Species richness, but not phylogenetic diversity, influences community biomass production and temporal stability in a re-examination of 16 grassland biodiversity studies

Venail, Patrick; Gross, Kevin; Oakley, Todd H.; Narwani, Anita; Allan, Eric; Flombaum, Pedro; Isbell, Forest; Joshi, Jasmin; Reich, Peter B.; Tilman, David; van Ruijven, Jasper; Cardinale, Bradley J.

How to cite

VENAIL, Patrick et al. Species richness, but not phylogenetic diversity, influences community biomass production and temporal stability in a re-examination of 16 grassland biodiversity studies. In: Functional ecology, 2015, vol. 29, n° 5, p. 615–626. doi: 10.1111/1365-2435.12432

This publication URL: https://archive-ouverte.unige.ch/unige:54883
Publication DOI: 10.1111/1365-2435.12432

Species richness, but not phylogenetic diversity, influences community biomass production and temporal stability in a re-examination of 16 grassland biodiversity studies

Patrick Venail[*, †, 1, 2], Kevin Gross[3], Todd H. Oakley[4], Anita Narwani[1,5], Eric Allan[6], Pedro Flombaum[7], Forest Isbell[8], Jasmin Joshi[9,10], Peter B. Reich[11,12], David Tilman[13,14], Jasper van Ruijven[15], Bradley J. Cardinale[1]

[1] *School of Natural Resources and Environment, University of Michigan, 440 Church Street, Ann Arbor, Michigan 48109, USA.*

[2] *Section of Earth and Environmental Sciences, Institute F.-A. Forel, University of Geneva, Versoix, Switzerland.*

[3] *Statistics Department, North Carolina State University, 2311 Stinson Drive, Raleigh, North Carolina 27695-8203, USA.*

[4] *Department of Ecology, Evolution and Marine Biology, University of California, Santa Barbara, California 93106-9620, USA.*

[5] *Aquatic Ecology, Eawag (Swiss Federal Institute of Aquatic Science and Technology), Dübendorf 8600, Switzerland.*

[6] *Institute of Plant Sciences, University of Bern, Altenbergrain 21, Bern, Switzerland.*

[7] *Centro de Investigaciones del Mar y la Atmósfera, Conicet/Universidad de Buenos Aires, C1428EGA Buenos Aires, Argentina.*

[8] *Department of Plant Biology, University of Georgia, 2502 Miller Plant Sciences, Athens, Georgia 30602, USA.*

[9] *Institute of Biochemistry and Biology, Biodiversity Research/Systematic Botany, University of Potsdam, Maulbeerallee 1, 14469 Potsdam, Germany.*

[10] *Berlin-Brandenburg Institute of Advanced Biodiversity Research (BBIB), Altensteinstr 6, 14195 Berlin, Germany.*

[11] *Department of Forest Resources, University of Minnesota, 1530 Cleveland Avenue N, Saint Paul, Minnesota, USA.*

[12] *Hawkesbury Institute for the Environment, University of Western Sydney, Penrith NSW 2751, Australia.*

[13] *College of Biological Sciences, University of Minnesota, 1987 Upper Buford Circle , Saint Paul, Minnesota 55108, USA.*

[14] *Bren School of Environmental Science and Management, University of California, Santa Barbara, CA 93106, USA.*

[15] *Department of Nature Conservation and Plant Ecology, Wageningen University, Droevendaalsesteeg 3, 6708PB Wageningen University, The Netherlands.*

*Corresponding author: E-mail: Patrick.Venail@unige.ch

†Current address: Institut F.A. Forel, University of Geneva, Route de Suisse 10, 1290 Versoix, Switzerland.

Summary

1. Hundreds of experiments have now manipulated species richness of various groups of organisms and examined how this aspect of biological diversity influences ecosystem functioning. Ecologists have recently expanded this field to look at whether phylogenetic diversity among species, often quantified as the sum of branch lengths on a molecular phylogeny leading to all species in a community, also predicts ecological function. Some have hypothesized that phylogenetic divergence should be a superior predictor of ecological function than species richness because evolutionary relatedness

represents the degree of ecological and functional differentiation among species. But studies to date have provided mixed support for this hypothesis.

