RUNNING TITLE: **Richness and abundance distribution**

The biogeographical patterns of species richness and abundance distribution in stream diatoms are driven by climate and water chemistry

Sophia I. Passy[1]*, Chad A. Larson[2], Aurélien Jamoneau[3], William Budnick[1], Jani Heino[4], Thibault Leboucher[3], Juliette Tison-Rosebery[3], Janne Soininen[5]

[1]Department of Biology, University of Texas at Arlington, Box 19498, Arlington,
Texas 76019-0498, USA
[2]Washington State Department of Ecology, Environmental Assessment Program,
300 Desmond Drive SE, Lacey, WA 98503
[3]Aquatic Ecosystems and Global Changes Research Unit, IRSTEA, 50 avenue de Verdun,
33612 Cestas, France
[4]Finnish Environment Institute, Biodiversity Centre
P.O. Box 413, FI-90014 Oulu, Finland
[5]Department of Geosciences and Geography, University of Helsinki, PO Box 64, Helsinki,
FIN-00014, Finland
*Corresponding author, e-mail: sophia.passy@uta.edu
phone: +1 (817)-272-2415, fax: +1 (817)-272-2855

Keywords: climatic tolerance hypothesis, energy variability hypothesis, latitudinal diversity gradient, metabolic theory, species abundance distribution, species-energy theory

Abstract

In this inter-continental study of stream diatoms, we asked three important but still unresolved ecological questions: 1) What factors drive the biogeography of species richness and species abundance distribution (SAD); 2) Are climate-related hypotheses, which have dominated the research on the latitudinal and altitudinal diversity gradients, adequate in explaining spatial biotic variability; and 3) Is the SAD response to the environment independent of richness? We tested a number of climatic theories and hypotheses (i.e., the species-energy theory, the metabolic theory, the energy variability hypothesis, and the climatic tolerance hypothesis) but found no support for any of these concepts as the relationships of richness with explanatory variables were non-existent, weak or unexpected. Instead, we demonstrated that diatom richness and SAD evenness generally increased with temperature seasonality and at mid- to high total phosphorus concentrations. The spatial patterns of diatom richness and the SAD—mainly longitudinal in the US, but latitudinal in Finland—were defined primarily by the covariance of climate and water chemistry with space. The SAD was not entirely controlled by richness, emphasizing its utility for ecological research. Thus, we found support for the operation of both climate and water chemistry mechanisms in structuring diatom communities, which underscores their complex response to the environment and the necessity for novel predictive frameworks.

Introduction

The spatial variability in species richness along latitudinal, longitudinal, and elevational gradients has drawn a continued interest among ecologists for centuries, beginning with the works of Alexander von Humboldt and Carolus Linnaeus (Pianka 1966; Gaston 2000; Willig et al. 2003; Rahbek 2005; Mittelbach et al. 2007). There are numerous hypotheses about the nearly universal poleward decline in species richness with ecological and evolutionary rates, energy, and climatic favorability and stability emerging as important underlying factors. The elevation-richness relationship is more variable, most frequently conforming to a monotonically declining or a hump-shaped form, but there are some commonalities in the driving forces of the latitudinal and elevational diversity gradients (Rahbek 2005). Thus, warmer temperatures in tropical and low elevation regions shorten the generation times and accelerate the metabolic and mutation rates, leading to greater speciation (Rohde 1992; Currie et al. 2004; Allen et al. 2006).

There are several prominent climate-based frameworks explaining the geographic variability in richness. The species-energy theory proposed that areas with greater total available energy, such as the tropics, have more diverse communities because they can maintain larger species populations with lower likelihood of extinction (Wright 1983). This theory is also known as the more individuals hypothesis, because larger species populations result in a greater total number of individuals (Srivastava and Lawton 1998). It predicts that richness scales positively with energy because richness is a function of the number of individuals, which is proportional to the available energy, i.e. communities are energy-limited. Further elaboration of the species-energy theory recognized that temporal variability in energy may directly impact the richness of an area, given that periods of low energy support fewer individuals and are consequently prone to higher extinction rates. Thus, models that included both total energy and energy variability explained the patterns in bird and mammal species richness better than models using a single energy variable (Carrara and Vazquez 2010). The Metabolic Theory of Ecology (MTE) predicted a positive relationship of species richness with temperature (described in more detail in Theory testing below) because higher temperatures increase the rates of speciation (Allen et al. 2002; Brown et al. 2004).

Finally, according to the climatic tolerance hypothesis, the tropics harbor greater richness because their more benign warm and humid conditions fall within the physiological tolerance of many more species compared to the cold and dry extra-tropical regions (Currie et al. 2004). However, the latitudinal species richness patterns can be complicated by longitudinal effects, such as an east-west heterogeneity in rainfall, which is particularly distinct at mid-latitudes (Terborgh 1973). In the US, longitudinal effects on biodiversity are expected to emerge as a result of a strong ocean influence, whereby coastal regions, especially along the Pacific, experience much milder temperatures than inland regions of the same or lower latitudes (http://planthardiness.ars.usda.gov/PHZMWeb/).

There are, notably, deviations from the classical latitudinal diversity pattern. For example, a bimodal latitudinal distribution of species richness in aquatic systems has been attributed in part to resource supply and productivity (Passy 2010; Chaudhary et al. 2016). Since temperature variability and severity in aquatic systems are much lower compared to terrestrial habitats, it is conceivable that aquatic communities are less sensitive to climate. On the other hand, global changes in water chemistry and primary production as a result of anthropogenic eutrophication have strong ecological and evolutionary consequences (Smith and Schindler 2009; Alexander et al. 2017). Indeed, a comprehensive review of the freshwater literature concluded that water chemistry was a stronger predictor of diatom distributions than temperature (Soininen 2007). Therefore, factors other than climate may have profound influence on aquatic biogeography and merit further research.

While the spatial variability of species richness has been extensively studied, we know substantially less about the spatial and environmental dependence of the species abundance distribution (SAD), defined here as the number of individuals across species in a community. The SAD underlies broadly studied macroecological patterns, including the relationships of number of species with area and with number of individuals (Preston 1962; May 1975; Keeley 2003; McGill et al. 2007) and is, therefore, of fundamental significance in ecology. However, due to a more theoretical and statistical emphasis in the study of the SAD over the past seventy years (May 1975; McGill et al. 2007; Ulrich et al. 2010), the empirical latitudinal, longitudinal, and elevational patterns of the SAD and their underlying mechanisms, including environmental variability, have remained largely unknown (Matthews et al. 2017).

It is possible that the latitudinal gradient of richness is paralleled by a corresponding gradient of the SAD. Brown (2014) suggested that as richness increases toward lower latitudes, the SAD may transition from a less even log linear (a few very abundant and some very rare species) to a more even curvilinear pattern (greater numbers of both intermediate and low abundance species) as a result of intensified enemy effects. There is evidence that the SADs of forest communities do become less even at high latitudes due to strong environmental filtering that favors only a few well-adapted species (Qiao et al. 2015). However, there is no broad empirical support for a tendency in the SAD toward greater evenness and lognormality at low latitudes. On the contrary, a recent study of forest trees showed the opposite trend—preponderance of log-series (i.e., less even) SADs at low latitudes but lognormal (i.e., more even) SADs at high latitudes, possibly because of overall greater dispersal in the tropics, and subsequent accumulation of rare species (Ulrich et al. 2016a). A global study of dryland plant communities reported an overall prevalence of lognormal compared to log-series SADs, which was linked to environmental variability and stress but not to latitude (Ulrich et al. 2016b). There are even fewer and generally terrestrial studies of the SAD along elevational gradients, which either have not reported any distinct patterns (Ulrich et al. 2016a; 2016b) or shown a transition from log-series to lognormal SADs with altitude (Arellano et al. 2017). Therefore, there is clearly a need for further research on the large-scale spatial patterns of the SAD, especially in aquatic ecosystems, which have been largely neglected in this context.

Our overall goal was to explore the latitudinal, longitudinal, and altitudinal patterns in diatom species richness and abundance distribution and, given the correlation between species richness and the SAD (Locey and White 2013; Passy 2016), determine if they are driven by the same environmental factors and whether these factors are climatic and/or chemical. To achieve this goal, we tested several climate-based hypotheses, water chemistry models, and climate + water chemistry models (Table 1).

Then, we implemented variance partitioning to assess if the sources of spatial variation in species richness and the SAD are climatic, chemical or both.

We examined the shape of the SAD by calculating the standard deviation (parameter σ) of a Poisson lognormal distribution model and skewness (fig. 1). Parameter σ indicates how equitably abundances are distributed across abundance classes with lower values suggesting higher equitability. However, it has been empirically shown for freshwater diatoms that parameter σ is inversely related to species richness (Passy 2016). To determine whether spatial and environmental factors control the shape of the SAD independently of species richness, and if so, whether species richness and the SAD have common environmental underpinning, we explored the response of parameter σ to space and environment after partialling out the effect of species richness. Additionally, we assessed skewness, which measures the symmetry of the SAD compared to a lognormal distribution. Negative values (left skew) indicate prevalence of rare species, while positive values (right skew), greater frequency of abundant species.

In summary, we had the following objectives with respect to species richness and the SAD: i) assess their spatial patterns and underlying environmental variability; ii) examine their responses to climate within the framework of several climate-based theories and hypotheses (i.e., the species-energy theory, the metabolic theory, the energy variability hypothesis, and the climatic tolerance hypothesis) and determine whether they are driven by climate and/or water chemistry; and iii) evaluate if the SAD responds to spatial and environmental factors independently of species richness.

Materials and Methods

Datasets

US. Data on stream water chemistry and diatom composition were collected from 526 distinct stream localities in the US (fig. A1) by the National Water-Quality Assessment (NAWQA) Program (http://water.usgs.gov/nawqa). Diatoms were sampled from a defined area of hard substrate or macrophytes. Water chemistry, including total phosphorus, nitrate + nitrite, ammonia, pH, and specific conductance, was measured for the month of algal collection (Table A1). Samples were taken in July and

August from 1993 to 2009 along a latitudinal range of 36 degrees, longitudinal range of 83 degrees, and altitudinal range of 2448 m. Climatic variables, including mean annual temperature (T_{mean}), temperature seasonality (standard deviation, T_{SD}), minimum temperature of the coldest month (T_{min}), and maximum temperature of the warmest month (T_{max}), were obtained from the WorldClim database (Hijmans et al. 2005). In each sample, about 600 diatom cells were counted and identified primarily to species. The total cell count was converted to total density ($cells \cdot cm^{-2}$).

Finland. Data on climate (the same variables as in the US) and water chemistry, including total phosphorus, pH, and specific conductance, were available for 100 streams in Finland (fig. A1). Diatoms were sampled from a total area of 90 cm^2 of stream substrate in July and August from 2001 to 2004 along a latitudinal range of 10 degrees, longitudinal range of 7 degrees, and altitudinal range of 302 m. Diatoms were identified primarily to species in counts of about 500 cells. In both the US and Finland, the numbers of counted cells were consistent with international protocols; therefore, we are confident we have good estimates of community species richness.

Analysis of the species abundance distribution

Using cell counts, the SAD of each community was fit with a Poisson lognormal distribution model (fig. 1), which is comparatively insensitive to sampling effort and performs equally well to other commonly used models (Sæther et al. 2013; Baldridge et al. 2016). We estimated the standard deviation, parameter σ, with the poilog R-package (Grøtan and Engen 2008). A comparison of the rank of the observed log likelihood with the log likelihood derived from 1000 bootstraps provided a goodness of fit metric. Goodness of fit values between 0.05 and 0.95 indicate good fit. We also calculated the proportion of the species pool revealed by the sample, which represents the unveiled proportion of the Poisson lognormal distribution. We calculated the skewness (γ) of the $\log_2$-transformed counts of individuals and the standard error of skewness (SES) as $(6/n)^{0.5}$, where n = number of species. Skewness is considered significant if the absolute value of the ratio γ/SES is greater than 2 (SYSTAT Software, Inc. 2009).

Environmental and biotic data are deposited in Dryad Digital Repository: https://datadryad.org/resource/doi:10.5061/dryad.v1v7856 (Passy et al. 2018).

Theory testing

We performed a series of regression analyses to explore the spatial patterns in species richness and the SAD and test the outlined theories and hypotheses (Table 1). To assess whether the SAD had responses to spatial and environmental predictors that were independent of species richness, we regressed parameter σ against richness (fig. A2), obtained the residuals (σ_{res}), and then treated σ_{res} as a dependent variable in all subsequent regressions of the SAD. If richness controlled the behavior of the SAD along spatial and environmental gradients, then these gradients would have little to no effect on σ_{res}. Considering that spatial trends could be more complex due to patchiness, the spatial predictors included linear, quadratic, and cubic terms. The environmental predictors, on the other hand, encompassed only linear and quadratic terms.

