. 2001, Cardinale et al. 2011); thus allowing us to compare the results of mechanistic, highly replicated BEF relationships performed in relatively simplistic environments to the BEF relationship that occurs in more natural ecosystems.

Several studies have now explicitly used large observational datasets to characterize the relationship between

![Figure 4](image.png)

**Figure 4.** Comparison of the mean quantitative magnitude of the diversity \( (x) \) – productivity \( (y) \) relationship of experiments across 4 taxa-specific groupings to the mean quantitative magnitude of the diversity \( (x) \) – productivity \( (y) \) relationship in natural lakes from this study. Quantitative magnitude is measured by calculating the mean scaling coefficient, \( b \), from the equation \( \log(\text{producer biomass}) = a + b \log(\text{richness}) \). The scaling coefficient, \( b \), is significantly larger in natural lakes in this study than that of prior experiments across all taxa-specific groups (\( p < 0.05 \)). Data for mean \( b \) values of prior experiments came from the dataset of Cardinale at al. (2011), and the mean \( b \) value of natural lakes came directly from our best-fitting SEM.
biodiversity and various ecosystem-level properties; but the conclusions of these studies have proven to be mixed. Paquette and Messier (2011) used a large observational dataset from the Quebec Forest Survey to examine the relationship between tree biodiversity and wood production. They found a significant, positive relationship between tree functional richness and forest productivity that was strong even after statistically controlling for several important climatic variables that influence wood production. Mora et al. (2011) used a global survey of 1906 reefs to evaluate the relationship between reef fish diversity and standing biomass of reef fishes. Using structural equation modeling to account for the potential influence of environmental, physiographic, and anthropogenic variables, Mora et al. (2011) suggested the relationship between fish biodiversity and standing biomass was stronger than that documented in prior experiments. In contrast, Grace et al. (2007) used structural equation modeling to investigate the relationship between plant biodiversity and community biomass, while statistically controlling for environmental variables in 12 natural grassland systems spanning four countries. They found no significant relationship between plant diversity and community biomass in natural grasslands, and instead suggested that producer biomass was more consistently and strongly controlled by abiotic variables.

Findings from this study are consistent with the proposition that biodiversity’s impact on community biomass in natural, unmanipulated systems is significantly stronger than revealed by prior experiments (Fig. 1A). Indeed, the mean scaling coefficient relating biomass to diversity in natural lakes exceeds mean experimental estimates (Fig. 4). It is, however, worth noting that the mean scaling coefficient relating algal biomass to diversity in natural lakes fell well-within the range of experimental estimates (Fig 4), suggesting that experiments do, in fact, potentially capture the nature of this relationship. Of course, the validity of this comparison hinges on the assumption that observational gradients in species richness across sites can be used to predict the functional consequences of biodiversity in much the same way that experimental manipulations reveal the functional role of diversity within individual sites. There are any number of reasons why this assumption could be incorrect. For example, if the biological mechanism(s) that maintain(s) diversity within a site differ from those that generate variation among sites, then the consequences of changing diversity could also differ. Changes in diversity from communities that have co-evolved for thousands of years could also have fundamentally different consequences for the functioning of ecosystems than changing diversity in experimental plots where species have interacted for a short time. Even so, it is unrealistic to think we will ever perform experimental manipulations of algal diversity at the scale of real lakes to achieve a direct comparison to biodiversity experiments. While a comparison between experiments and observational gradients is, perhaps, the best we can expect, our confidence in such comparisons could be improved if future work would focus on explicitly testing these underlying assumptions.