2. Here, we re-analyze data from 16 experiments that have manipulated plant species richness in grassland ecosystems and examined the impact on aboveground biomass production over multiple time points. Using a new molecular phylogeny of the plant species used in these experiments, we quantified how the phylogenetic diversity of plants impacts average community biomass production as well as the stability of community biomass production through time.
3. Using four complementary analyses we show that, after statistically controlling for variation in species richness, phylogenetic diversity (the sum of branches in a molecular phylogenetic tree connecting all species in a community) is neither related to mean community biomass nor to the temporal stability of biomass. These results run counter to past claims. However, after controlling for species richness, phylogenetic diversity was positively related to variation in community biomass over time due to an increase in the variances of individual species, but this relationship was not strong enough to influence community stability.
4. In contrast to the non-significant relationships between phylogenetic diversity, biomass, and stability, our analyses show that species richness *per se* tends to increase the mean biomass production of plant communities, after controlling for phylogenetic diversity. The relationship between species richness and temporal variation in community biomass was either positive, non-significant or negative depending on which analysis was used. However, the increases in community biomass with species richness, independently of phylogenetic diversity, always led to increased stability. These results

suggest that phylogenetic diversity is no better as a predictor of ecosystem functioning than species richness.

5. *Synthesis.* Our study on grasslands offers a cautionary tale when trying to relate phylogenetic diversity to ecosystem functioning suggesting that there may be ecologically important trait and functional variation among species that is not explained by phylogenetic relatedness. Our results fail to support the hypothesis that the conservation of evolutionarily distinct species would be more effective than the conservation of species richness as a way to maintain productive and stable communities under changing environmental conditions.

Key-words: biodiversity**,** community biomass**,** data-synthesis**,** ecosystem functioning, grasslands, phylogenetic diversity, relatedness, stability.

Introduction

Over the past few decades, ecologists have completed hundreds of experiments exploring how the variety of life forms influences the fluxes of carbon and cycling of elements that control how ecosystems ‘function’ (Schulze & Mooney 1993, Tilman & Downing 1994). To date, the field of biodiversity and ecosystem functioning (BEF for short) has been dominated by studies that used species richness as their sole measure of biodiversity (Loreau *et al.* 2001, Hooper *et al.* 2005, Cardinale *et al.* 2011). However, ecologists have recently begun to explore how other aspects of diversity like genetic and trait variation can influence the functioning of ecosystems, and begun to ask whether certain measures of diversity are better predictors of ecosystem functioning than others (Diaz & Cabido 2001, Petchey & Gaston 2006, Cadotte *et al.* 2008, Cadotte *et al.* 2012). One form of diversity that has received a growing amount of attention is phylogenetic diversity. Phylogenetic diversity is a measure of how much evolutionary divergence has occurred among the species in a community, often measured as the cumulative branch length differences that separate species on their molecular phylogeny. There are several reasons why ecologists have become interested in using phylogenetic diversity to predict ecosystem-level processes. First, this interest is part of a general trend to understand contemporary ecological patterns by looking at the evolutionary history of organisms in a community (Webb *et al.* 2002, Johnson & Stinchcombe 2007). As in the field of 'community phylogenetics', researchers in the field of biodiversity and ecosystem functioning have begun to think about how ecological and evolutionary processes might interact to control the functioning of ecosystems. Second and more importantly, ecologists have been enticed by the simplicity of using phylogenetics to predict ecological function. While it is difficult and time consuming to run manipulative experiments of species richness, and equally difficult to identify and measure the myriad of species traits that control the functioning of ecosystems, getting genetic information needed to characterize species relationships and thus to measure phylogenetic diversity, has become an increasingly straightforward task.

The ability to use phylogenetic diversity to predict ecological function is predicated on a sequence of assumptions that have rarely been tested directly, especially in an integrated fashion. The first assumption is that the biological traits that control ecological functions show a phylogenetic signal, meaning they tend to be more similar among closely related species than between distantly related species (Prinzing *et al.* 2001, Losos 2008, Wiens *et al.* 2010, Cavender-Bares & Reich 2012). The second assumption is that, when traits do show a phylogenetic signal, the trait variation leads to functional differentiation among species. The third and final assumption is that such functional differentiation enhances the productivity or stability of an entire community. Productivity might be enhanced if, for example, expression of a greater variety of traits allows species to better exploit all of the available resources (Tilman *et al.* 1997, Loreau 2004, Reich *et al.* 2012, Srivastava *et al.* 2012). To date, the influence of phylogenetic diversity on ecosystem functioning has been explored in just twelve studies that we know of, and these span a relatively small number of systems (see **Table 1**). Eight of these have found a positive relationship between phylogenetic diversity and various aspects of ecosystem functioning, one found a negative relationship, and three showed either mixed results or more complex non-linear relationships. In the instances where phylogenetic diversity was positively related to ecosystem functioning, it tended to explain only a small fraction more of the variation than species richness (Cadotte *et al.* 2008, Cadotte *et al.* 2009, Cadotte 2013, but see Paquette & Messier 2011, Cadotte *et al.* 2012). Nonetheless, authors of these studies tend to strongly advocate for the importance of PD for ecosystem functioning.