In Tables 2 and 3, models 1-4 tested the spatial effects on species richness and σ_{res}, models 5-8, climate-related theories and hypotheses, models 9-11, climatic and/or chemistry effects, and models 12-13, the metabolic theory. We tested the predictions of the species-energy theory that species richness (S) increases with energy (here mean temperature, T_{mean}) because S is a positive function of the number of individuals (N), which in turn, is proportional to the amount of energy (Srivastava and Lawton 1998). Testing this required several equations, including models 5 and 6 from Table 1, and equation (1) below, which contained a quadratic term to account for potential nonlinearity in the density (N) response to mean temperature (T_{mean}).

$$\ln(\mathrm{N}) = \ln(T_{\mathrm{mean}}) + \ln(T_{\mathrm{mean}})^2 \qquad (1)$$

The metabolic theory expresses species richness (S) as a function of temperature, according to the equation $\ln(S) = -E_A/kT + I$, where E_A = activation energy with an expected value of about −0.65, k = Boltzmann constant, 8.62×10^{-5} eV K^{-1}, T = temperature in Kelvin, and I = intercept (Allen et al. 2002; Brown et al. 2004). Here, we used T_{mean} and tested if this prediction holds and $-E_A$ lies between −0.6 and

−0.7. However, since the richness response to the Boltzmann temperature factor ($1/kT_{mean}$) can be curvilinear (Algar et al. 2007), we added a quadratic term (model 13).

For each multiple regression model, we performed backward selection of significant terms only to reduce redundancy and collinearity (Tables 2, 3). All models were compared using the Akaike Information Criterion (AIC). Models with lower AIC provided better fits. Regression trees were calculated to assess interactions among climatic and water chemistry predictors and potential non-linear responses of species richness and σ_{res}. Variance partitioning estimated whether the spatial responses of species richness and σ_{res} were driven by climate, water chemistry, and/or their covariance. Specifically, we estimated how much of the variance explained by the overall spatial model (model 4, Table 1) was contributed by covariance of the spatial predictors with climatic and chemistry predictors from models 9 and 10, respectively. The effects of climatic and water chemistry predictors on skewness were examined with multiple regressions using backward stepping procedure. Regressions, regression trees, and variance partitioning were performed with SYSTAT 13 (SYSTAT Software, Inc. 2009).

Results

Spatial patterns of environmental factors, species richness, and the SAD

Temperature-related variables and, to a lesser extent, water chemistry variables exhibited complex spatial patterns (figs 2A-D, 3A-E, A3-A6). For example, in the conterminous US, the Upper Midwest experienced low mean and minimum temperatures and high temperature seasonality, while coastal areas, even of similar latitude, had much greater T_{mean} and T_{min} and much lower T_{SD}. Total phosphorus and specific conductance peaked at mid-longitudes. In Finland, total phosphorus and all climatic variables but T_{SD} declined with latitude and altitude, while T_{SD} was the highest at intermediate latitudes and the highest altitudes. Regressions including linear, quadratic, and cubic terms of latitude, longitude, and altitude explained 72-95% of the variability in the four climatic variables in the US and 94-99% in Finland, but 14-30% of the variability in water chemistry in the US and 67-71% in Finland.

In both the US and Finnish datasets, the Poisson lognormal distribution model fit the abundance data well ($0.36 \leq$ goodness of fit ≤ 0.53). Therefore, this distribution model provided a reliable estimate of the species abundance variability and was used in further analyses to determine the drivers of the SAD. Parameter σ was a negative function of the ln-transformed species richness (fig. A2), i.e. rich communities had lower parameter σ and were thus more equitable. To assess the richness-independent effects of spatial and environmental predictors on the SAD, we calculated residual parameter σ (σ_{res}) from the regression of parameter σ against ln(S) and treated it as a dependent variable in subsequent regressions (Tables 2, 3, fig. 4B, D). Richness and parameters σ and σ_{res} displayed broad variability in both the US and Finland (fig. A7).

In the US, the most pronounced spatial trend in species richness and σ_{res} was longitudinal (fig. 2E, F), while latitude and altitude generated weaker responses. The highest richness and the lowest σ_{res} were detected in streams with longitudes between approximately −87 and −97 degrees across a wide range of latitudes, but primarily concentrated in the Midwest (figs 2E, F, A4C, D). In Finland, latitude was the only spatial gradient with a prominent effect on both species richness and σ_{res} (fig. 3F, G). The best spatial model in the US, especially for richness, included all spatial predictors (but the longitudinal terms had the highest standardized regression coefficients) and captured 14-17% of the biotic variance (Table 2, model 4). In Finland, latitude and longitude captured 24% of the variance in species richness, while just latitude explained 15% of the variance in σ_{res} (Table 3, model 4).

Responses of species richness and the SAD to environmental factors and their covariance with space

To assess whether variability in mean temperature, temperature seasonality or temperature extremes contributed the most to the variability in richness, as predicted by the species-energy theory, the energy variability hypothesis, and the climatic tolerance hypothesis, respectively, we calculated models 5, 7, and 8 (Tables 2, 3). Additionally, in the US we examined the response of richness to density to more fully evaluate the species-energy theory (Table 2, model 6). Temperature seasonality and temperature extremes emerged as stronger predictors of richness in both countries and σ_{res} in the US than mean temperature, while mean and extreme temperatures defined σ_{res} in Finland equally well. In the US, the climatic variables generated a more or less pronounced U-shaped response in richness, contrary to the predictions of the aforementioned theories and hypotheses, but primarily a linear response in σ_{res} (Table 2, fig. A8). A non-significant relationship of richness with density further indicated that the effect of temperature on richness was not due to a temperature dependence of density, which is inconsistent with the species-energy theory. The relationship of density with mean temperature was weak and unimodal ($R^2 = 0.03$, $n = 524$, $p < 0.00005$). In Finland, the greatest richness was observed at intermediate T_{max} and high T_{SD}, and the lowest σ_{res}, at the highest T_{mean} and T_{max} (Table 3, fig. A9). Despite high correlations among the climatic variables (Table A2), the model including all significant climatic variables (Tables 2 and 3, model 9) improved to various extents the predictability of richness and σ_{res}, indicating that in some cases, these community properties were products of multiple climatic influences.

In the US, water chemistry captured a greater proportion of the variance in species richness than climate, whereas climate outperformed water chemistry in the remaining models of richness in Finland and σ_{res} in both countries (Tables 2, 3, models 9, 10). In both countries, total phosphorus was the best water chemistry predictor of richness, and specific conductance, of σ_{res} (figs A8, A9). Notably, the best model in both countries included both climate and water chemistry variables and explained 19-38% and 13-22% of the variance in species richness and σ_{res}, respectively (Tables 2, 3, model 11).

The relationship of ln-species richness with the Boltzmann temperature factor ($1/kT_{\text{mean}}$) was not significant in the US and convex in Finland, inconsistent with the prediction of the MTE (Tables 2, 3 models 12, 13). Only in Finland, σ_{res} exhibited a notable relationship with the Boltzmann temperature factor (negative).

Regression tree analyses of the US data revealed the highest species richness and the lowest σ_{res} at high temperature seasonality (fig. 4A, B). At lower temperature seasonality, richness was greater at higher total phosphorus levels. In Finland, rich and poor communities were separated only by temperature seasonality with rich communities found at higher seasonality (fig. 4C). Parameter σ_{res} was differentiated by mean temperature and specific conductance (fig. 4D). The lowest σ_{res} (highest equitability) was detected at high mean temperature. At lower T_{mean}, streams of lower conductance had lower σ_{res}.

We asked next to what extent the spatial patterns of richness and the SAD were driven by climate vs. water chemistry. To answer this question, we performed variance partitioning (fig. 5), which revealed that in both countries the spatial effect on both species richness and σ_{res}, which captured 14-24% of their variance, was generated by covariance of space with climate and water chemistry (4-12% explained variance), followed by covariance of space with climate (4-10% explained variance).

Skewness was positive in all communities in both the US and Finland indicating a prevalence of common species. Of the skewness values, 51% were significant (γ/SES > 2) in the US and 71% in Finland. Skewness was either negatively (Pearson $r = -0.24$, $p < 0.000001$, the US) or non-significantly (Finland) related to richness. In both countries, skewness was positively correlated with σ_{res} (Pearson r = 0.38-0.52, $p < 0.00002$), i.e. equitable communities were more symmetric with lower skewness and these communities tended to be rich in the US. In the US, skewness responded primarily to climate but weakly (Table A3). In Finland, skewness was determined by climatic variables alone and declined at high values of mean, minimum, and maximum temperature.

Discussion

On both continents, diatom richness and the SAD exhibited distinct spatial patterns, which were attributed to spatially structured climate and water chemistry. By examining different environmental factors, we were able to test several climate-based theories and hypotheses for the spatial variability in richness, and ultimately, to develop a climate-water chemistry model (Tables 2 and 3, model 11), outperforming existing frameworks (Tables 2 and 3, models 5-8, 12-13). A similar model was formulated for the SAD, which exhibited variability along environmental and spatial gradients that was independent of richness. Next, we give an overview of the spatial patterns of richness and the SAD and discuss their potential origins.

In accordance with our first objective, we report a number of interesting findings about the spatial variability in species richness and the SAD. First, in the US, no single spatial factor captured much of the variance in richness, but a more complex polynomial model of latitude, longitude, and altitude was necessary to better describe richness variability. In this model, the strongest spatial effect was longitudinal. The most pronounced spatial gradient of the SAD in the US was also longitudinal. Second, the best spatial predictors of richness in Finland were latitude and longitude, but only latitude had a comparatively strong effect on the SAD. Third, both in the US and Finland, diatom richness did not conform to the classical pattern of monotonic latitudinal decrease, consistent with prior diatom studies (Passy 2010; Soininen et al. 2016). These studies attributed the deviation from the classical pattern to the overriding effect of wetlands and their impact on micronutrient availability. Here, we identify both climatic and water chemistry factors that further contribute to the unique spatial distribution of diatom richness. Fourth, the expectation for increased evenness in the SAD at low latitudes (Brown 2014) was confirmed only in Finland, while in the US, the most equitable communities were observed at mid-latitudes, where specific conductance, temperature seasonality, and total phosphorus were the greatest. Fifth, the altitudinal response of richness—bimodal in the US but peaking at high elevations in Finland—did not follow the common monotonic decline or hump-shaped patterns. Admittedly, the elevational gradient in Finland was short and not sufficient to reveal the full variability in richness. By examining different sources of climatic and water chemistry variability, here we provide a more comprehensive explanation for the non-conventional diatom spatial patterns.

Similar to species richness and the SAD, the four studied climatic factors—mean, minimum, and maximum temperature, and temperature seasonality, and to a lesser extent, water chemistry (e.g., total phosphorus and specific conductance), displayed complex spatial distributions, generally driven by all three spatial gradients—latitudinal, longitudinal, and altitudinal, but to a various degree. To assess which environmental factors contributed to the spatial structuring of species richness and the SAD and what biotic responses they generated, we performed a series of multiple regressions, regression tree analysis, and variance partitioning.

In pursuit of our second objective, we tested climatic and water chemistry effects on species richness and the SAD. In the US, species richness was constrained most strongly by temperature seasonality and total phosphorus—the highest richness was recorded in streams with the highest seasonality, whereas in streams of lower seasonality, high levels of total phosphorus promoted greater richness (fig. 4A). Temperature seasonality was a comparatively strong and positive predictor of richness in Finland. In fact, in Finland, the latitudinal richness distribution, with a mid-latitude maximum, was best approximated by the latitudinal distribution of temperature seasonality, also reaching high values at mid-latitudes. These results contradict the energy variability hypothesis, predicting a negative effect of temperature seasonality on richness due to reduced energy availability during the cold months. A possible explanation for this unexpected pattern can be derived from competition theory, which postulates that temporal heterogeneity in resource supply increases biodiversity because it allows coexistence of species that alternate between dominance and persistence at different times (Sommer 1985; Tilman and Pacala 1993). The shape of the SAD in the US was also determined by temperature seasonality—communities with higher abundance equality were found in streams of higher seasonality. This finding provides further evidence that the increased species coexistence under variable climatic conditions is potentially maintained by resource partitioning and diminished competition.

In the US, minimum temperature was among the strongest predictors of richness, which showed a U-shaped response with high values at the lowest minimum temperature. In Finland, richness was the highest at lower minimum temperature. These results contradict the expectation of the climatic tolerance hypothesis for reduced richness at low temperature minima. Given that low minimum temperature in this study correlated with high temperature seasonality (Table A2), these unexpected richness patterns are likely a consequence of the positive effect of temperature seasonality on richness. It is also possible that diatoms deviate from the expectation of the climatic tolerance hypothesis because they inhabit the comparatively milder aquatic environment, less prone to extreme fluctuations. Conversely, in terrestrial habitats, where temperatures reach much greater extremes, the climatic tolerance hypothesis was supported (Šímová et al. 2011). Furthermore, low temperatures may favor diatoms over cyanobacteria

(Peterson and Grimm 1992; van der Grinten et al. 2005) or green algae (Roberts et al. 2003); therefore, lower temperature minima and stronger temperature seasonality may prevent diatom exclusion and promote higher biodiversity of this algal group. High temperature seasonality, low temperature minima, and high temperature maxima were also associated with the highest abundance equality (lowest σ_{res}) in the US, which suggests more equitable resource partitioning at higher temperature variability and extremes with positive consequences for diatom biodiversity. These results suggest that a decrease in the temporal variability of temperature and an increase in the temperature minima due to global warming (Xu et al. 2013), may potentially lead to lower biodiversity and abundance equality in some stream diatoms.

