Ecology, 92(11), 2011, pp. 2096–2107

Large-scale biodiversity patterns in freshwater phytoplankton

Maayke Stomp,[1,2] Jef Huisman,[2,4] Gary G. Mittelbach,[3] Elena Litchman,[3] and Christopher A. Klausmeier[1]

[1]*Kellogg Biological Station and Department of Plant Biology, Michigan State University, Hickory Corners, Michigan 49060 USA*
[2]*Aquatic Microbiology, Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, P.O. Box 94248, 1090 GE Amsterdam, The Netherlands*
[3]*Kellogg Biological Station and Department of Zoology, Michigan State University, Hickory Corners, Michigan 49060 USA*

Abstract. Our planet shows striking gradients in the species richness of plants and animals, from high biodiversity in the tropics to low biodiversity in polar and high-mountain regions. Recently, similar patterns have been described for some groups of microorganisms, but the large-scale biogeographical distribution of freshwater phytoplankton diversity is still largely unknown. We examined the species diversity of freshwater phytoplankton sampled from 540 lakes and reservoirs distributed across the continental United States and found strong latitudinal, longitudinal, and altitudinal gradients in phytoplankton biodiversity, demonstrating that microorganisms can show substantial geographic variation in biodiversity. Detailed analysis using structural equation models indicated that these large-scale biodiversity gradients in freshwater phytoplankton diversity were mainly driven by local environmental factors, although there were residual direct effects of latitude, longitude, and altitude as well. Specifically, we found that phytoplankton species richness was an increasing saturating function of lake chlorophyll *a* concentration, increased with lake surface area and possibly increased with water temperature, resembling effects of productivity, habitat area, and temperature on diversity patterns commonly observed for macroorganisms. In turn, these local environmental factors varied along latitudinal, longitudinal, and altitudinal gradients. These results imply that changes in land use or climate that affect these local environmental factors are likely to have major impacts on large-scale biodiversity patterns of freshwater phytoplankton.

Key words: altitude; biodiversity gradient; climate change; elevation; lake; latitude; microbial diversity; phytoplankton; productivity; species richness; structural equation model; temperature.

Introduction

The increases in species richness of plants and animals from the poles to the tropics, and from the mountains to the lowlands provide two of the most striking biodiversity patterns on our planet (Rosenzweig 1995, Gaston 2000, Willig et al. 2003). Biologists have studied these large-scale biodiversity gradients in macroorganisms for centuries, leading to many insights on the biogeographical variation of species richness. For instance, temperature is considered to be a major determinant of latitudinal and altitudinal gradients in plant and animal diversity (Rohde 1992, Allen et al. 2002), diversity is often (but not always) highest at intermediate levels of ecosystem productivity (Grime 1973, Waide et al. 1999, Mittelbach et al. 2001), and species richness generally increases with habitat area (Arrhenius 1921, MacArthur and Wilson 1967, Rosenzweig 1995).

Until recently, it was controversial whether similar biodiversity patterns apply to microorganisms as well. Microorganisms have been hypothesized to show only weak geographic variation in diversity compared to macroorganisms. Their small size, high abundance, fast population growth, and long-range dispersal are often thought to increase the chances for microorganisms to reach new habitats and establish new populations, which could smooth out any diversity gradient (Fenchel and Finlay 2004, Hillebrand 2004). However, recent studies indicate that spatial diversity patterns do exist for microorganisms (Martiny et al. 2006). Fossil records (Stehli et al. 1969) and two recent surveys of contemporary pelagic marine bacteria (Pommier et al. 2007, Fuhrman et al. 2008) document a gradient of decreasing microbial diversity from low to high latitudes. Likewise, Rutherford et al. (1999) found that the diversity of marine planktonic foraminifera displays a latitudinal gradient, where diversity peaks at mid-latitudes, is lowest at high latitudes and is intermediate in the tropics. Altitudinal diversity gradients have been described for soil bacteria (Bryant et al. 2008) and diatoms (Wang et al. 2011). Surprisingly, relatively few studies have quantified spatial diversity gradients of entire algal communities (Dodson et al. 2000, Irigoien et al. 2004, Ptacnik et al. 2010), although diversity patterns of specific taxonomic groups, such as diatoms, have been studied in more detail (Vyverman et al. 2007, Passy 2008, 2010, Cermeño and Falkowski 2009, Soininen et al. 2009, Wang et al. 2011).

Manuscript received 20 May 2010; revised 10 May 2011; accepted 20 May 2011. Corresponding Editor: P. R. Leavitt.
[4] Corresponding author. E-mail: j.huisman@uva.nl

Phytoplankton form a highly diverse group of prokaryotic and eukaryotic microorganisms, and have served as one of the paradigm systems for studies on the maintenance of species diversity (the "paradox of the plankton"; Hutchinson 1961, Huisman and Weissing 1999, Stomp et al. 2004). Several factors are known to affect phytoplankton species coexistence at a local scale, such as productivity (Leibold 1996), nutrient supply ratios (Tilman 1982, Sommer 1993) and the underwater light climate (Huisman et al. 2004, Stomp et al. 2007). Geographical variation in these environmental factors has been implicated as a major determinant of large-scale patterns in phytoplankton diversity. For instance, earlier studies have shown that phytoplankton diversity is a unimodal function of productivity, both in freshwater lakes (Dodson et al. 2000) and marine ecosystems (Irigoien et al. 2004). A global ocean model predicts a decrease in phytoplankton diversity with increasing latitude (Barton et al. 2010), but these model results are under debate (Huisman 2010) and empirical studies failed to find a latitudinal diversity gradient in marine phytoplankton (Cermeño et al. 2008). Recent work by Ptacnik et al. (2010) showed that patterns of phytoplankton diversity across Scandinavian lakes are related to regional variation in phosphorus availability. Other studies point at the importance of historical processes such as migration and dispersal in determining large-scale biodiversity patterns of freshwater diatoms (Vyverman et al. 2007, Soininen et al. 2009), but dispersal limitation seems less important in marine diatoms (Cermeño and Falkowski 2009). In total, these studies have provided many new insights on phytoplankton biodiversity patterns, but a comprehensive understanding of the relative importance of different environmental drivers of phytoplankton diversity across large spatial scales is still lacking.

We took advantage of an extensive data set to examine possible patterns of phytoplankton diversity across the continental United States and to evaluate the environmental factors driving these patterns. Samples from 540 lakes and reservoirs from across the United States were collected by the Environmental Protection Agency (EPA). In each sample, phytoplankton were counted, usually to the species level. In addition, the EPA measured several chemical and physical parameters in each lake (e.g., nutrients, temperature, turbidity, lake area, lake depth). This extensive data set provides a unique opportunity to investigate large-scale biodiversity gradients in freshwater phytoplankton, and to test whether they are governed by similar underlying mechanisms as those for terrestrial plants and animals.