The hypothesis that diversity has a stronger impact on community biomass in natural ecosystems than in controlled experiments is based on the assumption that natural, unmanipulated ecosystems have more spatial and temporal heterogeneity. In turn, this additional heterogeneity allows more species niche differences to be expressed in unmanipulated systems than in relatively homogenous experiments. This is an assumption that has yet to be directly tested or verified in empirical studies of natural systems, and we have no ability to determine if this is indeed the underlying cause of a stronger diversity–biomass relationship in the NLA dataset. But it is worth noting that a select group of experimental studies have suggested that diversity effects grow increasingly strong at larger spatial and temporal scales, which could be due to increasing heterogeneity. For example, Reich et al. (2012) analyzed data from two long-term (13 + years) biodiversity experiments performed at the Cedar Creek Natural History area and found that effects of plant diversity on plant biomass grew increasingly strong through time. The authors attributed this trend to a divergence of the biological traits of the focal species, which is consistent with, though by no means conclusive evidence that, niche differences are more fully expressed at longer-time scales. Similarly, Stachowicz et al. (2008) compared the relationship between macroalgal diversity and community biomass accrual in marine intertidal zones over a three year period, and proposed that the positive relationship between diversity and productivity became stronger over time because of facilitation and differential use of the heterogeneous environment. While these, and other, studies suggest that the strength of diversity–function relationships grows stronger as the spatial or temporal scale of experiments increase (see Cardinael et al. 2011 or Griffin et al. 2013 for summaries), the role of heterogeneity in promoting a stronger relationship has more often been presumed than experimentally or statistically controlled for.

A number of variables, such as light and zooplankton abundance, which would have been expected to be important drivers of algal biomass production based on studies by others (Hill and Knight 1988, Feminella and Hawkins 1995, Cyr and Pace 1993), were not included in the best-fit model. We can only speculate on reasons why these variables were not significant. One possible reason for their lack of effects could be the relatively coarse measures that were available for analyses. Light obtained from NASA represented the total solar radiation incident on a horizontal surface at Earth’s surface for a given month, with that value representing the average of a 22-year collection period. Thus, light intensities did not correspond to actual conditions on, or immediately preceding, the day of sampling for the NLA project, and that may explain why there was no signal on this particular date. The lack of a signal of zooplankton is perhaps more perplexing given the well-known influence of grazers on algal biomass and turnover rates in freshwater ecosystems (Cyr and Pace 1993). The NLA only had measures of zooplankton abundance available, but did not have estimates of biomass. Abundance may not necessarily be a good proxy of consumption since consumption of algae by zooplankton is a function of the body size of the zooplankton. As such, biomass would likely be a more appropriate measure of the potential influence of herbivores on algal biomass. Unfortunately, measures available in the NLA dataset do not allow us to examine these possibilities; as such, we suggest interpreting the non-significant impacts of zooplankton (and light) cautiously.

As with all biogeographic studies, our work has several obvious limitations and boundaries. We address just a few
such limitations here. First, while this dataset has unrivaled spatial resolution (variables from ~1200 lakes across the Continental US measured in a consistent way), all lakes were sampled at just a single time point. Because the data is only a snapshot in time, we cannot test the temporal dynamics of the relationship between algal taxa richness and community biomass, and cannot rule out the possibility that the data expressed a particularly strong relationship between algal richness and biomass at the particular time of sampling. Second, structural equation modeling (SEM) is a means to test hypotheses about multivariate causal relationships; in the case of our study, the causal relationship between richness and community biomass while accounting for the influence of other environmental variables that affect biomass. But conclusions from these SEMs are only as reliable as the suite of hypothesized models. If the suite of hypothesized models does not contain the ‘true’ set of causal relationships, then our analyses would identify a ‘best-fit’ model – but one that is fundamentally incorrect. We made every attempt to evaluate a variety of hypothetical models showing how variables influence algal biomass; nevertheless, future work should continue to pit new, alternative hypotheses against the best fit obtained in Fig. 3.

In conclusion, our study shows that, on average, experiments tend to underestimate the diversity–productivity relationship in unmanipulated, natural lake ecosystems. Yet, the quantitative magnitude of the relationship falls within the range predicted from two decades of prior BEF experiments, suggesting that while experiments underestimate, they are not vastly off. There have been many hypotheses for why the relationship between biodiversity and ecosystem functioning may be stronger in natural environments than in small-scale, short-term experiments, including that greater spatial and temporal heterogeneity in natural environments allows more species niche differences to be expressed. This hypothesis has yet to be adequately tested, and the use of a large observational dataset such as the NLA does not allow us to differentiate between mechanisms that may be causing this stronger relationship. We would argue that even more comparisons between BEF experiments and natural systems are needed to verify the robustness of predictions from these experiments. At the same time, when differences between natural and experimental systems do become apparent, determining the cause of these discrepancies and incorporating them back into new experiments is the best way to improve predictions about the functional consequences of diversity loss.

References

Grace, J. B. 2006. Structural equation modeling and natural systems. – Cambridge Univ. Press.