The response of richness to the Boltzmann temperature factor and mean temperature contradicted the MTE and the species-energy theory, respectively, predicting a positive relationship (both theories) with a specific rate of increase (MTE). Deviations from the predictions of the MTE were previously reported for both aquatic and terrestrial macroorganisms, where the richness-temperature relationship was found to be curvilinear, linear with a slope significantly different from predictions, or not significantly different from zero (Algar et al. 2007; Hawkins et al. 2007; Pinel-Alloul et al. 2013). In microbes, including lake phytoplankton and soil bacteria, the richness response to temperature was also variable, following respectively, a segmented (Segura et al. 2015) or linear pattern, less pronounced than this observed in macroorganisms (Zhou et al. 2016). A systematic analysis of multiple datasets across a broad range of terrestrial macroorganisms, testing the predictions of the MTE, found that the richness-temperature relationship was positive in datasets that included areas with colder winters but non-existent or negative in datasets from tropical, subtropical, and warm temperate regions (Hawkins et al. 2007). A review of the species-energy relationship noted that it is scale-dependent and transitions from unimodal at small scales to monotonically increasing at large scales (Evans et al. 2005). Here, the relationship of richness with the Boltzmann temperature factor and mean temperature was non-significant or weak in the US ($R^2 \leq 0.01$, Table 2) but comparatively stronger and unimodal in Finland ($R^2 = 0.07$, Table 3) even though both datasets covered regional to subcontinental scales and included cold climate streams. Ultimately, our results show that species richness is more strongly related to temperature seasonality, extremes, and total phosphorus, than to mean temperature. The SAD response to the Boltzmann temperature factor and mean temperature was weak and U-shaped in the US, but stronger and monotonic in Finland, where inequality (parameter σ_{res}) decreased with temperature. It is thus possible that mean temperature becomes a more important determinant of the SAD at higher latitudes.

The present results were also inconsistent with the species-energy theory, which so far, has received mixed support (Srivastava and Lawton 1998; Kaspari et al. 2000; Mönkkönen et al. 2006; Šímová et al. 2011). The hump-shaped behavior of density along the temperature gradient in the US dataset indicated that higher temperatures actually decrease the number of individuals, which is in stark contrast with the prediction of this theory for a positive relationship. Although this pattern was weak, it could potentially be a consequence of intensified grazing at higher temperatures due to accelerated herbivore metabolic and consumption rates (O'Connor and Bruno 2009; West and Post 2016). Finally, the richness-density relationship was non-significant, indicating that the temperature effect on richness was not through density.

Water chemistry, total phosphorus and specific conductance in particular, emerged as some of the best predictors of richness and the SAD with a comparable effect to this of the best climatic predictors. In the two countries, species richness was higher at intermediate or high nutrient values. An increase of diatom richness with nutrient supply has been previously documented and explained with the ability of more functional groups (e.g., tolerant and sensitive to nutrient limitation) to coexist at high nutrients (Passy 2008; Soininen et al. 2016). Here, we further report that in the US, the SAD was also constrained by nutrient supply. A recent study on the SAD of stream diatom communities, sampled along a land use gradient, revealed that their equitability increased (parameter σ decreased) with the transition from forest to agriculture and suggested that nutrient enrichment was responsible for this pattern (Passy 2016). In the present investigation, we found support for this hypothesis and showed that communities in the US indeed became more even at higher nutrient supply. In Finland, which contained mostly oligotrophic streams, the SAD (σ_{res}) did not respond to total phosphorus, probably because of the limited variability of this predictor (Table A1).

Remarkably, in the US, where water chemistry exhibited broad variability, the overall water chemistry model outperformed the climate model (for richness) or was comparable to it (for σ_{res}) in terms of R^2 (Table 2). Moreover, the best multiple regression model for richness and the SAD in the two countries included both climatic and water chemistry variables (model 11, Tables 2, 3). Regression trees further elucidated the interactive effects of climate and water chemistry, although in some cases this approach selected only climatic variables, because these variables correlated with the best water chemistry predictors. In the US, the longitudinal distributions of species richness and the SAD, which were the most distinct spatial patterns, appeared to be driven primarily by total phosphorus and specific conductance, respectively, with some influence of temperature seasonality (fig. 2). Notably, all three predictors, i.e. temperature seasonality, total phosphorus, and specific conductance, varied most strongly along the longitudinal gradient (figs A3B, A4A, B), which explains the weak latitudinal and altitudinal biotic patterns. We showed by variance partitioning that much of the spatial variability of richness and the SAD was due to covariance of space with both climate and water chemistry. Therefore, an understanding of the biogeography of microorganisms requires models that explicitly include water chemistry. While a combined effect of nutrients and temperature on microbial biodiversity has been reported before (Wang et al. 2016), this study is the first to systematically test different climate theories and hypotheses about diatom biodiversity and demonstrate that climate (particularly temperature seasonality) and water chemistry co-controlled the variability in both diatom richness and the SAD.

Skewness of the $\log_2$-transformed abundances was positive in both datasets in contrast with a large body of literature, documenting negative skew as the dominant pattern (Gregory and Gaston 2000; Hubbell 2001; Magurran and Henderson 2003). This discrepancy may be due to both sampling intensity and environmental influences. Thus, when too few individuals are sampled, the left side of the SAD remains veiled and the right side exhibits a positive skew (McGill 2003). Although in both datasets a comparatively large number of individuals was sampled following standard protocols, the proportion of species revealed by the sample had a median of 44-75%, indicating that some species remained veiled. However, skewness had detectable correlations with environmental factors (R^2 = 0.07-0.22, Table A3), pointing to some ecological constraints on the SAD symmetry as well. Skewness responded most strongly to temperature seasonality and extremes in the US, while temperature mean and extremes were the underlying factors in Finland. Determining the drivers controlling the distribution of common species, i.e. the SAD symmetry, has important practical implications, given that these species are primarily responsible for delivering ecosystem services (Winfree et al. 2015). Here we show that these drivers are mainly climatic; therefore, climate change may have a strong effect on diatom community functioning in stream ecosystems.

It is noteworthy that the SAD (parameter σ_{res}) exhibited distinct environmental and spatial responses that were independent of richness, consistent with our third objective. There were also some differences in the environmental and spatial predictors that entered the regression models of species richness and σ_{res} (Tables 2, 3, fig. 4) as well as in the shape of the species richness and σ_{res} responses to common predictors (figs A8, A9). The form of the SAD is strongly dependent on richness and total abundance (Locey and White 2013, Passy 2016) and the usefulness of the SAD for ecological research has been questioned (Yen et al. 2013), in part because of its dependence on richness. The results presented here demonstrate that the abiotic environment has distinct impacts on richness and the SAD, and that the shape of the SAD can be predicted to some extent by climatic and water chemistry variables, emphasizing the utility of the SAD in unravelling ecological mechanisms.

Acknowledgements

We thank D. Bolnick, M. Duffy, and two anonymous reviewers for insightful comments, which greatly improved the clarity of our work. SIP gratefully acknowledges support from the National Science Foundation (grant NSF DEB-1745348).

Appendix

Table A1. Basic statistics of spatial, climatic, and water chemistry variables included in the regression analyses (Tables 2, 3). T_{mean} = mean annual temperature, T_{SD} = temperature seasonality (standard deviation), T_{min} = minimum temperature of the coldest month, and T_{max} = maximum temperature of the warmest month.

Variable	**Minimum**	**Maximum**	**Median**	**Mean**	**SD**
US					
Latitude (degrees)	25.40	61.77	41.64	41.45	4.64
Longitude (degrees)	−154.17	−70.74	−92.38	−95.81	17.15
Altitude (m)	4.00	2452.00	245.00	424.47	502.46
T_{mean} (°C)	0.50	23.70	9.70	10.08	3.69
T_{SD}	3.19	12.58	8.99	8.73	1.74
T_{min} (°C)	−22.20	13.40	−8.40	−8.09	5.87
T_{max} (°C)	15.20	36.30	29.15	29.40	2.85
Total phosphorus ($mg \cdot L^{-1}$)	0.002	3.40	0.06	0.17	0.34
Nitrite + nitrate ($mg \cdot L^{-1}$)	0.005	13.95	0.43	1.27	2.14
Ammonia ($mg \cdot L^{-1}$)	0.002	5.80	0.03	0.06	0.27
pH	2.95	9.60	7.90	7.79	0.58
Specific conductance ($\mu S \cdot cm^{-1}$)	10.00	23594.65	421.00	525.91	1075.29
Finland					
Latitude (degrees)	60.27	70.06	66.26	65.91	2.81
Longitude (degrees)	23.28	30.44	27.02	27.04	2.09
Altitude (m)	18.00	320.00	180.00	177.49	81.91
T_{mean} (°C)	−2.70	5.00	−0.50	0.28	2.03
T_{SD}	6.93	10.19	9.52	9.29	0.80
T_{min} (°C)	−21.30	−9.10	−17.15	−16.60	3.29
T_{max} (°C)	14.80	21.00	18.75	18.78	1.66

Total phosphorus (mg·L^{-1})	0.00	0.19	0.02	0.03	0.04
pH	4.50	8.20	7.00	6.80	0.81
Specific conductance (µS·cm^{-1})	9.40	277.00	44.75	65.37	59.24

Table A2. Pearson correlation matrix of climate and water chemistry variables in the US and Finland. Boldfaced values are not significant ($p \geq 0.05$). T_{mean} = mean annual temperature, T_{SD} = temperature seasonality (standard deviation), T_{min} = minimum temperature of the coldest month, T_{max} = maximum temperature of the warmest month, TP = total phosphorus, and Cond = specific conductance.

Variable	T_{mean}	T_{SD}	T_{min}	T_{max}	ln(TP)	pH	ln(Cond)	ln(NO_x)
US, *n* = 526								
T_{SD}	−0.56	1.00						
T_{min}	0.85	−0.89	1.00					
T_{max}	0.70	**−0.05**	0.37	1.00				
ln(TP)	**−0.01**	0.20	−0.10	**0.08**	1.00			
pH	−0.23	0.25	−0.26	0.10	0.11	1.00		
ln(Cond)	**0.01**	0.42	−0.23	0.28	0.38	0.44	1.00	
ln(NO_x)	0.09	0.16	**0.02**	0.11	0.41	**0.08**	0.38	1.00
ln(NH_4)	0.24	**0.00**	0.15	0.16	0.39	−0.15	0.21	0.31
Finland, *n* = 100								
T_{SD}	−0.54	1.00						
T_{min}	0.95	−0.76	1.00					
T_{max}	0.71	0.20	0.48	1.00				
ln(TP)	0.68	**−0.04**	0.52	0.74	1.00			
pH	**−0.18**	−0.24	**−0.04**	−0.41	**−0.19**	1.00		
ln(Cond)	0.48	−0.38	0.50	0.26	0.39	0.58	1.00	

Table A3. Regression models of skewness in the US and Finland. Std. = standardized. An outlier sample was removed from the Finnish dataset. T_{mean} = mean annual temperature, T_{SD} = temperature seasonality (standard deviation), T_{min} = minimum temperature of the coldest month, and T_{max} = maximum temperature of the warmest month.

Effect	**Std. coefficient**	**Standard error**	***t-value***	***p*-value**
US, R^2 = 0.07, *n* = 526				
T_{mean}	0.21	0.01	2.75	0.006
T_{SD}^2	−0.23	0.00	−4.56	0.000006
T_{max}	−0.24	0.01	−3.34	0.0009
NO_x	−0.11	0.01	−2.54	0.01
pH	−0.11	0.03	−2.38	0.02
Finland, R^2 = 0.22, *n* = 99				
T_{mean}^2	2.87	0.08	2.34	0.02
T_{SD}^2	0.97	0.17	2.15	0.03
T_{min}^2	−3.01	0.03	−2.47	0.02
T_{max}	−0.60	0.04	−3.45	0.0009
T_{max}^2	−0.99	0.04	−2.30	0.02

Figure A1. Maps of the study streams in the US and Finland.