Methods

Data collection

Data were collected from 540 lakes and reservoirs sampled from 1973–1975 as part of the National Eutrophication Survey conducted by the U.S. EPA. Here we summarize their methods (see Taylor et al. 1979 for detailed information). The lakes were sampled one to four times, with most of them sampled three times within the same year, in spring, summer and fall. The depth-integrated water samples were taken from the surface to 4.6 meters depth, or from the surface to the photic depth (depth at which light intensity falls to 1% of the surface light intensity), whichever was greater, at the deepest point in each lake and reservoir. If the depth of the sampling site was less than 4.6 m, the sample was taken from just above the sediment to the surface. Phytoplankton species abundances and a number of physical and chemical parameters were all measured on the same water sample. Subsamples for phytoplankton counts were preserved with Lugol's solution (IKI), and were shipped to the Environmental Monitoring and Support Laboratory, Las Vegas, Nevada, where phytoplankton were identified to species or genus level and enumerated by microscopes using a Neubauer counting chamber (Hausser Scientific, Horsham, Pennsylvania, USA) at 400×. If greater detail was essential to accurately identify diatom species, a phase-contrast microscope was used. The count was stopped when a minimum of 100 fields (corresponding to a total volume of 0.4 μL) had been viewed or when the dominant species had been observed a minimum of 100 times. Quality control checks were performed on 5% of the samples by G. W. Prescott of the University of Montana to verify species identifications. The agreement between the original counts and the quality control checks for species identification was good. The phytoplankton taxonomy results were published in EPA reports (see Appendix A for references). We used the total number of phytoplankton species encountered in each lake over the course of a year as our measure of phytoplankton diversity (species richness).

Chlorophyll *a* concentration, nutrient concentration, turbidity of the water column (as measured by Secchi depth), water temperature, lake depth, lake surface area, and geographical locations of the lakes were determined using EPA methods (Taylor et al. 1979). These environmental data were originally published in EPA NES Working Papers (see Appendix A for references), and are freely available at the EPA Legacy and Storage and Retrieval data system (*available online*).[5]

For our analysis, we used the mean values for each of these variables, averaged over depth and year. Chlorophyll *a* is the core photosynthetic pigment present in all phytoplankton species, and its concentration was therefore used as a simple proxy of total phytoplankton biomass. Total chlorophyll *a* is often used as a proxy of primary productivity as well (Eppley et al. 1985, Falkowski and Raven 1997), assuming that the rate of carbon fixation is positively correlated with the chlorophyll *a* concentration. Altitudes of the sampled lakes were retrieved from GPS Visualizer, using their geo-

[5] ⟨http://www.epa.gov/storet⟩

graphical coordinates (*available online*).[6] The data of the 540 lakes and reservoirs used in this paper are provided in the Supplement.

Data analysis

Our working hypothesis is that large-scale patterns of phytoplankton diversity can be interpreted as a multileveled biogeographical pattern. That is, we envision that the species richness of a lake is to a large extent determined by local environmental factors (e.g., nutrients, water temperature), and that these local environmental factors in turn vary along biogeographical gradients. This nested structure is reflected in our data analysis.

First, we used stepwise multiple regression with forward selection of variables to identify the most important local environmental variables explaining species richness in the 540 lakes and reservoirs. The dependent variable, species richness, was log-transformed to homogenize its variance. Environmental variables were log-transformed if this resulted in a more uniform spread of data points. Environmental variables tested in the multiple regression analysis were surface area of the lake (km^2), lake depth (m), water temperature (°C), total phosphorus (mg/L), total nitrogen (mg/L), Secchi depth (m), and chlorophyll *a* concentration (μg/L), as well as the square of chlorophyll *a* concentration consistent with the nonlinear relationship observed between chlorophyll *a* and species richness. In addition, the interactive effect of total nitrogen (TN) and total phosphorus (TP) concentration (TN × TP) was used to test whether the TN:TP ratio had an impact on species richness (Sommer 1990). Second, we used the same stepwise multiple regression approach to investigate how each of these local environmental variables varied with the geographical coordinates of latitude, longitude and altitude (km). Variables were selected in the multiple regression only if $P < 0.001$. The analysis was performed using SPSS version 16.0 (SPSS Inc., Chicago, Illinois, USA).

To gain further understanding of the linkages between species richness, environmental variables and geographical location, we developed a structural equation model (SEM). Structural equation modeling is a robust multivariate statistical method that allows for hypotheses testing of complex networks of relationships and its use is increasing in a wide range of ecological applications (e.g., Shipley 2002, Grace et al. 2010). We started with an initial SEM that included all plausible pathways between species richness, our set of local environmental variables, and the geographical coordinates of the lakes. In addition, the pathway between chlorophyll *a* and its quadratic term was included to represent their intercorrelation (Grace 2006). We corrected for uneven sampling effort by including the number of days sampled as a cause of species richness and a covariate of all environmental parameters in the SEM. Number of days sampled was positively related to species richness, negatively correlated with latitude, longitude, and altitude and positively correlated with lake area. By including the number of days sampled in the SEM, the presented results are exclusive of these spurious influences.

Our first attempt with the initial model revealed that it was under-identified, meaning that there was some redundancy such that it was not possible to estimate all of the model's parameters. We therefore investigated the statistical relationships among the variables included in the model to identify possible redundancies. This revealed high collinearity between total phosphorus and total nitrogen (Pearson correlation, $\rho = 0.73$, $P < 0.001$), with very similar effects on chlorophyll *a* concentration. Furthermore, Secchi depth was strongly and inversely correlated to chlorophyll *a* concentration, but was unrelated to the other variables once chlorophyll *a* concentration was incorporated in the model. Removal of total nitrogen and Secchi depth from the SEM eliminated the under-identification of the initial model.

Subsequently, the significance of each path-coefficient was tested by its critical ratio ($P < 0.001$), and nonsignificant paths were removed in a stepwise fashion until all remaining paths were significant. This procedure of stepwise reduction of the initial model is called model trimming (Kline 2005). In addition, we removed pathways from the model if they were not directly or indirectly linked to species richness. The overall fit of the final model was evaluated with the adjusted goodness-of-fit index (AGFI) (Schermelleh-Engel et al. 2003). We calculated both standardized and unstandardized path-coefficients for each pathway in the final model. The standardized path coefficients are independent of the units of measurement of the different variables, thus enabling comparison of the relative contributions of the different paths in the structural equation model. However, standardized coefficients are known to depend on the range of variation within each variable, which could bias its interpretation (Grace and Bollen 2005). Therefore, we also report the unstandardized coefficients and their standard errors (Appendix B). The structural equation analyses was performed using R version 2.8.1 (R Development Core Team 2008) with package sem version 0.9–21 (*available online*).[7]

Results

Biogeographical distribution of phytoplankton diversity

We found a distinct biogeographical distribution of phytoplankton diversity (Fig. 1A), which showed striking similarities with the topography of the USA (Fig. 1B). High phytoplankton diversity was found in lakes of the southern and eastern part of the United