Many studies have also shown that diverse communities have more temporally stable biomass production than less diverse communities. In most cases the temporal stability of community biomass production is commonly measured as the inverse of its coefficient of variation over time (Tilman 1999, Jiang & Pu 2009, Hector *et al.* 2010, Campbell *et al.* 2011), which is the biomass of the community averaged over time divided by its standard deviation through time. The standard deviation of community biomass can be influenced by changes in variances of individual species' biomasses as well as by changes in the synchrony of species' biomass fluctuations over time. Thus, diversity can influence temporal community biomass stability through the average biomass production of the community or through individual species' biomasses (e.g., their synchrony). Higher community biomass, lower sum of species variances and more asynchronized fluctuations of species' biomasses would increase community stability. Assuming communities with higher phylogenetic diversity result in the expression of a greater variety of traits allowing species to better exploit resources, it can be predicted that the average community biomass will increase with phylogenetic diversity. Similarly, a greater variety of traits (assumed to be represented by a higher phylogenetic diversity) should allow communities to show a greater array of compensatory dynamics (Tilman 1999, Hector *et al.* 2010, Violle *et al.* 2011, Cadotte *et al.* 2012, Verdu *et al.* 2012), reducing the standard deviation of community biomass over time. Overall, the temporal stability of community biomass, measured as the average community biomass divided by its standard deviation is expected to increase as phylogenetic diversity increases. To date, only three studies have explored the influence of phylogenetic diversity on the temporal stability of ecosystem function (**Table 1**). One found a positive effect of phylogenetic diversity on the stability of plant biomass in grasslands (Cadotte *et al.* 2012), one found a negative effect on the stability of algal biomass in microcosms (Venail *et al.* 2013), and one found a non-linear (U-shaped) relationship between phylogenetic diversity and the stability of protists' biomass in microcosms (Pu *et al*. 2014). The relatively small number of studies and their equivocal results suggest more comprehensive studies are needed.

Here, we reanalyze data from 16 biodiversity experiments using grassland plants to better assess how phylogenetic diversity influences the production of biomass and its stability over time. Twelve of the studies used here are a subset of the 29 studies used by Cadotte *et al.* (2008) to examine how phylogenetic diversity impacts biomass production, and all 16 studies are the same studies whose time-series were used by Cardinale *et al.* (2011) and Gross *et al.* (2014) to examine how species richness influences the stability of biomass production through time. The primary advance of our study is that we use four different complementary analytical methods to separate the effects of phylogenetic diversity and species richness on community productivity and stability. These two forms of diversity are inherently correlated since a greater number of species almost always correlates with greater summed genetic divergence on a phylogeny. However, this correlation has not been adequately dealt with in prior studies and, as we will show, our analyses lead to several modified conclusions about the role of species vs. phylogenetic diversity in ecosystem functioning.

Methods

Data

Our study represents a new data synthesis of 16 previously published studies that have examined the relationship between plant biodiversity and the production and stability of population and community-level biomass in grasslands. Data from these studies were previously compiled for use in other data syntheses (Cadotte *et al.* 2008, Cardinale *et al.* 2011, Gross *et al.* 2014) where studies were chosen based on the following criteria: 1) experiments had to be performed in grasslands, which is the system most frequently studied in BEF research and for which the most data are available; 2) studies had to include estimates of net annual aboveground plant biomass production or aerial coverage; 3) studies had to include at least three sampling occasions performed over time, thus allowing estimation of temporal stability; and 4) studies had to include species-level data for each experimental plot, thus allowing measurement of responses to environmental fluctuations of individual species in polycultures (which is necessary for certain calculations of stability). Only sixteen studies met all four of these criteria (**Table S1**). All the data used in the current analysis are available in dryad (http://datadryad.org/).