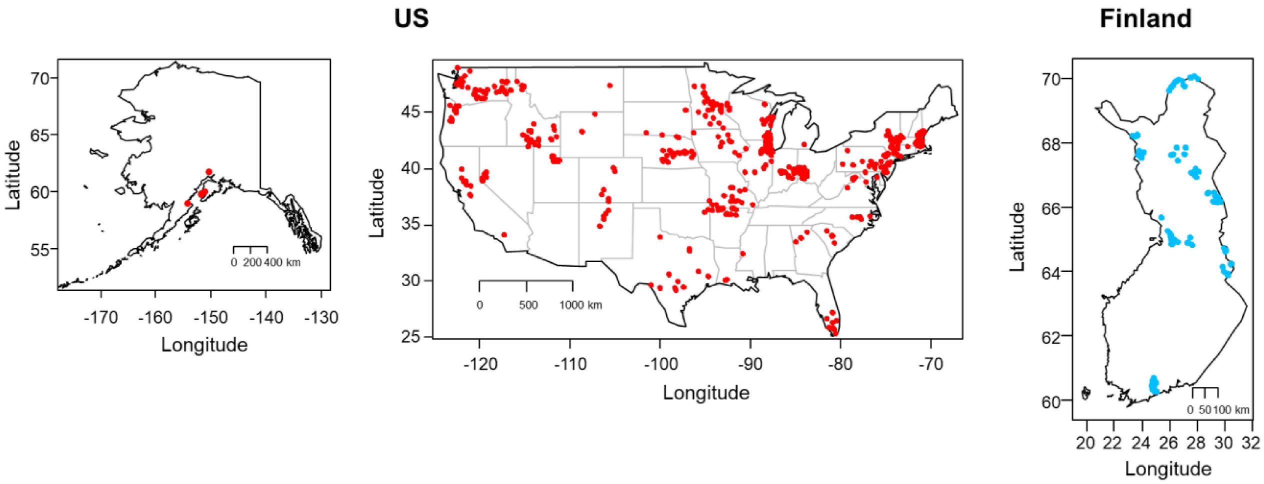


Figure A2. Scatter plots of parameter *σ* against species richness in the US, *n* = 526 (A) and Finland, *n* = 100 (B). The regression fits, models, and statistics are shown in the panels.

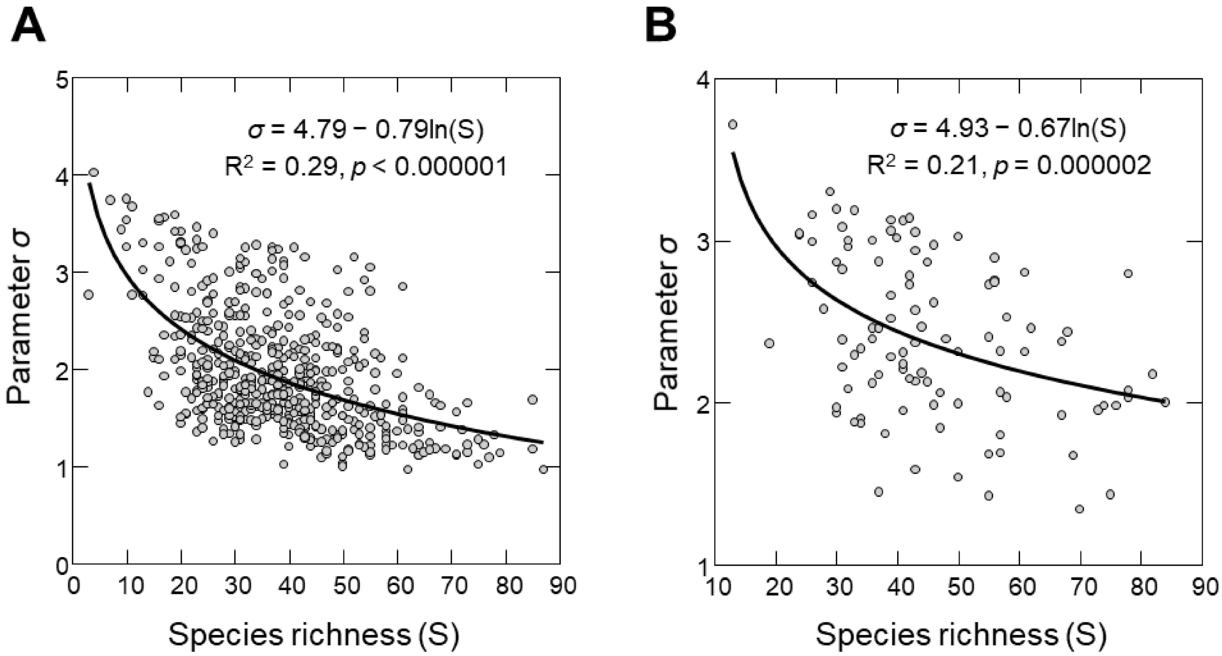


Figure A3. Maps showing the spatial variability of mean temperature, T_{mean} (A), temperature seasonality, T_{SD} (B), minimum temperature, T_{min} (C), and maximum temperature, T_{max} (D) in the US. The insets show sites in Alaska. Regressions against the significant ($p < 0.05$) linear, quadratic, and cubic terms of latitude (Lat), longitude (Long), and altitude (Alt) explained 95% of the variance in T_{mean} ($-1.05\text{Lat} + 0.10\text{Lat}^2 + 0.13\text{Lat}^3 - 0.68\text{Long} + 0.23\text{Long}^2 + 0.45\text{Long}^3 - 0.35\text{Alt} - 0.50\text{Alt}^2 - 0.38\text{Alt}^3$), 84% of the variance in T_{SD} ($0.53\text{Lat} + 0.40\text{Lat}^3 + 0.75\text{Long} - 0.93\text{Long}^2 - 0.41\text{Long}^3 + 0.27\text{Alt} + 0.10\text{Alt}^2$), 89% of the variance in T_{min} ($-0.82\text{Lat} - 0.19\text{Lat}^3 - 0.84\text{Long} + 0.70\text{Long}^2 + 0.48\text{Long}^3 - 0.37\text{Alt} - 0.42\text{Alt}^2 - 0.28\text{Alt}^3$), and 72% of the variance in T_{max} ($-1.08\text{Lat} + 0.55\text{Lat}^3 - 0.83\text{Long} + 0.82\text{Long}^3 - 0.38\text{Alt}^2 - 0.36\text{Alt}^3$).

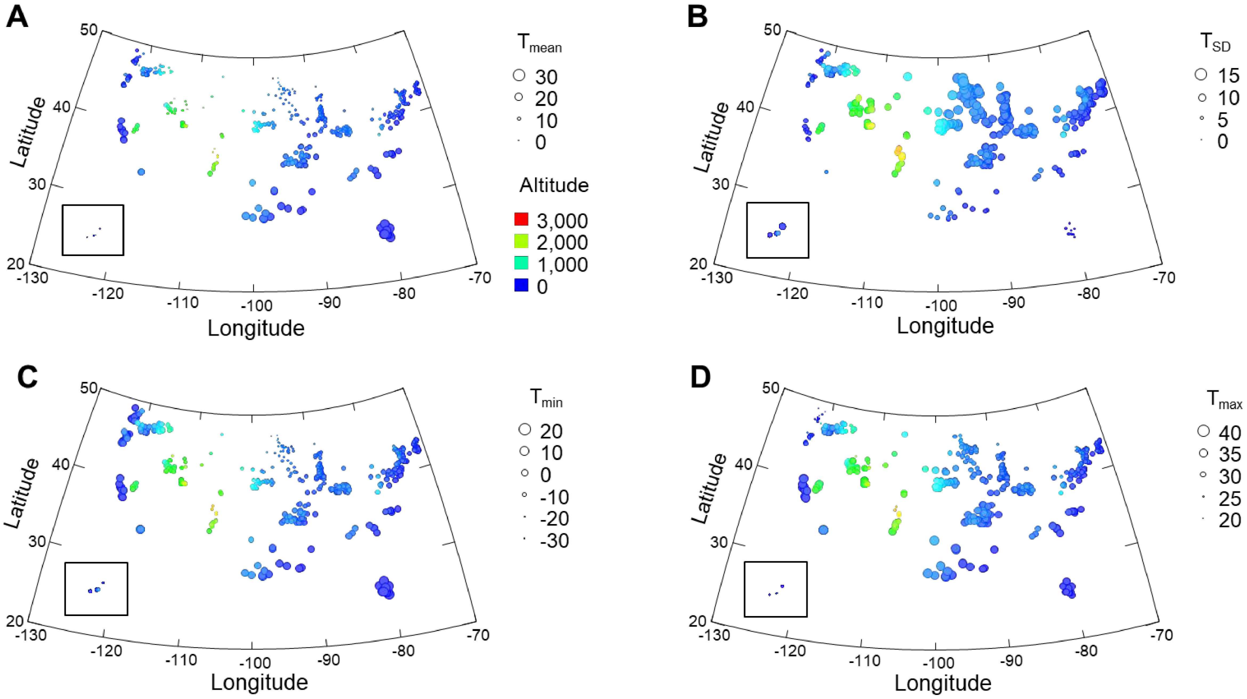


Figure A4. Maps showing the spatial variability of ln-transformed total phosphorus, ln(TP) (A), ln-transformed specific conductance (B), species richness (C), and σ_{res} (D) in the US. The insets show sites in Alaska. Regressions against the significant ($p < 0.05$) linear, quadratic, and cubic terms of latitude (Lat), longitude (Long), and altitude (Alt) explained 17% of the variance in ln(TP) ($-0.34\text{Lat}^2 - 0.71\text{Long}^2 - 0.84\text{Long}^3 - 0.29\text{Alt}$), 30% of the variance in ln(Conductance) ($0.13\text{Lat} - 0.34\text{Lat}^2 - 0.36\text{Lat}^3 + 0.35\text{Long} - 0.81\text{Long}^2 - 1.07\text{Long}^3$), 14% of the variance in species richness, and 17% of the variance in σ_{res}. The models for species richness and σ_{res} are given in Table 2 (model 4).

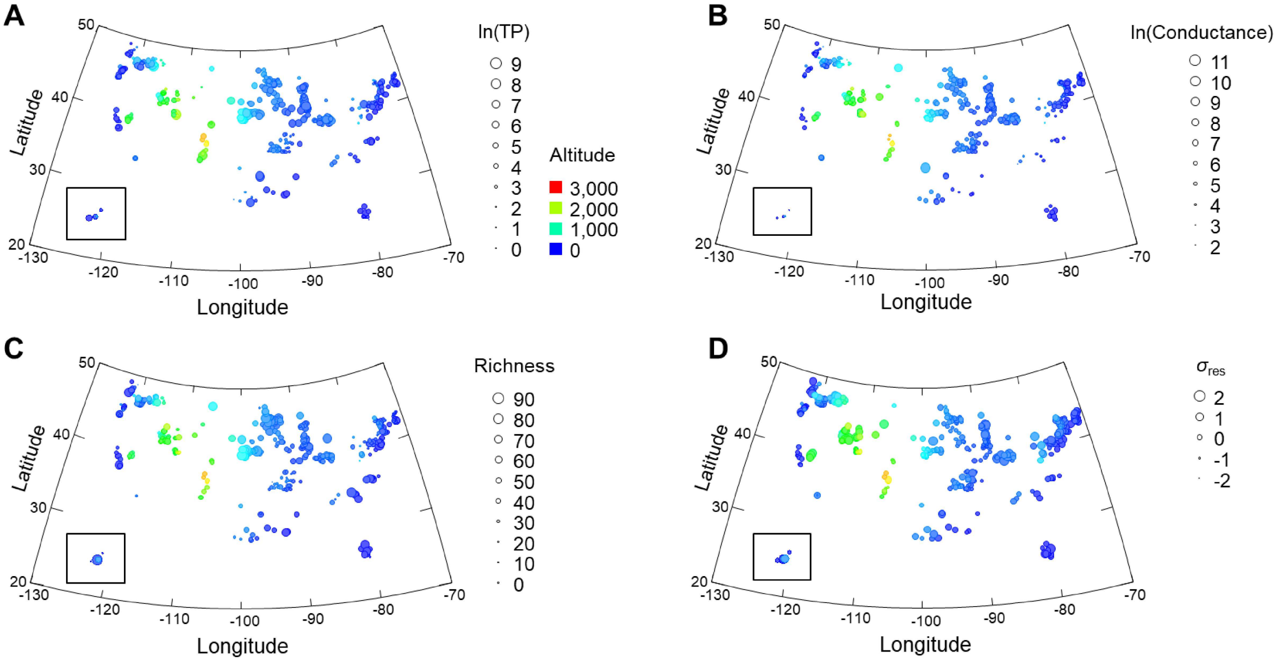


Figure A5. Maps showing the spatial variability of mean temperature, T_{mean} (A), temperature seasonality, T_{SD} (B), minimum, T_{min} (C), and maximum temperature, T_{max} (D) in Finland. Regressions against the significant ($p < 0.05$) linear, quadratic, and cubic terms of latitude (Lat), longitude (Long), and altitude (Alt) explained 98% of the variance in T_{mean} ($-0.98\text{Lat} + 0.25\ \text{Lat}^2 + 0.26\ \text{Lat}^3 + 0.38\text{Long} - 0.06\text{Long}^2 - 0.34\text{Long}^3 - 0.41\text{Alt} + 0.07\text{Alt}^2 + 0.13\text{Alt}^3$), 94% of the variance in T_{SD} ($-0.92\text{Lat}^2 - 0.25\text{Lat}^3 - 1.18\ \text{Long} + 0.13\text{Long}^2 + 0.80\ \text{Long}^3 + 0.71\text{Alt} - 0.29\text{Alt}^2 - 0.45\text{Alt}^3$), 97% of the variance in T_{min} ($-0.74\text{Lat} + 0.54\text{Lat}^2 + 0.26\text{Lat}^3 + 0.85\text{Long} - 0.08\text{Long}^2 - 0.60\text{Long}^3 - 0.56\text{Alt} + 0.18\text{Alt}^2 + 0.28\text{Alt}^3$), and 99% of the variance in T_{max} ($-0.95\ \text{Lat} - 0.58\text{Lat}^2 - 0.19\ \text{Lat}^3 - 0.38\text{Long} - 0.03\text{Long}^2 + 0.21\text{Long}^3 - 0.14\text{Alt}^2 - 0.11\text{Alt}^3$).