[6] ⟨http://www.gpsvisualizer.com/elevation⟩

[7] ⟨http://cran.r-project.org/web/packages/sem/index.html⟩

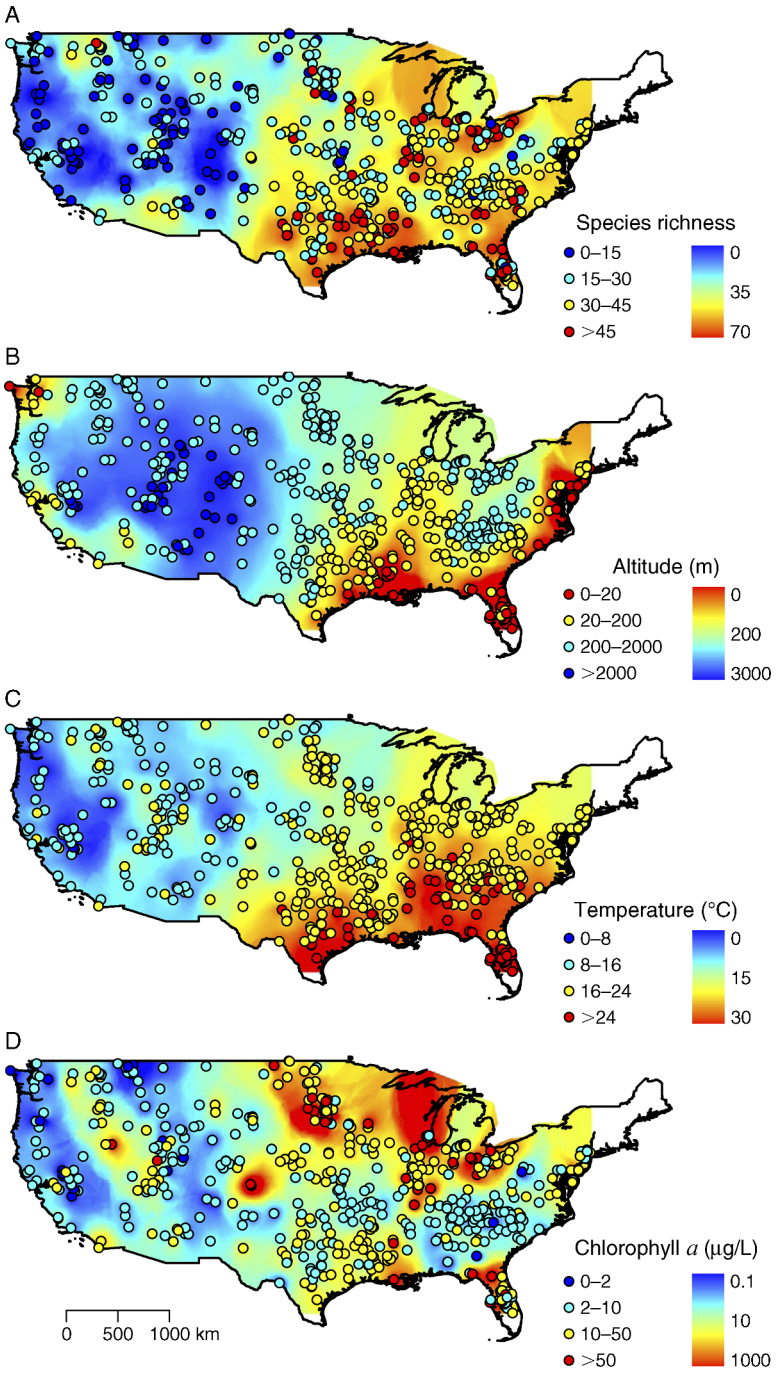

FIG. 1. Geographical distribution of phytoplankton species richness and associated explanatory variables across the continental United States: (A) phytoplankton species richness, (B) altitude, (C) annual mean water temperature, (D) annual mean chlorophyll *a* concentration. The spatial maps are based on data from 540 lakes (circles). The colors in all four maps represent spatially interpolated values obtained by Inverse Distance Weighting using ArcMap 9.2 (ESRI, Redlands, California, USA).

States, particularly in lowland areas such as in Florida and the Mississippi delta. Low diversity was found at higher altitudes in the mountainous west, and also in some lakes in the east (e.g., in the Appalachian Mountains; Fig. 1A, B). Indeed, linear regression showed that phytoplankton diversity decreased with latitude and altitude (Fig. 2A, C), but increased with longitude (Fig. 2B).

The biogeographical distribution of phytoplankton diversity (Fig. 1A) also showed many similarities with maps of local environmental variables such as water temperature (Fig. 1C) and lake chlorophyll *a* (Fig. 1D).

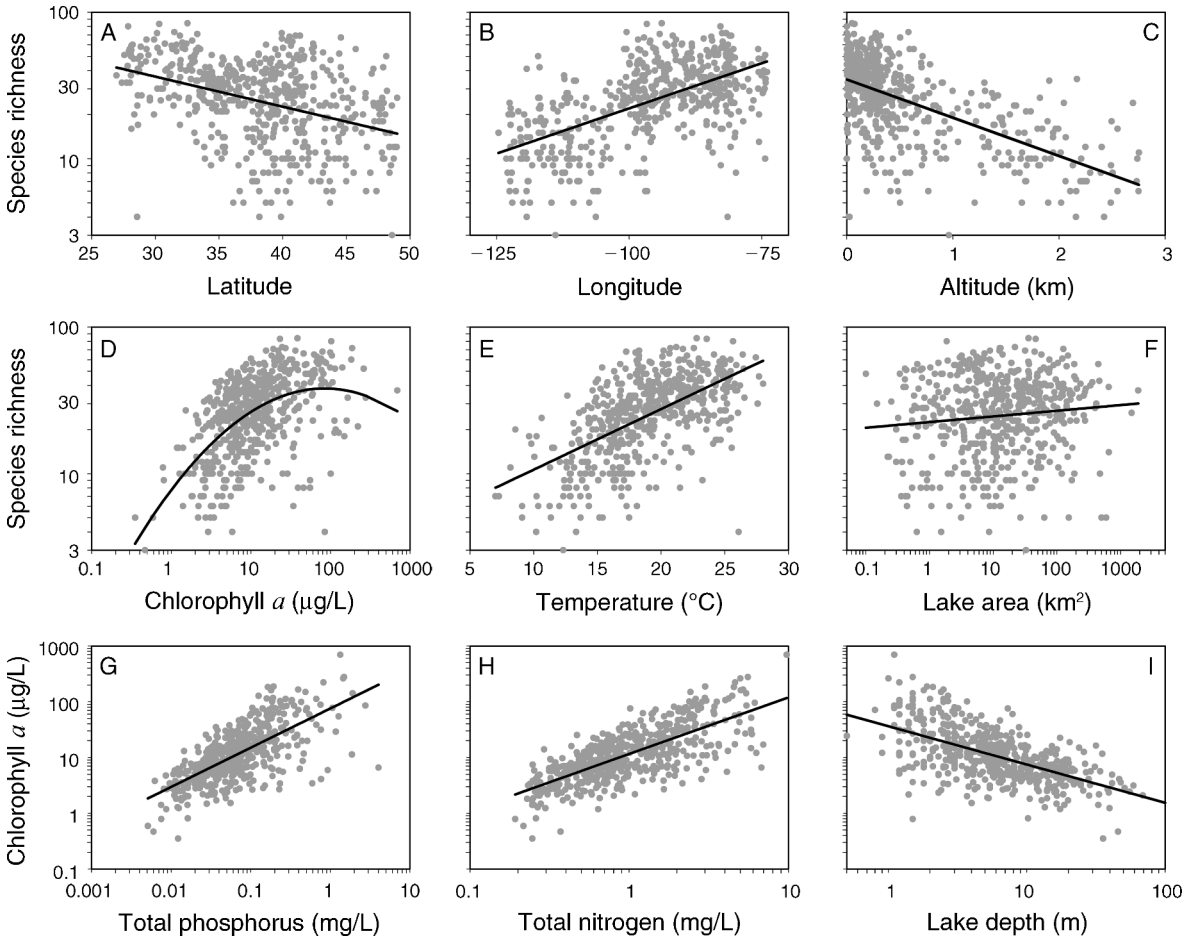

FIG. 2. Patterns of species richness and phytoplankton biomass. Species richness of freshwater phytoplankton as a function of (A) latitude, (B) longitude, (C) altitude, (D) productivity, (E) temperature, and (F) lake area. Phytoplankton biomass (measured as chlorophyll *a*) as a function of (G) total phosphorus concentration, (H) total nitrogen concentration, and (I) lake depth. Data points correspond to individual lakes. Regression lines show best linear fits (in A–C, E–I) and quadratic fit (D). All regression lines are highly significant ($P < 0.001$).