Molecular phylogeny

We estimated phylogenetic relationships of 141 plant species used in the experimental plots plus two outgroups (*Magnolia grandiflora* and *Amborella trichopoda,* **Fig. S1**). For this, we used publicly available genetic data from 6 gene sequences commonly used in angiosperm phylogenetics: *matk*, *rbcl*, *ndhf*, *its1*, *its2*, and *5.8s*. All but 14 species had publicly available genetic data from at least one of the target genes. To represent each species that had none of these genes available (*Amorpha canadensis, Anemone cylindrica, Bothriochloa laguroides, Conyza albida*, *Dalea villosa, Medicago varia, Mulinum spinosum, Nassella leucotricha, Pimpinella major, Poa ligularis, Salvia azurea, Sporobolus compositus, Stipa speciosa, Symphyotrichum oolentangiense*), we randomly chose a representative congener with target genes publicly available (*Amorpha fruticosa, Anemone patens, Bothriochloa insculpta, Conyza gouanii, Dalea brachystachya, Medicago sativa, Mulinum chillanense, Nassella pampagrandensis*, *Pimpinella betsileensis*, *Poa sichotensis, Salvia przewalskii, Sporobolus atrovirens, Stipa stenophylla, Symphyotrichum ericoides*). Accession numbers for all genes used are reported in **Table S2**. We aligned all sequences of each gene using MUSCLE (Edgar 2004). We concatenated all genes using phylocatenator (Oakley *et al.* 2014) and estimated a maximum likelihood phylogeny using RAxML (Stamatakis & Ott 2008), along with 100 bootstrap pseudoreplications to gauge nodal support. We conducted all phylogenetic analyses in the Osiris package (Oakley *et al.* 2014) of Galaxy, which allows us to easily share all data and analyses with a web link (http://galaxy-dev.cnsi.ucsb.edu/osiris/u/ostratodd/h/plant-pd-venail).

To estimate the evolutionary relatedness among species in a plot we used phylogenetic diversity, defined as the total phylogenetic distance among two or more species (Faith 1992, Cadotte *et al.* 2008). Thus, the phylogenetic diversity of an assemblage (plot) is influenced both by the number of species and by their level of evolutionary relatedness. Phylogenetic diversity is inversely proportional to the evolutionary relatedness of the species, thus the more distantly related a set of species becomes, the higher the phylogenetic diversity will be. We used Picante in R (Kembel *et al.* 2010) to calculate different phylogenetic diversity metrics including phylogenetic distance (PD, Cadotte et al. 2008), mean phylogenetic distance (MPD, Webb *et al.* 2008), mean nearest taxon distance (MNTD, Webb *et al.* 2008) and phylogenetic species variability (PSV, Helmus *et al.* 2007) for each plot (data available in dryad). We assessed the sensitivity of our estimates of phylogenetic diversity to different phylogenies by comparing our values with those obtained using a recently published plant phylogeny (Zanne *et al.* 2014). That phylogeny, like ours, is based on a Maximum Likelihood analysis of GenBank data. The Zanne *et al.* tree used seven gene regions from GenBank, so there is substantial overlap of primary data with our phylogeny. The Zanne *et al.* analysis differs from ours in that those authors constrained major clades (families and orders), partitioned data by gene regions, and smoothed their tree using divergence time estimates.

Productivity & Stability

We focus on the influence of biodiversity on both 1) the production and 2) temporal stability of biomass produced by mixtures of grassland plant species grown in polyculture. At each time point community biomass production was estimated as the sum of the biomass produced by all the species in a plot. Then, we averaged community (plot) biomass over time. Most estimates of biomass production in the datasets are in units of mass per area; however, two studies used estimates of aerial plant coverage instead (studies 12 and 15, **Table S1**). For consistency with previous data-synthesis (Cardinale *et al.* 2011, Gross *et al.* 2014) we did not transform the data from these two studies.

The most commonly used measure of temporal variability in community biomass is the coefficient of variation (Jiang & Pu 2009, Hector *et al.* 2010, Campbell *et al.* 2011), which is the standard deviation of community biomass through time scaled to account for the average biomass of the community. Temporal community stability is then the inverse of the coefficient of variation:

$$stability = \frac{\bar{x}}{sd}$$ **Equation 1**

Thus, community stability can be influenced both by changes in the average biomass production (numerator of **equation 1**) or by changes in the temporal standard deviation of biomass production (denominator of **equation 1**). The standard deviation can be further decomposed into the sum of population-level variances and the covariances among species' biomasses through time. The covariance in species biomasses is frequently used as a measure of the degree of synchrony in the temporal variation of species' population responses (Jiang & Pu 2009). However, when more than two species are present in an assemblage, it is now known that covariance is an inappropriate measure of species synchrony because the covariance depends on both the number of species and the synchrony among them (Loreau & de Mazancourt 2008, 2013). This limitation has hindered interpretation of what most contributes to stability, and has led to efforts to develop new metrics of species synchrony (Loreau & de Mazancourt 2008, Gross *et al.* 2014). Here, we used the most recent metric developed by Gross *et al.* (2014), which measures synchrony among species' biomasses as the average correlation between the biomass of each species (Y_i) and the total biomass of all other species in the group ($\sum_{j \neq i} Y_j$).

$$synchrony = \left(\frac{1}{n}\right) \sum_i corr\left(Y_i, \sum_{j \neq i} Y_j\right) \qquad \textbf{Equation 2}$$

A *synchrony* value close to -1 suggests species are maximally asynchronized, a value close to +1 that species are maximally synchronized and values close to 0 that species fluctuate independently.