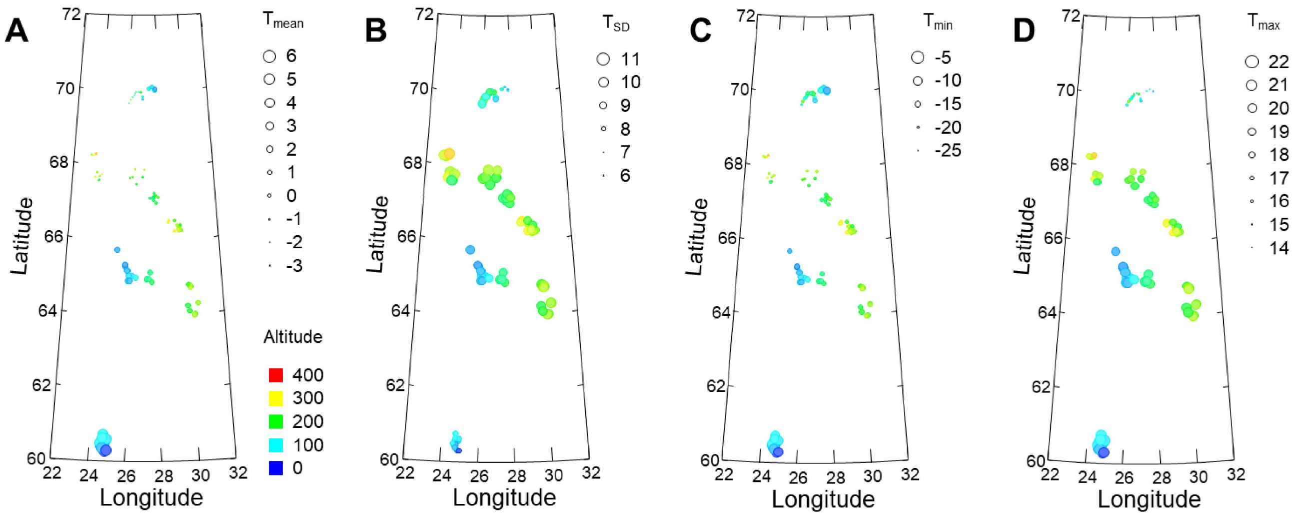


Figure A6. Maps showing the spatial variability of ln-transformed total phosphorus, ln(TP) (A), ln-transformed specific conductance, ln(Cond) (B), species richness (C), and σ_{res} (D) in Finland. Regressions against the significant ($p < 0.05$) linear, quadratic, and cubic terms of latitude (Lat), longitude (Long), and altitude (Alt) explained 67% of the variance in ln(TP) ($-0.87\text{Lat} - 0.16\text{Lat}^2 - 0.34\text{Long}^3$), 71% of the variance in ln(Cond) ($2.13\text{Lat} - 0.35\text{Lat}^2 - 2.90\text{Lat}^3 + 0.48\text{Long} - 0.48\text{Alt} + 0.26\text{Alt}^3$), 24% of the variance in species richness, and 15% of the variance in σ_{res}. The models for species richness and σ_{res} are given in Table 3 (model 4).

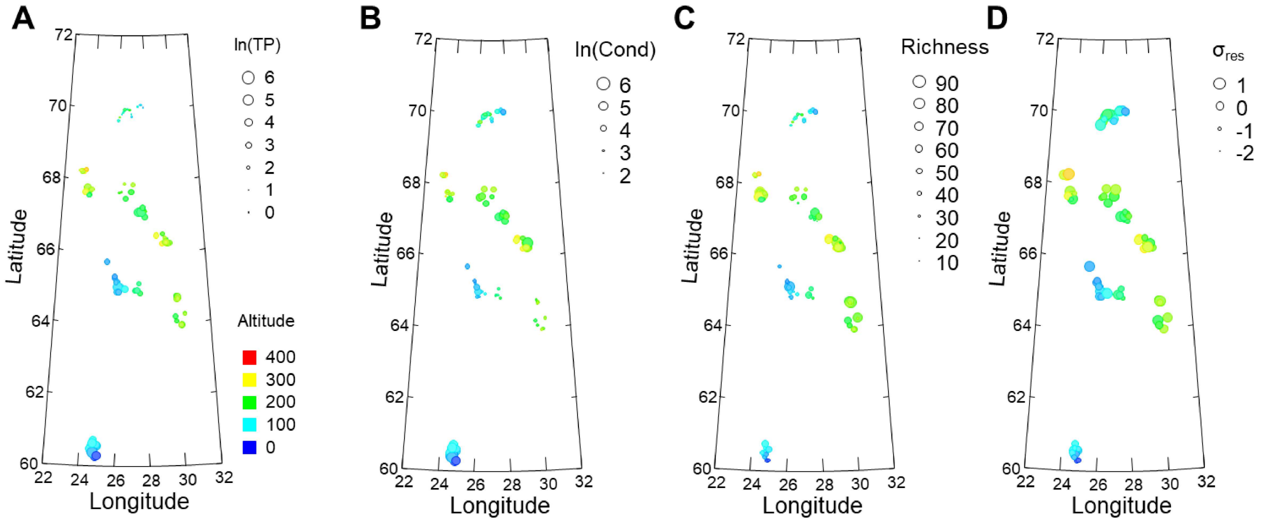


Figure A7. Frequency distributions of species richness and parameters σ and σ_{res} in the US (A-C) and Finland (D-F).

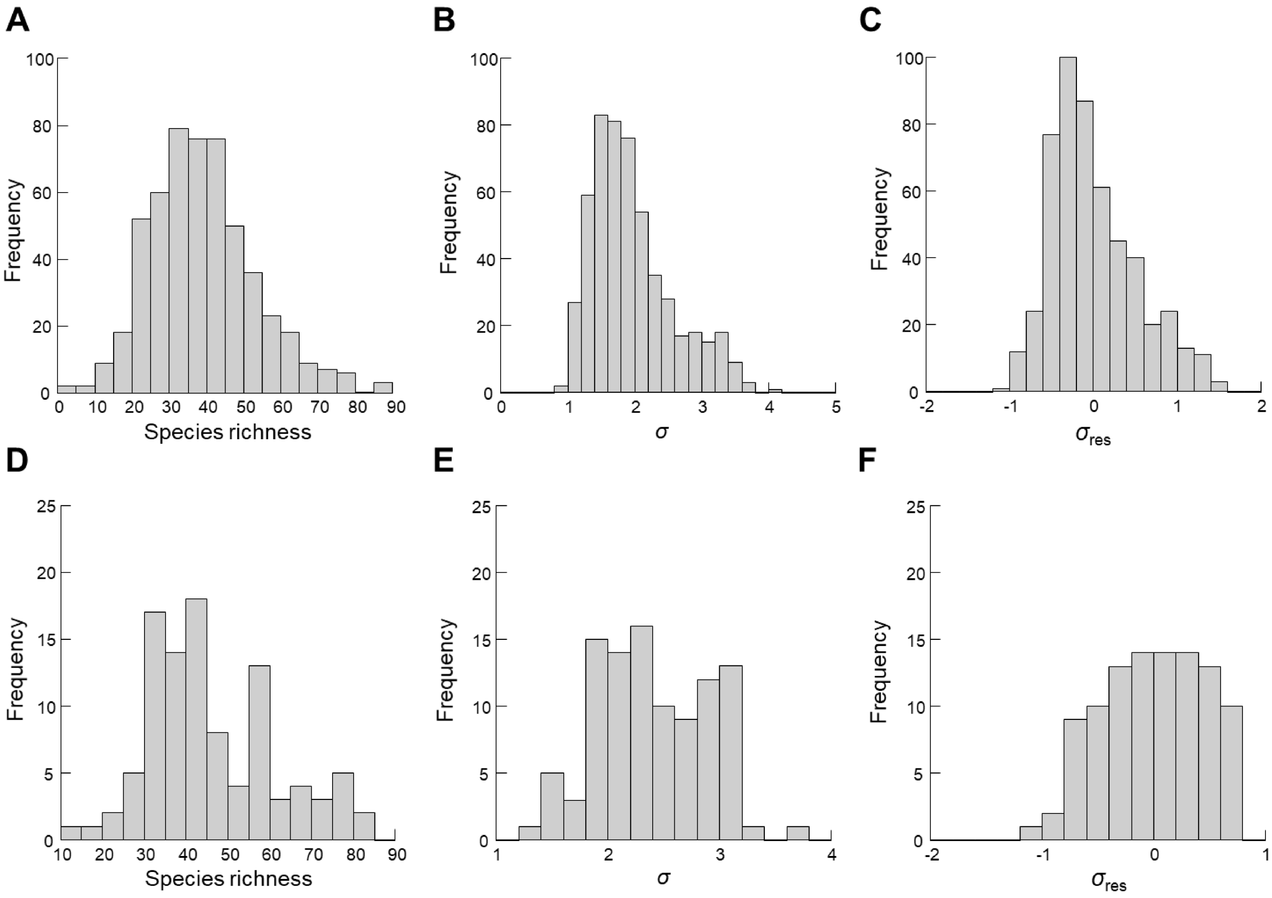


Figure A8. Relationships of species richness (A-C) and σ_{res} (D-F) with their respective strongest water chemistry and climatic predictors in the US. The regression fits and statistics are shown in the panels. n = 526.