Multiple regression analysis confirmed that temperature, chlorophyll *a*, lake area, and lake depth all had significant effects on phytoplankton diversity, explaining more than 50% of the total variation in the data set (Table 1). More specifically, phytoplankton diversity showed an increasing saturating relationship with lake chlorophyll *a* (Fig. 2D), increased with temperature (Fig. 2E), increased with lake area (Fig. 2F), and decreased with lake depth (Table 1). Interestingly, nitrogen and phosphorus failed to enter the regression equation once lake chlorophyll *a* was included (Table 1), probably because the chlorophyll *a* concentration showed high collinearity with total phosphorus (Fig. 2G) and total nitrogen (Fig. 2H). Furthermore, chlorophyll *a* decreased with lake depth (Fig. 2I).

All local environmental variables showed large-scale variation across the landscape of the USA. For instance, temperature decreased with latitude and altitude but increased with longitude, chlorophyll *a* decreased with altitude but increased with longitude, and lake depth decreased with longitude (Table 2).

Drivers of phytoplankton diversity

We used a structural equation model to disentangle the complex network of direct and indirect effects on phytoplankton diversity. Our initial model included all plausible pathways between phytoplankton diversity, the local environmental variables and geographical coordinates (Fig. 3A). We trimmed the initial model by removing all non-significant pathways, until we arrived at a final model in which all pathways were significant (Fig. 3B; Appendix B: Table B2). The order of deletion of the nonsignificant pathways and their significance values are presented in Appendix B (Table B1). The adjusted goodness-of-fit index (AGFI) of the trimmed model was AGFI = 0.92, indicating a good fit of the model to the data (Schermelleh-Engel et al. 2003).

The path-coefficients of the final model showed that most environmental variables, except total phosphorus,

TABLE 1. Multiple regression analysis of phytoplankton species richness (dependent variable) as a function of environmental variables.

Regression variables	Partial regression coefficients	Standard error	Standardized coefficients	P
Constant	0.539	0.059		<0.001
Temperature	0.028	0.002	0.401	<0.001
Chlorophyll *a*	0.483	0.060	0.826	<0.001
(Chlorophyll *a*)^2	−0.120	0.024	−0.490	<0.001
Lake area	0.063	0.012	0.174	<0.001
Lake depth	−0.110	0.027	−0.163	<0.001
Model summary	$R^2 = 0.54$; $P < 0.001$			

Notes: We used stepwise multiple regression with forward selection of variables. Environmental variables investigated in the regression analysis were lake area (km^2), lake depth (m), water temperature (°C), total phosphorus (TP; mg/L), total nitrogen (TN; mg/L), the interaction term TN × TP, Secchi depth (m), chlorophyll *a* concentration (μg/L), and the square of chlorophyll *a* concentration. Environmental variables were selected only if $P < 0.001$ and are listed in the order of entry into the model. Species richness, lake area, lake depth, total phosphorus, total nitrogen, Secchi depth, and chlorophyll *a* concentration were log-transformed prior to the analysis.

TABLE 2. Multiple regression analysis of each environmental variable as function of geographical location.

Variables	Partial regression coefficients	Standard error	Standardized coefficients	P
Temperature (model $R^2 = 0.73$, $P < 0.001$)				
Constant	41.896	0.908		<0.001
Longitude	0.105	0.009	0.346	<0.001
Latitude	−0.312	0.019	−0.413	<0.001
Altitude	−1.916	0.176	−0.316	<0.001
Chlorophyll *a* (model $R^2 = 0.13$, $P < 0.001$)				
Constant	1.842	0.169		<0.001
Longitude	0.008	0.002	0.211	<0.001
Altitude	−0.131	0.038	−0.181	0.001
Lake area (model $R^2 = 0.10$, $P < 0.001$)				
Constant	1.07	0.318		0.001
Latitude	−0.041	0.007	−0.284	<0.001
Longitude	−0.018	0.003	−0.303	<0.001
Altitude	−0.211	0.062	−0.182	0.001
Lake depth (model $R^2 = 0.16$, $P < 0.001$)				
Constant	−0.424	0.119		<0.001
Longitude	−0.012	0.001	−0.397	<0.001
Secchi depth (model $R^2 = 0.18$, $P < 0.001$)				
Constant	−0.595	0.126		<0.001
Altitude	0.135	0.028	0.246	<0.001
Longitude	−0.006	0.001	−0.219	<0.001
Total nitrogen (model $R^2 = 0.30$, $P < 0.001$)				
Constant	0.48	0.124		<0.001
Longitude	0.011	0.001	0.433	<0.001
Latitude	0.017	0.003	0.258	<0.001
Altitude	−0.119	0.024	−0.231	<0.001
Total phosphorus (model $R^2 = 0.06$, $P < 0.001$)				
Constant	−1.095	0.025		<0.001
Altitude	−0.173	0.029	−0.247	<0.001
TN × TP (model $R^2 = 0.33$, $P < 0.001$)				
Constant	−0.577	0.148		<0.001
Longitude	−0.013	0.001	−0.426	<0.001
Altitude	0.168	0.029	0.269	<0.001
Latitude	−0.017	0.003	−0.219	<0.001

Notes: We used stepwise multiple regression with forward selection of variables. Geographical location was characterized by three regression variables: latitude, longitude, and altitude (km). The regression variables were selected only if $P < 0.001$ and are listed in the order of entry into the model. Lake area, lake depth, total phosphorus, total nitrogen, Secchi depth, and chlorophyll *a* concentration were log-transformed prior to the analysis.