To summarize, in our analyses we used phylogenetic diversity and species richness as explanatory variables. Stability and its different components (average biomass and standard deviation, **equation 1**) as well as the sum of variances and synchrony (**equation 2**) were used as response (dependent) variables.

Statistical Analyses

Within the full dataset we assembled, which contains 824 experimental plots spread across 16 studies, measures of phylogenetic diversity (PD) and species richness (SR) were highly correlated with one another (**Fig. 1a**, *r* = 0.90). This high degree of correlation is not surprising given that PD is not only influenced by the branch lengths separating species on a phylogeny (i.e., their relatedness), but also by the number of species being considered. Importantly, these 16 experiments were not originally designed to produce a wide range of

PD values or to manipulate phylogenetic diversity independently of species richness. Therefore, the high degree of correlation leads to statistical problems of multi-collinearity in many forms of data analyses, making it difficult to draw robust conclusions about the influence of PD per se, or SR per se on biomass production and stability.

In an attempt to disentangle the effects of PD and SR on community biomass production and temporal stability in community biomass, we performed four unique analyses on this dataset (**Fig. 1**). These are described as follows:

Type 1 analysis: In this analysis (**Fig. 1a**) we quantified the effect of PD on community stability, community biomass production (**equation 1**), standard deviation, sum of variances and the synchrony metric (**equation 2**) within levels of species richness (i.e., holding SR constant). The original dataset included species assemblages that spanned a wide array of planted species richness levels (from 2 to 60). However, we focused on richness levels 2, 3, 4, 6, 8, 9, 11, 12 and 16 species (for 716 plots in total) because these were the richness levels for which multiple studies were represented and each level of richness had multiple values of phylogenetic distance (i.e., different species compositions). For each study and within each species richness level, we calculated the correlation between phylogenetic distance (PD) and each of five response variables: 1) temporally averaged community-level biomass (biomass summed across all species in a plot at each time point, then averaged over time; numerator in right side of **equation 1**), 2) the temporal standard deviation of community biomass (denominator in right side of **equation 1),** 3) the community-level temporal stability of biomass (left side of **equation 1**), 4) the summed variances of individual species' biomasses and 5) population-level temporal synchrony (as in **equation 2**). Correlation coefficients were weighted by the number of plots in each study to reduce the influence of poorly replicated studies. We normalized the distribution of data using Fisher's z-algorithm (Z_r; Balvanera *et al.* 2006) to test if for each of the five response variables the weighted/normalized correlation coefficients (Z_r) were significantly different from zero using double-tailed t-tests.

Type 2 analysis: Unlike the *type 1 analysis* where we were able to analyze the impact of phylogenetic diversity (PD) on production and stability with species richness (SR) held constant, a directly comparable analysis looking at the effects of SR with PD held constant is not straightforward. This is because PD represents a continuous measure that cannot be binned into categories in the same way SR can. Nevertheless, in our *type 2 analysis* (**Fig. 1a**) we were able to identify a large number of experimental plots that were relatively similar in values of PD, but which had differing levels of SR. For each study, we compared every plot to every other plot in the study. We found a total of 1,417 pairs of plots, with each pair belonging to the same study where PD differed by <10%, but for which species richness differed. When compared to random sampling of plots, these paired contrasts represent a highly constrained range of variation in PD, and come as close as is reasonably possible to holding PD constant while allowing SR to vary (**Fig. S3**). For each of these 1,417 pair-wise contrasts, we calculated the log ratios of community biomass and stability, $\ln(Y_{\text{high richness}}/Y_{\text{low richness}})$, where Y represents: a) total plot biomass, b) standard deviation of biomass, c) temporal stability of biomass through time or d) the sum of variances for the high vs. the low species richness plot within each pair. Positive log ratios indicate that the more speciose community either: produces more biomass, has a higher standard deviation in biomass through time, is more stable than the less speciose community or has more variable species. We used two tailed t-tests to evaluate if log-ratios for each metric were different from zero. We could not establish log ratios for synchrony because synchrony can have negative values and it is not possible to calculate a logarithm of negative values.