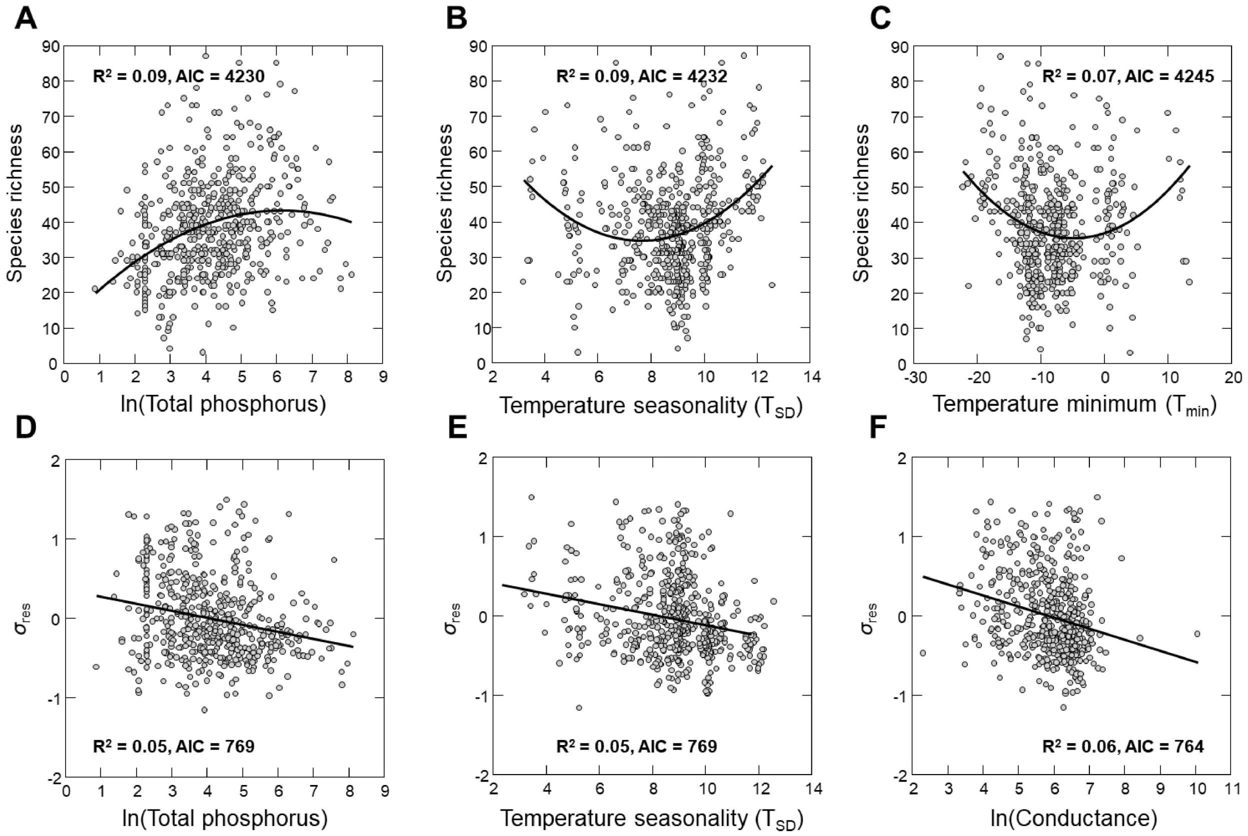
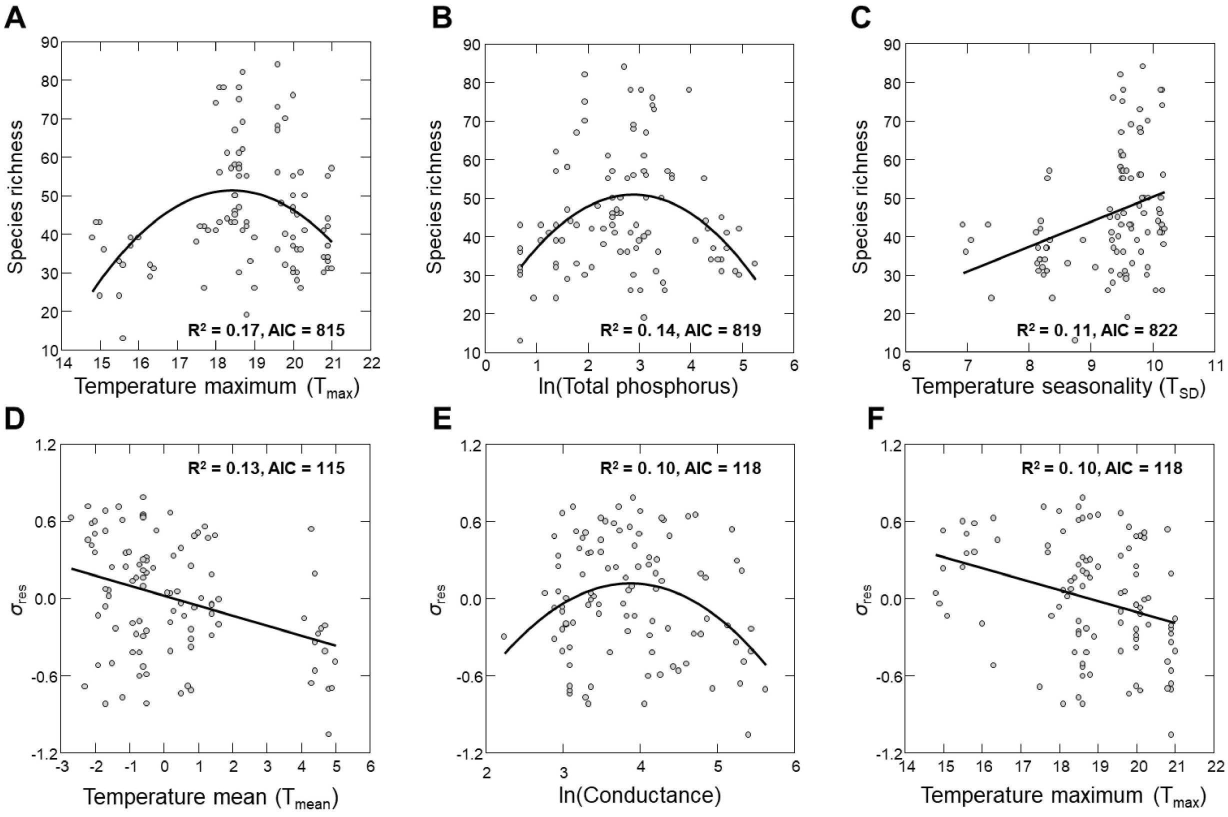


Figure A9. Relationships of species richness (A-C) and σ_{res} (DF) with their respective strongest climatic and water chemistry predictors in Finland. The regression fits and statistics are shown in the panels. *n* = 100.



Literature Cited

Alexander, T. J., P. Vonlanthen, and O. Seehausen. 2017. Does eutrophication-driven evolution change aquatic ecosystems? Philosophical Transactions of the Royal Society B-Biological Sciences 372:20160041.

Algar, A. C., J. T. Kerr, and D. J. Currie. 2007. A test of Metabolic Theory as the mechanism underlying broad-scale species-richness gradients. Global Ecology and Biogeography 16:170-178.

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

Allen, A. P., J. F. Gillooly, V. M. Savage, and J. H. Brown. 2006. Kinetic effects of temperature on rates of genetic divergence and speciation. Proceedings of the National Academy of Sciences (USA) 103:9130-9135.

Arellano, G., M. N. Umana, M. J. Macia, M. I. Loza, A. Fuentes, V. Cala, and P. M. Jorgensen. 2017. The role of niche overlap, environmental heterogeneity, landscape roughness and productivity in shaping species abundance distributions along the Amazon-Andes gradient. Global Ecology and Biogeography 26:191-202.

Baldridge, E., D. J. Harris, X. Xiao, and E. P. White. 2016. An extensive comparison of species-abundance distribution models. Peerj 4: e2823 https://doi.org/10.7717/peerj.2823.

Brown, J. H. 2014. Why are there so many species in the tropics? Journal of Biogeography 41:8-22.

Brown, J. H., J. F. Gillooly, A. P. Allen, V. M. Savage, and G. B. West. 2004. Toward a metabolic theory of ecology. Ecology 85:1771-1789.

Carrara, R. and D. P. Vazquez. 2010. The species-energy theory: a role for energy variability. Ecography 33:942-948.

Chaudhary, C., H. Saeedi, and M. J. Castello. 2016. Bimodality of latitudinal gradients in marine species richness. Trends in Ecology & Evolution 31:670-676.

Currie, D. J., G. G. Mittelbach, H. V. Cornell, R. Field, J. F. Guegan, 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.

Evans, K. L., P. H. Warren, and K. J. Gaston. 2005. Species-energy relationships at the macroecological scale: a review of the mechanisms. Biological Reviews 80:1-25.

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

Gregory, R. D. and K. J. Gaston. 2000. Explanations of commonness and rarity in British breeding birds: separating resource use and resource availability. Oikos 88:515-526.

Grøtan, V. and S. Engen. 2008. poilog: Poisson lognormal and bivariate Poisson lognormal distribution. R package version 0.4. http://cran.r-project.org/web/packages/poilog.

Hawkins, B. A., F. S. Albuquerque, M. B. Araujo, J. Beck, L. M. Bini, F. J. Cabrero-Sanudo, I. Castro-Parga, J. A. F. Diniz, D. Ferrer-Castan, R. Field, J. F. Gomez, J. Hortal, J. T. Kerr, I. J. Kitching, J. L. Leon-Cortes, J. M. Lobo, D. Montoya, J. C. Moreno, M. A. Olalla-Tarraga, J. G. Pausas, H. Qian, C. Rahbek, M. A. Rodriguez, N. J. Sanders, and P. Williams. 2007. A global evaluation of metabolic theory as an explanation for terrestrial species richness gradients. Ecology 88:1877-1888.

Hijmans, R. J., S. E. Cameron, J. L. Parra, P. G. Jones, and A. Jarvis. 2005. Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology 25:1965-1978.

Hubbell, S. P. 2001. The unified neutral theory of biodiversity and biogeography. Princeton University Press, Princeton, New Jersey.

Kaspari, M., S. O'Donnell, and J. R. Kercher. 2000. Energy, density, and constraints to species richness: Ant assemblages along a productivity gradient. American Naturalist 155:280-293.

Keeley, J. E. 2003. Relating species abundance distributions to species-area curves in two Mediterranean-type shrublands. Diversity and Distributions 9:253-259.

Locey, K. J. and E. P. White. 2013. How species richness and total abundance constrain the distribution of abundance. Ecology Letters 16:1177-1185.

Magurran, A. E. and P. A. Henderson. 2003. Explaining the excess of rare species in natural species abundance distributions. Nature 422:714-716.

Matthews, T. J., P. A. V. Borges, E. B. de Azevedo, and R. J. Whittaker. 2017. A biogeographical perspective on species abundance distributions: recent advances and opportunities for future research. Journal of Biogeography 44:1705-1710.

May, R. M. 1975. Patterns of species abundance and diversity. Pages 81-120 *in* M. L. Cody and J. M. Diamond, eds. Ecology and evolution of communities. Harvard University Press, Cambridge, MA.

McGill, B. J. 2003. Does Mother Nature really prefer rare species or are log-left-skewed SADs a sampling artefact? Ecology Letters 6:766-773.

McGill, B. J., R. S. Etienne, J. S. Gray, D. Alonso, M. J. Anderson, H. K. Benecha, M. Dornelas, B. J. Enquist, J. L. Green, F. L. He, A. H. Hurlbert, A. E. Magurran, P. A. Marquet, B. A. Maurer, A. Ostling, C. U. Soykan, K. I. Ugland, and E. P. White. 2007. Species abundance distributions: moving beyond single prediction theories to integration within an ecological framework. Ecology Letters 10:995-1015.

Mittelbach, G. G., D. W. Schemske, H. V. Cornell, A. P. Allen, J. M. Brown, M. B. Bush, S. P. Harrison, A. H. Hurlbert, N. Knowlton, H. A. Lessios, C. M. McCain, A. R. McCune, L. A. McDade, M. A. McPeek, T. J. Near, T. D. Price, R. E. Ricklefs, K. Roy, D. F. Sax, D. Schluter, J. M. Sobel, and M. Turelli. 2007. Evolution and the latitudinal diversity gradient: speciation, extinction and biogeography. Ecology Letters 10:315-331.

Mönkkönen, M., J. T. Forsman, and F. Bokma. 2006. Energy availability, abundance, energy-use and species richness in forest bird communities: a test of the species-energy theory. Global Ecology and Biogeography 15:290-302.

O'Connor, M. I. and J. F. Bruno. 2009. Predator richness has no effect in a diverse marine food web. Journal of Animal Ecology 78:732-740.

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.

Passy, S. I. 2016. Abundance inequality in freshwater communities has an ecological origin. American Naturalist 187:502-516.

Passy, S. I., C. A. Larson, A. Jamoneau, W. Budnick, J. Heino, T. Leboucher, J. Tison-Rosebery, and J. Soininen 2018. Data from: The biogeographical patterns of species richness and abundance distribution in stream diatoms are driven by climate and water chemistry. American Naturalist, Dryad Digital Repository, https://datadryad.org/resource/doi:10.5061/dryad.v1v7856.

Peterson, C. G. and N. B. Grimm. 1992. Temporal variation in enrichment effects during periphyton succession in a nitrogen-limited desert stream ecosystem. Journal of the North American Benthological Society 11:20-36.

Pianka, E. R. 1966. Latitudinal gradients in species diversity: a review of concepts. American Naturalist 100:33-46.

Pinel-Alloul, B., A. Andre, P. Legendre, J. A. Cardille, K. Patalas, and A. Salki. 2013. Large-scale geographic patterns of diversity and community structure of pelagic crustacean zooplankton in Canadian lakes. Global Ecology and Biogeography 22:784-795.

Preston, F. W. 1962. Canonical distribution of commonness and rarity: Part I. Ecology 43:185-215.

Qiao, X. J., F. Jabot, Z. Y. Tang, M. X. Jiang, and J. Y. Fang. 2015. A latitudinal gradient in tree community assembly processes evidenced in Chinese forests. Global Ecology and Biogeography 24:314-323.

Rahbek, C. 2005. The role of spatial scale and the perception of large-scale species-richness patterns. Ecology Letters 8:224-239.

Roberts, E., J. Kroker, S. Korner, and A. Nicklisch. 2003. The role of periphyton during the re-colonization of a shallow lake with submerged macrophytes. Hydrobiologia 506:525-530.

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

Sæther, B. E., S. Engen, and V. Grøtan. 2013. Species diversity and community similarity in fluctuating environments: parametric approaches using species abundance distributions. Journal of Animal Ecology 82:721-738.

Segura, A. M., D. Calliari, C. Kruk, H. Fort, I. Izaguirre, J. F. Saad, and M. Arim. 2015. Metabolic dependence of phytoplankton species richness. Global Ecology and Biogeography 24:472-482.

Šímová, I., D. Storch, P. Keil, B. Boyle, O. L. Phillips, and B. J. Enquist. 2011. Global species-energy relationship in forest plots: role of abundance, temperature and species climatic tolerances. Global Ecology and Biogeography 20:842-856.

Smith, V. H. and D. W. Schindler. 2009. Eutrophication science: where do we go from here? Trends in Ecology & Evolution 24:201-207.

Soininen, J. 2007. Environmental and spatial control of freshwater diatoms - A review. Diatom Research 22:473-490.

Soininen, J., A. Jamoneau, J. Rosebery, and S. I. Passy. 2016. Global patterns and drivers of species and trait composition in diatoms. Global Ecology and Biogeography 25:940-950.