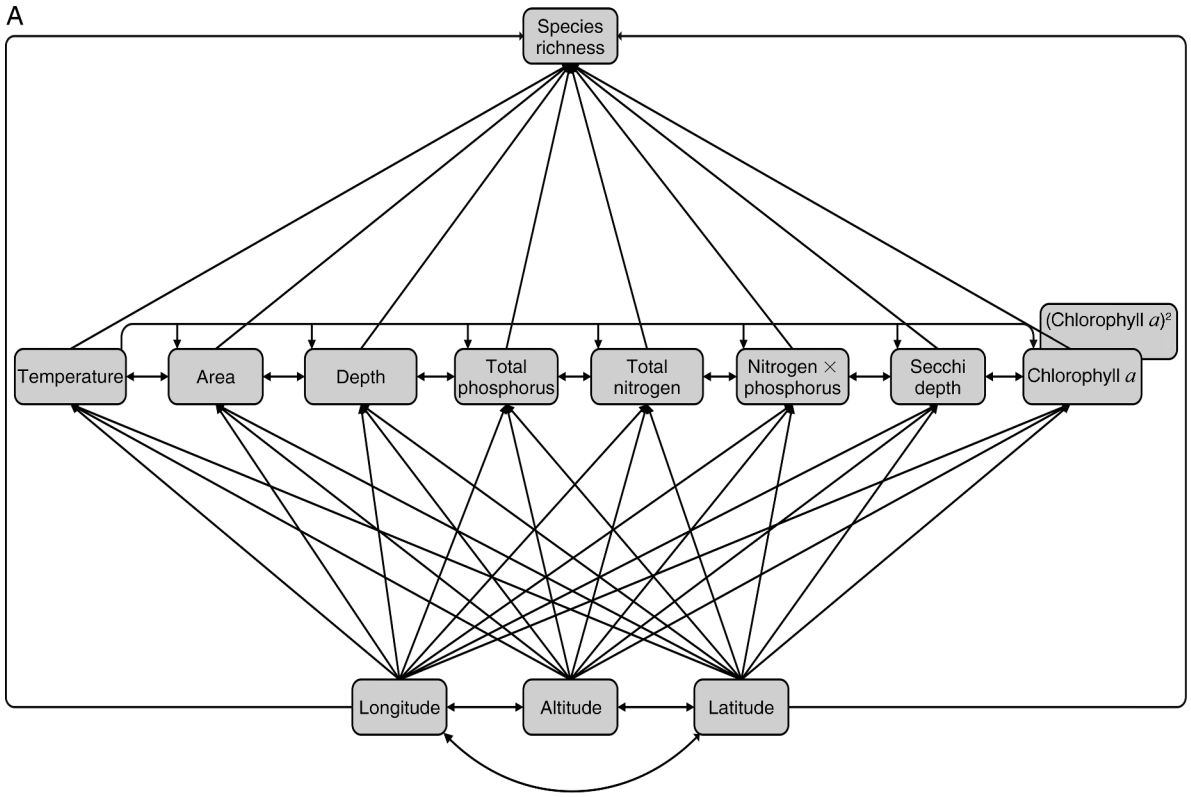

FIG. 3. Structural equation model explaining direct and indirect effects of latitude, longitude, and altitude on phytoplankton species richness. (A) Initial model illustrating all plausible interaction pathways in the study system. (B) Final model obtained by trimming of the initial model. (C) Final model obtained by trimming of the initial model without direct pathways from longitude, latitude and altitude to species richness. The values along the pathways represent standardized path coefficients, red arrows indicate negative effects, and blue arrows positive effects.

varied along latitudinal, longitudinal and/or altitudinal gradients. Furthermore, phytoplankton species richness increased with lake area, and showed a strong positive relationship with lake chlorophyll *a* (Fig. 3B; Appendix B: Table B2). In addition, lake chlorophyll *a* had a negative quadratic effect on species richness, consistent with the saturating relationship in Fig. 2D. The geographical variables latitude, longitude and altitude also had direct effects on species richness (Fig. 3B). Total phosphorus concentration and lake depth did not display a significant direct effect on phytoplankton diversity, but had indirect effects on phytoplankton diversity due to their effects on lake chlorophyll *a*. Interestingly, temperature did not have a significant effect on species richness in the structural equation model, neither directly nor indirectly, whereas it did have a significant effect in the multiple regression analysis.

To evaluate whether the direct effects of geographical variables (longitude, latitude and altitude) on species richness might suppress effects of environmental variables, we also performed a SEM analysis without direct pathways from the geographical variables to species richness (Fig. 3C). In this case, direct effects of temperature and lake depth on species richness became significant. Note that most other pathways in the two structural equation models (Fig. 3B, C) were very similar, which indicates that these pathways were robust.

DISCUSSION

The role of local environmental variables

Our results show that the diversity of freshwater phytoplankton displays strong biogeographical variation across the United States (Fig. 1, Fig. 2). Visual inspection of the data (Fig. 2D–F), multiple regression analysis (Table 1) and structural equation modeling (Fig. 3) all indicate that these large-scale biodiversity gradients are largely mediated by geographic variation in local environmental factors such as lake productivity (chlorophyll *a* concentration), lake area, and possibly lake depth and water temperature.

Of all variables included in our study, lake chlorophyll *a* had the largest effect on phytoplankton biodiversity. Moreover, both the multiple regression analysis and the structural equation model identified a significant quadratic term for lake chlorophyll *a*. This finding suggests either a saturating or unimodal productivity–diversity relationship, as has also been documented in many other

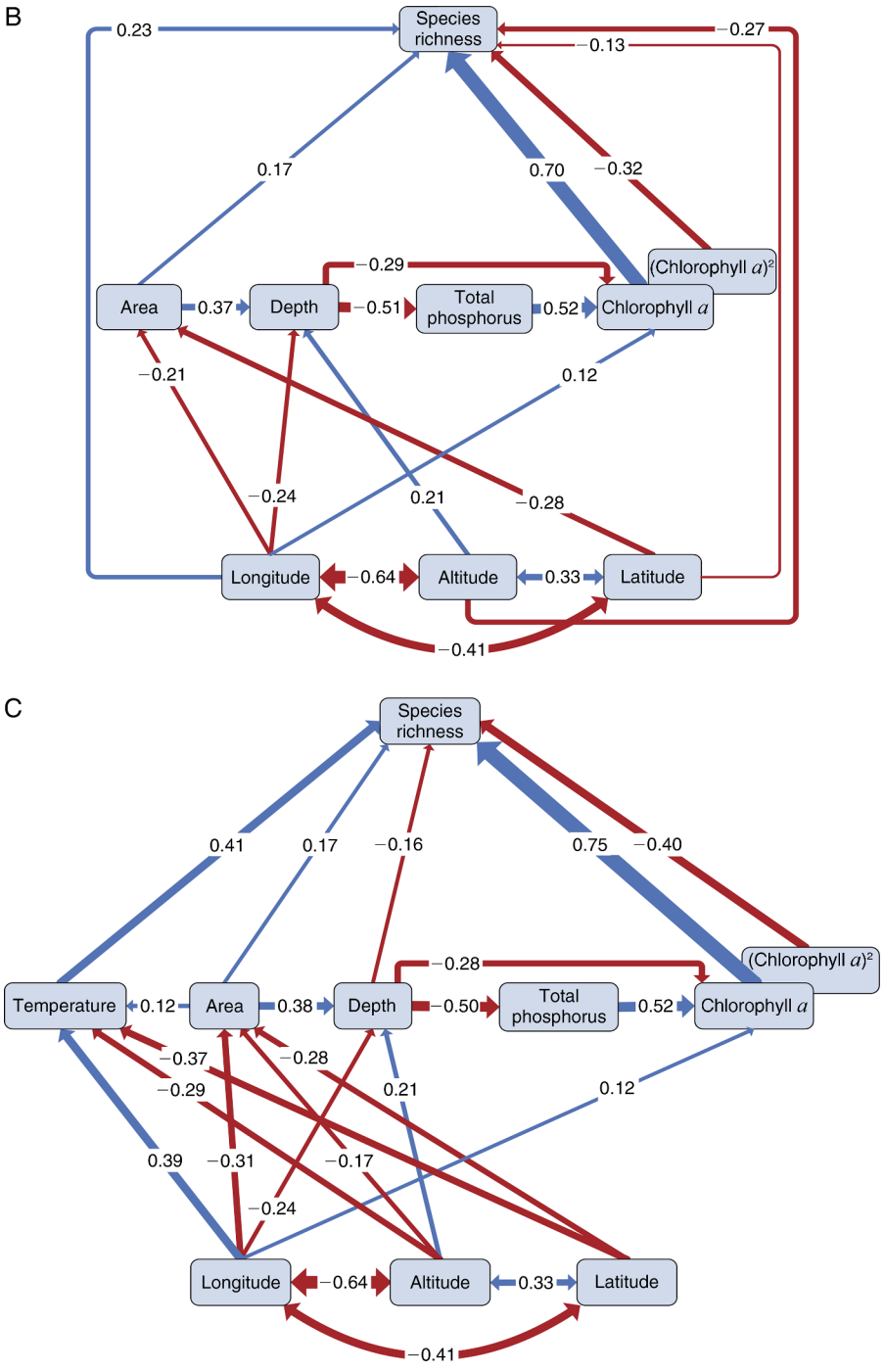

FIG. 3. Continued.

studies of terrestrial, marine and freshwater biodiversity (Grime 1973, Rosenzweig 1995, Waide et al. 1999, Dodson et al. 2000, Mittelbach et al. 2001, Irigoien et al. 2004). Earlier studies demonstrating unimodal productivity-diversity relationships in freshwater phytoplankton typically included very productive lakes (Dodson et al. 2000), whereas our data set contained only few highly productive waters with chlorophyll *a* concentrations > 100 μg/L. This may explain why a decrease of diversity at high chlorophyll *a* levels is not very pronounced in our data set (Fig. 2D). In turn, chlorophyll *a* concentration showed a strong positive relationship with total nitrogen and total phosphorus (Fig. 2G, H), which is expected because nitrogen and phosphorus are impor-

tant limiting nutrients in freshwater ecosystems (Schindler 1974, Elser et al. 2007). Furthermore, chlorophyll *a* concentration was negatively associated with lake depth (Fig. 2I). This is probably because deep mixing suppresses phytoplankton production due to light limitation (Petersen et al. 1997, Huisman 1999, Diehl et al. 2002) and because internal loading of nutrients from sediments into the water column is generally much higher in shallow lakes than in deep lakes (Wetzel 2001), as indicated by the negative relationship between total phosphorus and lake depth (Fig. 3B, C).