Type 3 analysis: In this analysis we used Structural Equation Modelling (SEM) to summarize data from five experiments (studies 5, 6, 9, 13 and 14 in **Table S1**) where the species pools used led to relatively low correlation coefficients relating PD and SR (**Fig. 1b**, N = 5 studies, r = 0.72 using 222 experimental plots). While these five studies represent but a subset of available data, the correlations between PD and SR in all other studies were well above 0.8, rendering them unusable in any attempts to statistically control for covariance among SR and PD in a single analysis. However, for this subset of five studies, it was possible to statistically control for the covariance between SR and PD. In turn, the SEM allowed us to calculate the partial regression coefficients that represent the unique coefficients relating both PD and SR to community biomass and the standard deviation of biomass through time. We did not incorporate the sum of variances and synchrony into *type 3* analyses because clear causal pathways have yet to be established.

Type 4 analysis: In *type 1, 2* and *3 analysis* we use PD as a metric of phylogenetic diversity, which is the metric used in most previously published studies (8 out of 12 listed in **Table 1** used it). However, other metrics of evolutionary relatedness have been developed; among the more common are the mean pairwise distance (MPD), mean nearest taxon distance (MNTD, Webb *et al.* 2008) and phylogenetic species variability (PSV, Helmus *et al.* 2007). Some of these have been proposed to be less correlated to species richness than PD (**Fig. 1c**, **Fig. S4**) and would, in principle, reduce statistical problems related to multi-collinearity. However, there are concerns about more advanced metrics like MPD and MNTD because they count each branch of the phylogenetic tree multiple times depending on the number of species in a plot (e.g., in a plot with n species each branch is counted *n-1* times). We complemented our three other types of analyses with *type 4 analysis* that used linear mixed effect (LME) models to explore the impact of species richness (SR) and mean pairwise distance (MPD) on all five dependent variables: stability, average biomass, standard deviation, sum of variances and synchrony. Analysis using MNTD and PSV would lead to the same results given their strong correlation with MPD (**Fig. S2**). All our LME models also included "study" as random effects.

Results

Phylogeny

The topology of the phylogeny of grassland plants included in the current study (**Fig. S1**) is very similar to a previous study that used similar methods ($\rho = 0.947$, $p < 0,001$; Cadotte *et al.* 2008). As expected, we found support for two major ingroup clades, Poales and eudicots. Forty-one nodes are supported by 100% bootstrap values. Twenty nodes showed lower than 50% bootstrap support, suggesting uncertainty in these nodes. In previous studies (e.g., Cadotte *et al.* 2008) sensitivity analyses using different phylogenetic approaches indicated that correlations between phylogenetic diversity and other variables were very minimally affected by differences in tree topology. Again, we found very similar values of phylogenetic diversity based on our new tree compared to values obtained with a recently published tree that used different (but overlapping) primary data and made different assumptions (Zanne *et al.* 2014, **Fig. S2**). Values for the four different phylogenetic diversity metrics assuming the two different phylogenetic analyses for each community are available in dryad.

Type 1 analysis: effect of PD within richness levels.

For each level of species richness considered, studies showed highly variable effects of phylogenetic diversity (PD) on stability, average biomass production, standard deviation (SD), the sum of species variances (sum. var.) and synchrony, ranging from negative to positive relationships (**Fig. 2a**). Of these, only a limited set of studies had any significant relationship between PD and community stability or its different components (**Fig. 2a**). When the correlation coefficients were weighted and averaged across all experiments, there was a tendency for PD to be negatively related to temporal community stability and positively related to average community biomass production, though neither of these trends were significantly different from zero at the $p = 0.05$ level of significance (**Fig. 2b**). PD was, however, positively correlated with temporal variation in community biomass (SD biomass), a trend that was driven by an increase in the summed variance across species, rather than by a change in the synchrony of species' biomasses through time (**Fig. 2b**).

Type 2 analysis: effect of SR within PD bins.

When we performed pair-wise comparisons among plots that differed in SR but had similar PD (values differing by less than 10%), the temporal stability of biomass and the average biomass both significantly increased as a function of SR (**Fig. 3**). In contrast, the standard deviation of community biomass through time (S.D.) was negatively influenced by SR. The sum of species variances (sum. var.) was not affected by species richness.

Type 3 analysis: Effect of PD and SR after accounting for their covariance.