Sommer, U. 1985. Comparison between steady state and non-steady state competition: experiments with natural phytoplankton. Limnology and Oceanography 30:335-346.

Srivastava, D. S. and J. H. Lawton. 1998. Why more productive sites have more species: An experimental test of theory using tree-hole communities. American Naturalist 152:510-529.

Terborgh, J. 1973. On the notion of favorableness in plant ecology. American Naturalist 107:481-501.

Tilman, D. and S. Pacala. 1993. The maintenance of species richness in plant communities. Pages 13-25 *in* R. E. Ricklefs and D. Schluter, eds. Species diversity in ecological communities. University of Chicago Press, Chicago.

Ulrich, W., B. Kusumoto, T. Shiono, and Y. Kubota. 2016a. Climatic and geographic correlates of global forest tree species-abundance distributions and community evenness. Journal of Vegetation Science 27:295-305.

Ulrich, W., M. Ollik, and K. I. Ugland. 2010. A meta-analysis of species-abundance distributions. Oikos 119:1149-1155.

Ulrich, W., S. Soliveres, A. D. Thomas, A. J. Dougill, and F. T. Maestre. 2016b. Environmental correlates of species rank abundance distributions in global drylands. Perspectives in Plant Ecology Evolution and Systematics 20:56-64.

van der Grinten, E., A. P. H. M. Janssen, K. de Mutsert, C. Barranguet, and W. Admiraal. 2005. Temperature- and light-dependent performance of the cyanobacterium *Leptolyngbya foveolarum* and the diatom *Nitzschia perminuta* in mixed biofilms. Hydrobiologia 548:267-278.

Wang, J. J., F. Y. Pan, J. Soininen, J. Heino, and J. Shen. 2016. Nutrient enrichment modifies temperature-biodiversity relationships in large-scale field experiments. Nature Communications 7:13960, doi:10.1038/ncomms13960.

West, D. C. and D. M. Post. 2016. Impacts of warming revealed by linking resource growth rates with consumer functional responses. Journal of Animal Ecology 85:671-680.

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.

Winfree, R., J. W. Fox, N. M. Williams, J. R. Reilly, and D. P. Cariveau. 2015. Abundance of common species, not species richness, drives delivery of a real-world ecosystem service. Ecology Letters 18:626-635.

Wright, D. H. 1983. Species-energy theory: an extension of species-area theory. Oikos 41:496-506.

Xu, L., R. B. Myneni, F. S. Chapin, T. V. Callaghan, J. E. Pinzon, C. J. Tucker, Z. Zhu, J. Bi, P. Ciais, H. Tommervik, E. S. Euskirchen, B. C. Forbes, S. L. Piao, B. T. Anderson, S. Ganguly, R. R. Nemani, S. J. Goetz, P. S. A. Beck, A. G. Bunn, C. Cao, and J. C. Stroeve. 2013. Temperature and vegetation seasonality diminishment over northern lands. Nature Climate Change 3:581-586.

Yen, J. D. L., J. R. Thomson, and R. Mac Nally. 2013. Is there an ecological basis for species abundance distributions? Oecologia 171:517-525.

Zhou, J. Z., Y. Deng, L. N. Shen, C. Q. Wen, Q. Y. Yan, D. L. Ning, Y. J. Qin, K. Xue, L. Y. Wu, Z. L. He, J. W. Voordeckers, J. D. Van Nostrand, V. Buzzard, S. T. Michaletz, B. J. Enquist, M. D. Weiser, M. Kaspari, R. Waide, Y. F. Yang, and J. H. Brown. 2016. Temperature mediates continental-scale diversity of microbes in forest soils. Nature Communications 7:12083, doi: 10.1038/ncomms12083.

Table 1. Tested effects, theories, and hypotheses with corresponding regression models (Eqn). Lat = latitude, Long = longitude, Alt = altitude, T_{mean} = mean annual temperature, T_{SD} = temperature seasonality (standard deviation), T_{min} = minimum temperature of the coldest month, T_{max} = maximum temperature of the warmest month, TP = total phosphorus, and Cond = specific conductance.

Eqn	Effect/theory/hypothesis	Regression model
1	Latitudinal effect	$Lat + Lat^2 + Lat^3$
2	Longitudinal effect	$Long + Long^2 + Long^3$
3	Altitudinal effect	$Alt + Alt^2 + Alt^3$
4	Spatial effect	$Lat + Lat^2 + Lat^3 + Long + Long^2 + Long^3 + Alt + Alt^2 + Alt^3$
5	Species-energy theory	$T_{mean} + T_{mean}^2$
6	Species-energy theory	$\ln(N) + \ln(N)^2$
7	Energy variability hypothesis	$T_{SD} + T_{SD}^2$
8	Climatic tolerance hypothesis	$T_{min} + T_{min}^2 + T_{max} + T_{max}^2$
9	Climate effect	$T_{mean} + T_{mean}^2 + T_{SD} + T_{SD}^2 + T_{min} + T_{min}^2 + T_{max} + T_{max}^2$
10	Chemistry effect	$TP + TP^2 + NO_x + NO_x^2 + NH_4 + NH_4^2 + pH + pH^2 + Cond + Cond^2$
11	Climate + chemistry effect	$T_{mean} + T_{mean}^2 + T_{SD} + T_{SD}^2 + T_{min} + T_{min}^2 + T_{max} + T_{max}^2 + TP + TP^2$ $+ NO_x + NO_x^2 + NH_4 + NH_4^2 + pH + pH^2 + Cond + Cond^2$
12	Metabolic theory	$1/kT_{mean}$
13	Metabolic theory	$1/kT_{mean} + (1/kT_{mean})^2$

Table 2. US diatoms. Regression models (Eqn) testing the responses of species richness (models 1-11), ln-transformed species richness, ln(S) (models 12, 13), and residual parameter σ (σ_{res}, models 1-13) to spatial and environmental variables. Parameter σ_{res} was obtained from a regression of parameter σ against ln(S) (fig. A2A). The parameters in each model are standardized regression coefficients with $0.000001 < p < 0.05$. $n = 526$. AIC = Akaike Information Criterion. The AIC of the best environmental model is boldfaced. Abbreviations as in Table 1. NS = non-significant.

Eqn	**Regression model for richness**	**R^2**	**AIC**	**Regression model for σ_{res}**	**R^2**	**AIC**
1	$0.20\text{Lat} + 0.15\text{Lat}^2 - 0.18\text{Lat}^3$	0.03	4269	$0.18\text{Lat} + 0.22\text{Lat}^2 - 0.15\text{Lat}^3$	0.05	776
2	$-0.34\text{Long}^2 - 0.35\text{Long}^3$	0.05	4256	$0.57\text{Long}^2 + 0.37\text{Long}^3$	0.12	729
3	-0.15Alt^3	0.02	4268	0.13Alt^2	0.02	787
4	$0.34\text{Lat} + 0.17\text{Lat}^2 - 0.69\text{Long}^2 - 0.31\text{Long}^3 - 0.21\text{Alt} + 0.16\text{Alt}^2$	0.14	4211	$0.30\text{Lat}^2 + 0.68\text{Long}^2 + 0.64\text{Long}^3 + 0.16\text{Alt}$	0.17	707
5	$0.12\text{T}_{mean}{}^2$	0.01	4273	$0.11\text{T}_{mean} + 0.20\text{T}_{mean}{}^2$	0.03	785
6	$\ln(N) + \ln(N)^2$	NS	–	$-0.11\ln(N) + 0.13\ln(N)^2$	0.04	778
7	$0.24\text{T}_{SD} + 0.31\text{T}_{SD}{}^2$	0.09	4232	-0.22T_{SD}	0.05	770
8	$-0.17\text{T}_{min} + 0.28\text{T}_{min}{}^2$	0.07	4245	$0.21\text{T}_{min} - 0.22\text{T}_{max}$	0.06	767
9	$-0.32\text{T}_{mean}{}^2 + 0.44\text{T}_{SD} + 0.24\text{T}_{min} + 0.52\text{T}_{min}{}^2 + 0.19\text{T}_{max}{}^2$	0.11	4225	$-0.28\text{T}_{SD} - 0.46\text{T}_{SD}{}^2 + 0.39\text{T}_{min}{}^2 - 0.26\text{T}_{max}$	0.11	740
10	$0.35\text{TP} - 0.13\text{TP}^2 - 0.20\text{NO}_x + 0.11\text{NH}_4$	0.13	4213	$-0.12\text{TP} - 0.12\text{NO}_x - 0.15\text{Cond}$	0.09	751
11	$-0.19\text{T}_{mean}{}^2 + 0.21\text{T}_{SD} + 0.40\text{T}_{min}{}^2 + 0.13\text{T}_{max}{}^2 + 0.28\text{TP} - 0.10\text{TP}^2 - 0.15\text{NO}_x + 0.11\text{NH}_4$	0.19	**4182**	$0.13\text{T}_{mean}{}^2 - 0.18\text{T}_{SD} + 0.22\text{T}_{min} - 0.25\text{T}_{max} - 0.11\text{TP} - 0.14\text{NO}_x$	0.13	**731**
12	$1/k\text{T}_{mean}$	NS	–	$1/k\text{T}$	NS	–
13	$1/k\text{T}_{mean} + (1/k\text{T}_{mean})^2$	NS	–	$0.13/(k\text{T})^2$	0.02	787

Table 3. Finnish diatoms. Regression models (Eqn) testing the responses of species richness (models 1-11), ln-transformed species richness, ln(S) (models 12, 13), and residual parameter σ (σ_{res}, models 1-13) to spatial and environmental variables. Parameter σ_{res} was obtained from a regression of parameter σ against ln(S) (fig. A2B). The parameters in each model are standardized regression coefficients with $0.000003 \leq p < 0.05$. $n = 100$. AIC = Akaike Information Criterion. The AIC of the best environmental model is boldfaced. Abbreviations as in Table 1. NS = non-significant.

Eqn	Regression model for richness	R^2	AIC	Regression model for σ_{res}	R^2	AIC
1	$-0.58\mathrm{Lat}^2 - 0.33\mathrm{Lat}^3$	0.20	813	$0.38\mathrm{Lat}^3$	0.15	113
2	$0.31\mathrm{Long} + 0.36\mathrm{Long}^2$	0.22	812	$\mathrm{Long} + \mathrm{Long}^2 + \mathrm{Long}^3$	NS	–
3	$0.52\mathrm{Alt} + 0.22\mathrm{Alt}^2$	0.20	813	$0.27\mathrm{Alt}^3$	0.07	121
4	$-0.35\mathrm{Lat}^2 + 0.32\mathrm{Long}^2$	0.24	809	$0.38\mathrm{Lat}^3$	0.15	113
5	$-0.26\mathrm{T_{mean}}^2$	0.07	827	$-0.35\mathrm{T_{mean}}$	0.13	115
6	$\ln(N) + \ln(N)^2$	NA	–	$\ln(N) + \ln(N)^2$	NA	–
7	$0.34\mathrm{T_{SD}}$	0.11	822	$\mathrm{T_{SD}} + \mathrm{T_{SD}}^2$	NS	–
8	$0.55\mathrm{T_{min}} - 0.74\mathrm{T_{max}} - 1.05\mathrm{T_{max}}^2$	0.23	812	$-0.40\mathrm{T_{min}} + 0.24\mathrm{T_{max}}^2$	0.14	115
9	$-0.31\mathrm{T_{mean}}^2 + 2.39\mathrm{T_{SD}} + 3.29\mathrm{T_{min}} - 2.33\mathrm{T_{max}} - 0.91\mathrm{T_{max}}^2$	0.33	801	$1.39\mathrm{T_{mean}}^2 + 0.33\mathrm{T_{SD}}^2 - 0.44\mathrm{T_{min}} - 1.47\mathrm{T_{min}}^2$	0.17	116
10	$-0.42\mathrm{TP}^2 - 0.23\mathrm{pH}^2$	0.19	815	$-0.32\mathrm{Cond}^2$	0.10	118
11	$-3.25\mathrm{T_{mean}} - 0.97\mathrm{T_{mean}}^2 + 2.05\mathrm{T_{SD}} + 4.67\mathrm{T_{min}} + 0.72\mathrm{T_{min}}^2 - 0.29\mathrm{pH}^2$	0.38	**796**	$-0.51\mathrm{T_{mean}} - 0.28\mathrm{T_{min}}^2 + 0.32\mathrm{T_{max}}^2 + 0.43\mathrm{TP}$	0.22	**109**
12	$1/k\mathrm{T_{mean}}$	NS	–	$0.35/kT$	0.13	115
13	$-0.26/(k\mathrm{T_{mean}})^2$	0.07	827	$0.35/kT$	0.13	115

Figures

Figure 1. Species abundance distribution with a Poisson lognormal fit in two sample communities from the US, differing in species richness (S), parameter σ, proportion of species revealed by the sample (psr), and skewness (γ). NS = non-significant, * = significant.