Phytoplankton diversity increased with lake area. This pattern resembles the common species-area relationship observed in many ecosystems (Arrhenius 1921, MacArthur and Wilson 1967, Rosenzweig 1995). According to the theory of island biogeography (MacArthur and Wilson 1967), large ecosystems are likely to harbor more species due to higher immigration rates and lower extinction rates. Indeed, lakes can be regarded as aquatic islands in a terrestrial world, thus offering an explanation for the positive species-area relation in our analysis. This finding is consistent with Horner-Devine et al. (2004) and Reche et al. (2005), who reported species–area relationships for bacterial communities in salt marshes and alpine lakes, respectively. Likewise, Smith et al. (2005) described a positive species–area relationship in phytoplankton communities across a broad size range from small experimental microcosms to ponds, lakes, and oceans.

Interestingly, some environmental factors did not reveal a consistent relationship with phytoplankton biodiversity. For instance, the multiple regression analysis (Table 1) points at a positive effect of water temperature on species richness. However, water temperature did not show a significant effect on species richness in the SEM analysis of Fig. 3B. This latter finding might be a bias arising from the structure of this SEM, as temperature showed by far the strongest statistical relationship ($R^2 = 0.73$; Table 2) with the geographical variables altitude, latitude, and longitude. Hence, temperature may have become a redundant variable when the SEM included direct effects of these geographical variables on species richness. Indeed, after removing the direct effects of these geographical variables from the model, temperature did have a significant positive effect on species richness (compare Figs. 3B and 3C). Alternatively, the lack of a strong temperature effect on phytoplankton species richness might be real. This would be in agreement with results of Cermeño et al. (2008), who found that water temperature did not have a significant effect on marine phytoplankton diversity. Yet, temperature is considered to be a good predictor of large-scale biodiversity patterns for many other species groups, including terrestrial plants and animals (Rohde 1992, Allen et al. 2002, Currie et al. 2004, Mittelbach et al. 2007) as well as marine organisms (Fuhrman et al. 2008, Rombouts et al. 2009, Tittensor et al. 2010). Diversity studies spanning a larger latitudinal gradient, from the poles to the tropics, may shed more light on the potential effects of temperature on phytoplankton species richness.

Our results also show a minor but significant negative effect of lake depth on phytoplankton species richness in the multiple regression analysis (Table 1), and in the SEM analysis without direct pathways from the geographical variables to species richness (Fig. 3C). This finding may reflect the sampling strategy used by the EPA, as deeper lakes may have additional biomass and perhaps additional diversity in deep chlorophyll *a* maxima not collected in samples from the upper 4.6 m. Alternatively, this negative relationship may indicate that shallow lakes host more species than do deep lakes. For example, phytoplankton communities of shallow lakes can be enriched with benthic algal species resuspended into the water column after storms (Schelske et al. 1995) or by activities of benthivorous fish disturbing the sediments (Roozen et al. 2007). Furthermore, some species that are commonly regarded as planktonic organisms may actually spend part of their life in the sediment (Preston et al. 1980, Karlsson-Elfgren and Brunberg 2004, Verspagen et al. 2005). In this way, lake sediments might offer additional niche space, contributing to higher phytoplankton biodiversity in shallow lakes, as anticipated by Hutchinson (1961) in his classic essay on the paradox of the plankton.

Limitations and caveats

A major contribution of our analysis is that we aimed to unravel the key environmental drivers of large-scale patterns in phytoplankton diversity using advanced statistical techniques like structural equation models. This approach revealed that geographical variation in phytoplankton diversity is not governed by a single master factor, e.g., temperature, but by a relatively complex network of multiple local environmental variables (Fig. 3).

The statistical analyses performed here are subject to the same caveats as simple regression analyses. Specifically, statistical relationships do not necessarily reveal the underlying mechanisms regulating phytoplankton biodiversity. For example, the structural equation model indicates that altitude has a direct negative effect on phytoplankton biodiversity (Fig. 3B), yet this relationship is probably driven by one or more environmental variables that covary with altitude but which were not measured in our study. For instance, a shorter growing season at high altitudes might reduce the number of species that can develop during seasonal succession (Litchman and Klausmeier 2001, Klausmeier 2010). Likewise, seasonal variation in environmental conditions increases at higher elevation, and could reduce species richness by excluding sensitive species with a narrow tolerance range (Currie et al. 2004). Reduced dispersal is an additional mechanism by which altitude may reduce phytoplankton diversity, because alpine

lakes are relatively isolated compared to most lakes at lower elevations (Stevens 1992, Vyverman et al. 2007). Moreover, organisms of high mountain lakes are challenged by high UV radiation and UV-sensitive species may get excluded, thus reducing diversity (Callieri et al. 2001). An important challenge for future studies is to disentangle these possible altitudinal effects on phytoplankton diversity.

Significant correlations between the geographical variables altitude, latitude, and longitude (Fig. 3B, C) may limit the statistical power of our analysis to distinguish between altitudinal, latitudinal and longitudinal diversity gradients. In particular, altitude co-varies strongly with longitude across the United States (Fig. 1B) due to mountain ranges of the Rocky Mountains and Sierra Nevada in the western part of the United States. This pattern makes it particularly difficult to separate altitudinal from longitudinal gradients. Similar biodiversity studies on other continents with less pronounced covariation of altitude and longitude (e.g., Europe, Asia) could offer a valuable addition to the data presented here.

Our analysis is also subject to common caveats concerning the relationship between measured and actual ("true") species diversity. For example, some taxa may be undetected because they are rare or show high levels of cryptic diversity (Hughes et al. 2001). Inevitably, our estimate of phytoplankton species richness is biased towards more easily detectable species and no doubt underestimates true species richness. However, a key strength of the EPA dataset is that the species of all lakes were identified and counted by the same investigators, which contributes to the internal consistency of the dataset. Further studies are needed to quantify the magnitude of the discrepancy between estimated and true species richness, possibly by comparison of microscopic counts vs. molecular techniques.

Conclusions and perspectives

Our study provides compelling evidence for strong latitudinal, longitudinal, and altitudinal gradients in freshwater phytoplankton diversity. This finding confirms recent studies that microorganisms can show substantial geographical variation in biodiversity (Martiny et al. 2006, Pommier et al. 2007, Fuhrman et al. 2008, Bryant et al. 2008, Ptacnik et al. 2010). In addition, we found that phytoplankton biomass (chlorophyll *a*) and lake area were important drivers of phytoplankton diversity in lakes, consistent with the impact of productivity and habitat area on diversity patterns observed in macroorganisms (MacArthur and Wilson 1967, Rohde 1992, Rosenzweig 1995, Mittelbach et al. 2001). The unique data set collected by the EPA provided the opportunity to examine patterns of phytoplankton diversity over an exceptionally large spatial scale. Yet, these data were gathered almost 40 years ago, and meanwhile the phytoplankton community structure in many of these lakes may have altered drastically due to changes in nutrient status, land use and climate change. It would therefore be highly interesting to reexamine phytoplankton diversity in these lakes. Such knowledge is crucial for our understanding of aquatic biodiversity, and we encourage long-term investment in similar large-scale biodiversity surveys in the next decades to assess impacts of worldwide changes in land use and climate.