After accounting for the covariance between SR and PD in the five experiments with the lowest correlations (mean $r = 0.72$, $p < 0.05$, $n = 222$), a path analysis suggested that SR was positively associated with mean plot biomass ($r = 0.39$, $p < 0.01$) and with the standard deviation of biomass over time ($r = 0.20$, $p < 0.05$, **Fig. 4**). Therefore, there were positive indirect effects of SR on community stability that were mediated through the increase in biomass ($r = 0.30$, $p < 0.01$) and variance ($r = -0.21$, $p < 0.01$, **Table S3**). In contrast, PD was not associated with the standard deviation of biomass over time ($r = 0.10$, $p > 0.05$) or with any change in the mean community biomass ($r = 0.003$, $p > 0.05$, **Fig. 4**). Therefore, there were no indirect effects of PD on community stability via biomass ($r = 0.002$, $p > 0.05$) or variance ($r = -0.11$, $p > 0.05$, **Table S3**).

Type 4 analysis: Effect of MPD and SR.

Linear mixed effect (LME) models with species richness (SR), mean pairwise distance (MPD, both as fixed effects) and study (as random effect) on the five different dependent variables revealed a positive effect of species richness on stability, average biomass, standard deviation (S.D.) and synchrony, but no effect on the sum of species variances (sum. var., **Table S4**). Phylogenetic diversity (measured as MPD) had positive effects on S.D. driven by a positive effect on the sum of species variances, but had no effect on stability, average biomass or synchrony.

Summary of results

Table 2 summarizes results of the different types of analyses, which were consistent in showing a positive relationship between species richness (SR) and biomass production after controlling for phylogenetic diversity. Analyses disagreed in how SR influences the standard deviation of biomass through time. Type 2 analyses showed a negative influence of SR on S.D. but with an absence of effect on the sum of variances. Type 3 showed a positive effect on S.D. whereas Type 4 showed no effect on S.D., with type 4 also revealing no effect on the sum of species variances but a positive effect of SR on synchrony. Ultimately all the analyses converged in showing that species richness has a positive influence on community biomass stability, via the increase in average community biomass.

Analyses were also consistent in showing that phylogenetic diversity (measured as PD for type 1-3 analyses, and as MPD for type 4 analysis), after controlling for species richness, did not explain any significant variation in mean community biomass. While there was a positive effect of phylogenetic diversity (either as PD in Type 1 and MPD in Type 4 analysis respectively) on the standard deviation of biomass over time, driven by a positive effect on the sum of variances but not on synchrony, this was not sufficiently large to generate a decrease in community stability as PD increased.

Discussion

Here, we re-analyzed data from sixteen experiments that manipulated plant species richness in grassland ecosystems to examine how species richness and phylogenetic diversity influence mean community biomass and its temporal stability. The primary advance of our study was to use a variety of analyses that attempt to control for the inherent positive covariance between species richness and phylogenetic diversity so that we could try to tease apart their effects. Consistent with the results of many individual studies (e.g., almost all of those referenced in **Table 1**, among others) and prior data syntheses (e.g., Balvanera *et al.* 2006, Cardinale *et al.* 2006, Cadotte *et al.* 2008, Cardinale *et al.* 2011, Flynn *et al.* 2011, Gross *et al.* 2014), our analyses confirmed that plant communities composed of more species tend to produce greater community level biomass and to be more stable over time. This result held true even after controlling for variation in the phylogenetic diversity of species, suggesting that the impact of species richness on biomass production and temporal stability cannot be explained fully by differences in phylogenetic diversity among communities.

Although our analyses confirmed prior conclusions about the positive effect of species richness on community biomass production and stability after controlling for variation in phylogenetic diversity, the reverse was not true. We found no evidence that, after controlling for variation in species richness, phylogenetic diversity was related to community biomass production or its temporal stability in grasslands. Despite this absence of any effect on the average community biomass and stability, two of our analyses revealed a positive effect of phylogenetic diversity on the standard deviation (S.D.) of community biomass. Examination of the sum of species variances and synchrony components suggest that the increase in community biomass standard deviation (S.D.) was driven by an increase in the sum of individual species variances and not by changes in the synchrony of their fluctuations. This suggests that closely related species share low biomass variation over time but these similarities vanish as species become less related, providing some evidence of a phylogenetic signal in the temporal variation of species' biomass. A recent study by Godoy *et al.* (2014) found that fitness differences among annual plants were higher and much more variable between distantly than closely related species, suggesting that the outcome of competition should be more variable between more distantly related species. It is possible that such increased competitive variability with increasing phylogenetic diversity lead to an increase in biomass variability over time. Though, the observed increase in the sum of variances with PD could also be due to a higher probability of the presence of species with higher biomass variability in plots with higher phylogenetic diversity (i.e., sampling effect).