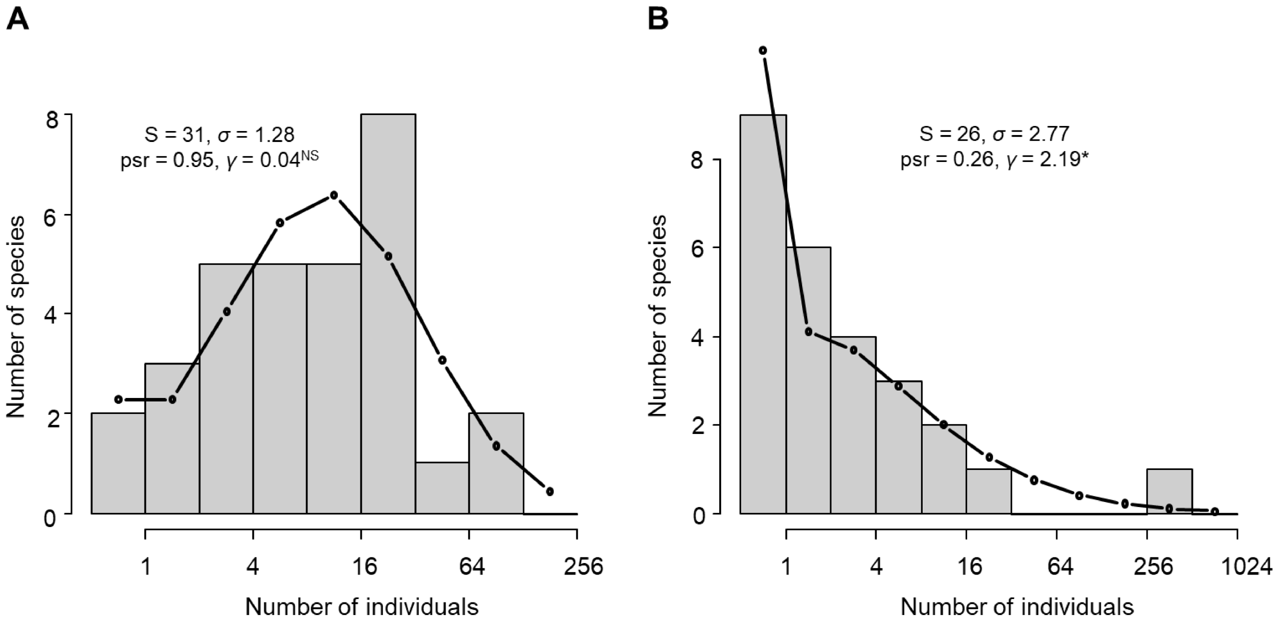


Figure 2. Longitudinal distributions of the best climatic and water chemistry predictors (A-D) of species richness and σ_{res} in the US (Table 2). The longitudinal trends of species richness (E) and σ_{res} (F) were the most pronounced spatial gradients of biotic variability in the US. The fits were generated by LOWESS with tension of 0.5. $n = 526$.

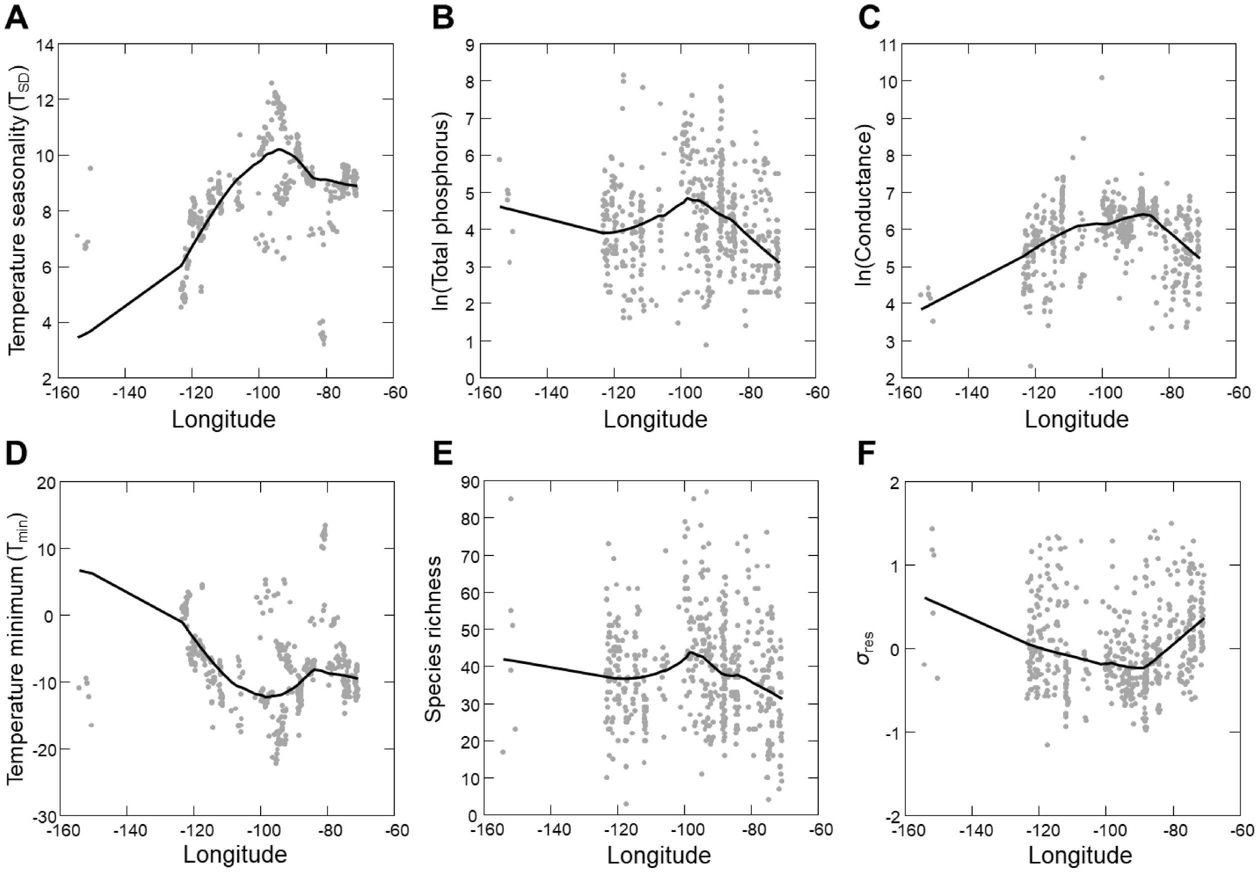


Figure 3. Latitudinal distributions of the best climatic and water chemistry predictors (A-E) of species richness and σ_{res} in Finland (Table 3). The latitudinal distributions of species richness (F) and σ_{res} (G) were among the strongest spatial trends of biotic variability in Finland. The fits were generated by LOWESS with tension of 0.5. $n = 100$.

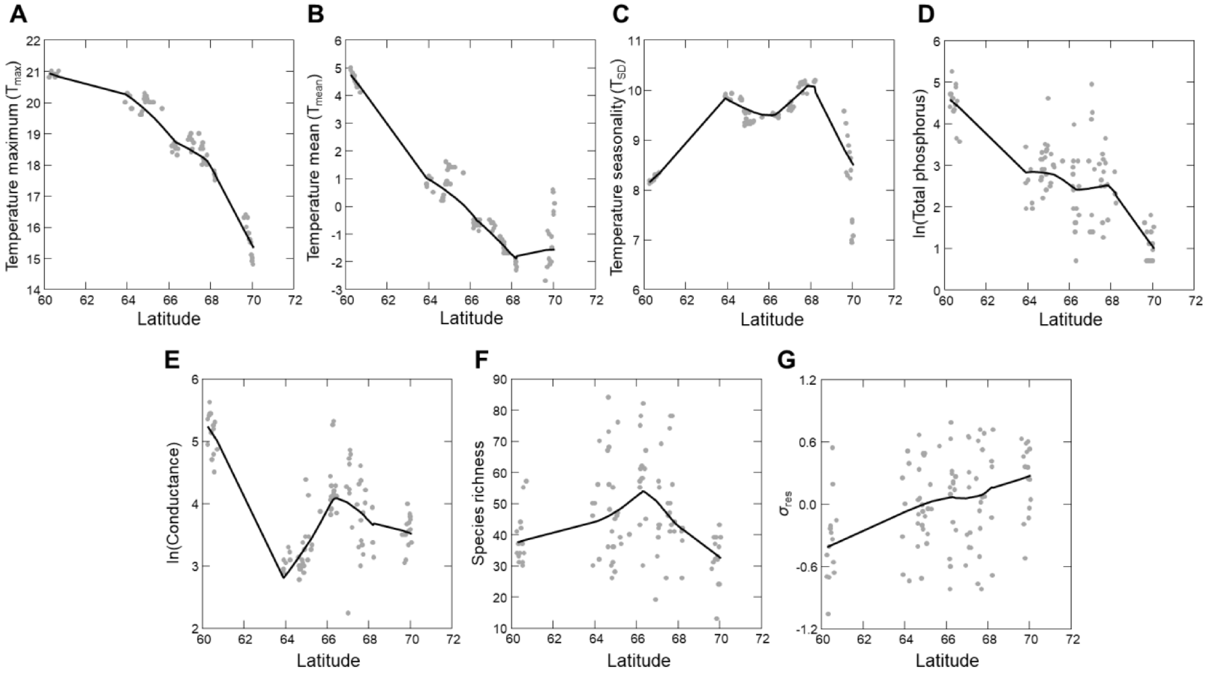


Figure 4. Regression trees of species richness, S (A) and σ_{res} (B) in the US, and S (C) and σ_{res} (D) in Finland, showing the splitting variables, their cut values, and the fit, derived from the proportional reduction in error. For both countries, the predictor set included T_{mean}, T_{SD}, T_{min}, T_{max}, ln(TP) (total phosphorus), ln(Cond) (specific conductance), and pH. For the US, ln(NO_x) and ln(NH_4) were also added. *n* = number of communities.

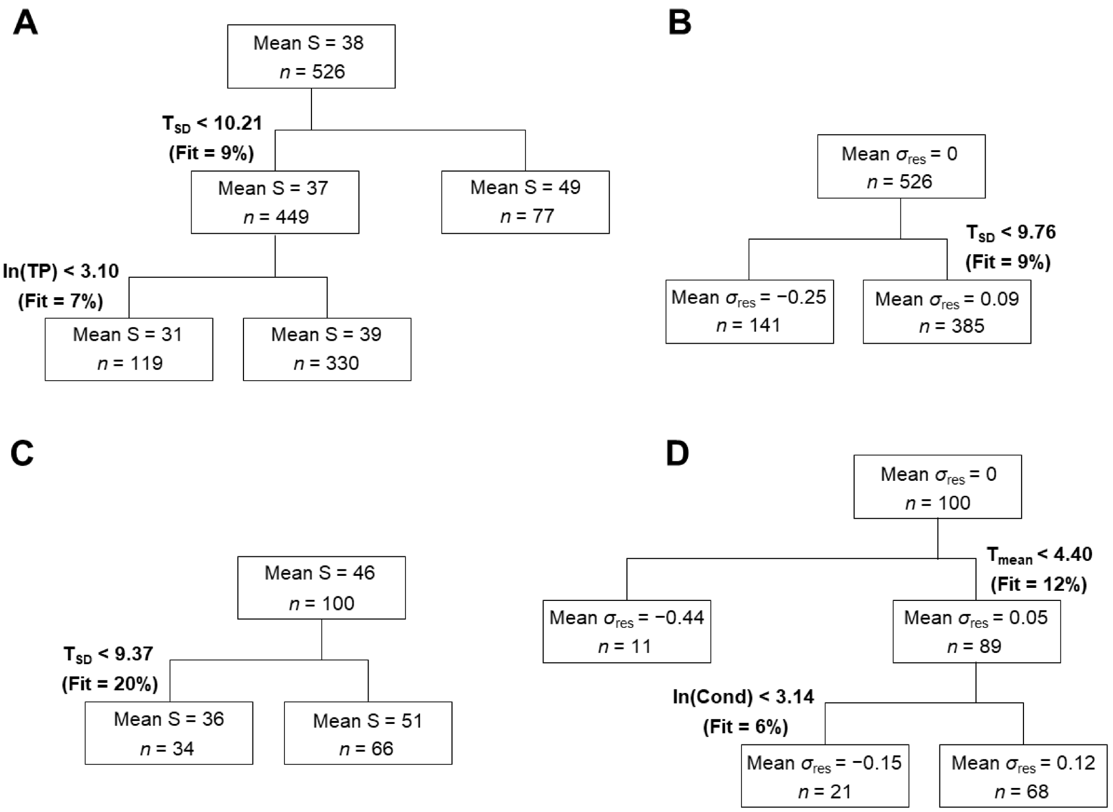


Figure 5. Variance partitioning of species richness (A) and σ_{res} (B) in the US, and species richness (C) and σ_{res} (D) in Finland using spatial, climatic, and water chemistry variables from models 4, 9, and 10, respectively, in Table 2 (A, B) and Table 3 (C, D). The numbers indicate percent explained variance.