Acknowledgments

We thank the U.S. EPA for providing the data set. In addition, we thank A. Hutchens for digitizing the EPA phytoplankton data, E. E. van Loon and J. Shamoun-Baranes for help with the spatial mapping, and the reviewers and members of the KBS microbial ecology discussion group for their constructive comments on the manuscript. The research was in part supported by the J. S. McDonnell Foundation to C. A. Klausmeier and E. Litchman. This is contribution #1589 from the Kellogg Biological Station.

Literature Cited

Allen, A. P., J. H. Brown, and J. F. Gillooly. 2002. Global biodiversity, biochemical kinetics, and the energetic-equivalence rule. Science 297:1545–1548.

Arrhenius, O. 1921. Species and area. Journal of Ecology 9:95–99.

Barton, A. D., S. Dutkiewicz, G. Flierl, J. Bragg, and M. J. Follows. 2010. Patterns of diversity in marine phytoplankton. Science 327:1509–1511.

Bryant, J., C. Lamanna, H. Morlon, A. J. Kerkhoff, B. J. Enquist, and J. L. Green. 2008. Microbes on mountainsides: contrasting elevational patterns of bacterial and plant diversity. Proceedings of the National Academy of Sciences USA 105:11505–11511.

Callieri, C., G. Morabito, Y. Huot, P. J. Neale, and E. Litchman. 2001. Photosynthetic response of pico- and nanoplanktonic algae to UVB, UVA and PAR in a high mountain lake. Aquatic Sciences 63:286–293.

Cermeño, P., and P. G. Falkowski. 2009. Controls on diatom biogeography in the ocean. Science 325:1539–1541.

Cermeño, P., E. Maranon, D. Harbour, F. G. Figueiras, B. G. Crespo, M. Varela, and R. P. Harris. 2008. Resource levels, allometric scaling of population abundance, and marine phytoplankton diversity. Limnology and Oceanography 53:312–318.

Currie, D. J., G. G. Mittelbach, H. V. Cornell, R. Field, J. F. Guégan, B. A. Hawkins, D. M. Kaufman, J. T. Kerr, T. Oberdorff, E. O'Brien, and J. R. G. Turner. 2004. Predictions and tests of climate-based hypotheses of broad-scale variation in taxonomic richness. Ecology Letters 7:1121–1134.

Diehl, S., S. A. Berger, R. Ptacnik, and A. Wild. 2002. Phytoplankton, light, and nutrients in a gradient of mixing depths: field experiments. Ecology 83:399–411.

Dodson, S. I., S. E. Arnott, and K. L. Cottingham. 2000. The relationship in lake communities between primary productivity and species richness. Ecology 81:2662–2679.

Elser, J. J., M. E. S. Bracken, E. E. Cleland, D. S. Gruner, W. S. Harpole, H. Hillebrand, J. T. Ngai, E. W. Seabloom, J. B. Shurin, and J. E. Smith. 2007. Global analysis of nitrogen and phosphorus limitation of primary producers in freshwater, marine, and terrestrial ecosystems. Ecology Letters 10:1135–1142.

Eppley, R. W., E. Stewart, M. R. Abbott, and U. Heyman. 1985. Estimating ocean primary production from satellite chlorophyll: introduction to regional differences and statistics for the Southern California Bight. Journal of Plankton Research 7:57–70.

Falkowski, P. G., and J. A. Raven. 1997. Aquatic photosynthesis. Blackwell Science, Malden, Massachusetts, USA.

Fenchel, T., and B. J. Finlay. 2004. The ubiquity of small species: patterns of local and global diversity. BioScience 54:777–784.

Fuhrman, J. A., J. A. Steele, I. Hewson, M. S. Schwalbach, M. V. Brown, J. L. Green, and J. H. Brown. 2008. A latitudinal diversity gradient in planktonic marine bacteria. Proceedings of the National Academy of Sciences USA 105:7774–7778.

Gaston, K. J. 2000. Global patterns in biodiversity. Nature 405:220–227.

Grace, J. B. 2006. Structural equation modeling and natural systems. Cambridge University Press, Cambridge, UK.

Grace, J. B., T. M. Anderson, H. Olff, and S. M. Scheiner. 2010. On the specification of structural equation models for ecological systems. Ecological Monographs 80:67–87.

Grace, J. B., and K. A. Bollen. 2005. Interpreting the results from multiple regression and structural equation models. Bulletin of the Ecological Society of America 86:283–295.

Grime, J. P. 1973. Control of species diversity in herbaceous vegetation. Journal of Environmental Management 1:151–167.

Hillebrand, H. 2004. On the generality of the latitudinal diversity gradient. American Naturalist 163:192–211.

Horner-Devine, M. C., M. Lage, J. B. Hughes, and B. J. M. Bohannan. 2004. A taxa–area relationship for bacteria. Nature 432:750–753.

Hughes, J. B., J. J. Hellmann, T. H. Ricketts, and B. J. M. Bohannan. 2001. Counting the uncountable: statistical approaches to estimating microbial diversity. Applied and Environmental Microbiology 67:4399–4406.

Huisman, J. 1999. Population dynamics of light-limited phytoplankton: microcosm experiments. Ecology 80:202–210.

Huisman, J. 2010. Comment on "Patterns of diversity in marine phytoplankton." Science 329:512.

Huisman, J., J. Sharples, J. M. Stroom, P. M. Visser, W. E. A. Kardinaal, J. M. H. Verspagen, and B. Sommeijer. 2004. Changes in turbulent mixing shift competition for light between phytoplankton species. Ecology 85:2960–2970.

Huisman, J., and F. J. Weissing. 1999. Biodiversity of plankton by species oscillations and chaos. Nature 402:407–410.

Hutchinson, G. E. 1961. The paradox of the plankton. American Naturalist 95:137–145.

Irigoien, X., J. Huisman, and R. P. Harris. 2004. Global biodiversity patterns of marine phytoplankton and zooplankton. Nature 429:863–867.

Karlsson-Elfgren, I., and A. K. Brunberg. 2004. The importance of shallow sediments in the recruitment of *Anabaena* and *Aphanizomenon* (Cyanophyceae). Journal of Phycology 40:831–836.

Klausmeier, C. A. 2010. Successional state dynamics: a novel approach to modeling nonequilibrium foodweb dynamics. Journal of Theoretical Biology 262:584–595.

Kline, R. B. 2005. Principles and practice of structural equation modelling. Second edition. Guilford Press, New York, New York, USA.

Leibold, M. A. 1996. A graphical model of keystone predators in food webs: trophic regulation of abundance, incidence and diversity patterns in communities. American Naturalist 147:784–812.

Litchman, E., and C. A. Klausmeier. 2001. Competition of phytoplankton under fluctuating light. American Naturalist 157:170–187.

MacArthur, R. H., and E. O. Wilson. 1967. The theory of island biogeography. Princeton University Press, Princeton, New Jersey, USA.

Martiny, J. B. H., et al. 2006. Microbial biogeography: putting microorganisms on the map. Nature Reviews Microbiology 4:102–112.

Mittelbach, G. G., et al. 2007. Evolution and the latitudinal diversity gradient: speciation, extinction and biogeography. Ecology Letters 10:315–331.

Mittelbach, G. G., C. F. Steiner, S. M. Scheiner, K. L. Gross, H. L. Reynolds, R. B. Waide, M. R. Willig, S. I. Dodson, and L. Gough. 2001. What is the observed relationship between species richness and productivity? Ecology 82:2381–2396.

Passy, S. I. 2008. Continental diatom biodiversity in stream benthos declines as more nutrients become limiting. Proceedings of the National Academy of Sciences USA 105:9663–9667.

Passy, S. I. 2010. A distinct latitudinal gradient of diatom diversity is linked to resource supply. Ecology 91:36–41.

Petersen, J. E., C. C. Chen, and W. M. Kemp. 1997. Scaling aquatic primary productivity: experiments under nutrient- and light-limited conditions. Ecology 78:2326–2338.

Pommier, T., B. Canbäck, L. Riemann, K. H. Boström, K. Simu, P. Lundberg, A. Tunlid, and Å. Hagström. 2007. Global patterns of diversity and community structure in marine bacterioplankton. Molecular Ecology 16:867–880.

Preston, T., W. D. P. Stewart, and C. S. Reynolds. 1980. Bloom-forming cyanobacterium *Microcystis aeruginosa* overwinters on sediment surface. Nature 288:365–367.

Ptacnik, R., T. Andersen, P. Brettum, L. Lepistö, and E. Willén. 2010. Regional species pools control community saturation in lake phytoplankton. Proceedings of the Royal Society B 277:3755–3764.

R Development Core Team. 2008. R version 2.8.1. R Project for Statistical Computing, Vienna, Austria. ⟨www.r-project.org⟩

Reche, I., E. Pulido-Villena, R. Morales-Baquero, and E. O. Casamayor. 2005. Does ecosystem size determine aquatic bacterial richness? Ecology 86:1715–1722.

Rohde, K. 1992. Latitudinal gradients in species diversity: the search for the primary cause. Oikos 65:514–527.

Rombouts, I., G. Beaugrand, F. Ibañez, S. Gasparini, S. Chiba, and L. Legendre. 2009. Global latitudinal variations in marine copepod diversity and environmental factors. Proceedings of the Royal Society B 276:3053–3062.

Roozen, F. C. J. M., M. Lürling, H. Vlek, E. A. J. Van der Pouw Kraan, B. W. Ibelings, and M. Scheffer. 2007. Resuspension of algal cells by benthivorous fish boosts phytoplankton biomass and alters community structure in shallow lakes. Freshwater Biology 52:977–987.

Rosenzweig, M. L. 1995. Diversity in space and time. Cambridge University Press, Cambridge, UK.

Rutherford, S., S. D'Hondt, and W. Prell. 1999. Environmental controls on the geographic distribution of zooplankton diversity. Nature 400:749–753.

Schelske, C. L., H. J. Carrick, and F. J. Aldridge. 1995. Can wind-induced resuspension of meroplankton affect phytoplankton dynamics? Journal of the North American Benthological Society 14:616–630.

Schermelleh-Engel, K., H. Moosbrugger, and H. Müller. 2003. Evaluating the fit of structural equation models: tests of significance and descriptive goodness-of-fit measures. Methods of Psychological Research 8:23–74.

Schindler, D. W. 1974. Eutrophication and recovery in experimental lakes: implications for lake management. Science 184:897–898.

Shipley, B. 2002. Cause and correlation in biology: a user's guide to path analysis, structural equations, and causal inference. Cambridge University Press, Cambridge, UK.

Smith, V. H., B. L. Foster, J. P. Grover, R. D. Holt, M. A. Leibold, and F. deNoyelles. 2005. Phytoplankton species richness scales consistently from laboratory microcosms to the world's oceans. Proceedings of the Natural Academy of Sciences USA 102:4393–4396.

Soininen, J., J. Heino, M. Kokocinski, and T. Muotka. 2009. Local-regional diversity relationship varies with spatial scale in lotic diatoms. Journal of Biogeography 36:720–727.

Sommer, U. 1990. Resource ratios or absolute concentrations: a reply to the comment by Prairie. Limnology and Oceanography 35:779–780.

Sommer, U. 1993. Phytoplankton competition in Plußsee: a field test of the resource-ratio hypothesis. Limnology and Oceanography 38:838–845.

Stehli, F. G., D. G. Douglas, and N. D. Newell. 1969. Generation and maintenance of gradients in taxonomic diversity. Science 164:947–949.

Stevens, G. C. 1992. The elevational gradient in altitudinal range: an extension of Rapoport's latitudinal rule to altitude. American Naturalist 140:893–911.

Stomp, M., J. Huisman, F. de Jongh, A. J. Veraart, D. Gerla, M. Rijkeboer, B. W. Ibelings, U. I. A. Wollenzien, and L. J. Stal. 2004. Adaptive divergence in pigment composition promotes phytoplankton biodiversity. Nature 432:104–107.

Stomp, M., J. Huisman, L. Vörös, F. R. Pick, M. Laamanen, T. Haverkamp, and L. J. Stal. 2007. Colourful coexistence of red and green picocyanobacteria in lakes and seas. Ecology Letters 10:290–298.

Taylor, W. D., L. R. Williams, S. C. Hern, V. W. Lambou, F. A. Morris, and M. K. Morris. 1979. Phytoplankton water quality relationships in U.S. lakes. Part I: methods, rationale, and data limitations. U.S. Environmental Protection Agency, Las Vegas, Nevada, USA.

Tilman, D. 1982. Resource competition and community structure. Princeton University Press, Princeton, New Jersey, USA.

Tittensor, D. P., C. Mora. W. Jetz, H. K. Lotze, D. Ricard, E. Vanden Berghe, and B. Worm. 2010. Global patterns and predictors of marine biodiversity across taxa. Nature 466:1098–1101.

Verspagen, J. M. H., E. O. F. M. Snelder, P. M. Visser, K. D. Jöhnk, B. W. Ibelings, L. R. Mur, and J. Huisman. 2005. Benthic-pelagic coupling in the population dynamics of the harmful cyanobacterium *Microcystis*. Freshwater Biology 50:854–867.

Vyverman, W., et al. 2007. Historical processes constrain patterns in global diatom diversity. Ecology 88:1924–1931.

Waide, R. B., M. R. Willig, C. F. Steiner, G. G. Mittelbach, L. Gough, S. I. Dodson, G. P. Juday, and R. Parmenter. 1999. The relationship between primary productivity and species richness. Annual Review of Ecology and Systematics 30:257–300.

Wang, J., J. Soininen, Y. Zhang, B. Wang, X. Jang, and J. Shen. 2011. Contrasting patterns in elevational diversity between microorganisms and macroorganisms. Journal of Biogeography 38:595–603.

Wetzel, R. G. 2001. Limnology. Third edition. Academic Press, San Diego, California, USA.

Willig, M. R., D. M. Kaufman, and R. D. Stevens. 2003. Latitudinal gradients of biodiversity: pattern, process, scale, and synthesis. Annual Review of Ecology, Evolution, and Systematics 34:273–309.

APPENDIX A

References describing EPA methods and their data (*Ecological Archives* E092-183-A1).

APPENDIX B

Development of the structural equation model (*Ecological Archives* E092-183-A2).

SUPPLEMENT

Data on phytoplankton species richness and environmental variables used in this study (*Ecological Archives* E092-183-S1).