Our general conclusion about the lack of effect of phylogenetic diversity on community biomass differs from the conclusions of two previous data-syntheses (Cadotte *et al.* 2008, Flynn *et al.* 2011). Cadotte *et al.* (2008) summarized data from 29 experiments that manipulated richness of terrestrial angiosperms and asked whether phylogenetic diversity could explain variation in a standardized diversity ‘‘effect size’’ (the log ratio of biomass in a polyculture / the mean biomass of the constituent species in monoculture). The authors concluded that *“the amount of phylogenetic diversity within communities explained significantly more variation in plant community biomass than other measures of diversity, such as the number of species or functional groups”.* In an attempt to deal with the strong covariance between species richness and phylogenetic diversity, Cadotte *et al.* (2008) examined how phylogenetic diversity related to diversity effect sizes within levels of species richness. They found that phylogenetic diversity was only related to diversity effects at the lowest levels of richness (i.e., 2 and 4 species), and suggested this was because researchers tended to use fewer species combinations at high levels of richness (i.e., 6 and 8 species), resulting in less variation in phylogenetic diversity.

The study by Cadotte *et al.* (2008) differs from ours in several aspects. First, only twelve of the studies included in our analyses overlapped with those included in their dataset. This is because we only included studies providing community biomass for at least three different time-points so that we could quantify temporal stability. Second, the phylogenetic trees used to calculate the phylogenetic diversity within plots in our study and the Cadotte *et al*. 2008 study were though similar but not exactly the same. Third, the response variables used in our studies were different; we used the mean biomass across time-series, as opposed to a log response ratio at a single time point used in Cadotte *et al.*’s study. Finally, the statistical analyses also differed among studies. Cadotte *et al.* used linear mixed effect models with species richness and phylogenetic diversity as explanatory variables despite the fact these two variables were strongly correlated. In our study, to avoid the problems related to covariance of the explanatory variables, we calculated correlation coefficients between phylogenetic diversity and community biomass at each level of species richness and for each individual study. Then we weighted and averaged the correlation coefficients among studies and richness levels.

In order to determine which of these four factors were responsible for the discrepancies in the results among both studies, we collated a dataset that contained the explanatory (i.e., phylogenetic diversity) and dependent variables (i.e., community biomass) from both studies. This resulted in an overlapping dataset that included 318 plots. We tested the effect of switching the two metrics of phylogenetic diversity, the two measures of community biomass and the two statistical analyses from both studies by performing a series of permutations using the collated dataset (see Supplementary Material **S4**). The permutations revealed that the conclusions from both studies about the effect of phylogenetic diversity on community biomass (i.e., positive for *Cadotte et al.'s* and no effect for this study) differed because they focused on different measures of community biomass and used different statistical approaches. This is not surprising, given that both studies were answering different questions related to the role of phylogenetic diversity as a predictor of community biomass as we explained before. We consider that for the purposes of our study, which was to separate the effects of SR and PD, the statistical approach based on coefficients of correlations is more appropriate because it avoids problems due to collinearity between species richness and phylogenetic diversity. Moreover, the lack of effect of phylogenetic diversity (as PD) on community biomass was confirmed by a linear mixed effect models using MPD as the explanatory variable. While useful for addressing questions related to the effect of diversity on ecosystem functioning, log ratios open the possibility that the observed differences in community biomass are due to differences in the monoculture biomasses of the constituent species, which seemed to be the case here (see Supplementary Material **S4)**. For instance, having monocultures with lower average biomass would result in higher community biomass if estimated as log ratios. Thus, to allow a clearer interpretation of the differences in biomass among communities we preferred to directly analyze raw community biomass.

Our results also deviate from the conclusions of another prominent data synthesis by Flynn *et al.* (2011), who added measures of functional diversity (i.e., trait variation among plant species on the phylogeny) to Cadotte *et al.*'s dataset and tested to see if functional diversity was a superior predictor of biomass production than phylogenetic diversity. The authors ran a variety of models comparing the explanatory power of phylogenetic diversity alone, functional diversity alone, both together, as well as in combination with species richness. They concluded that a model containing only phylogenetic diversity was the most likely explanation of variation in plant biomass among plots (see Table 2 in their paper). But Flynn *et al.* did not statistically control for the covariance between species richness and phylogenetic diversity when drawing their conclusions. Because none of their linear mixed models (Table 2 in Flynn *et al.* 2011) accounted for covariance among variables, nor did any of their multivariate analyses (see the Structural Equations models presented in their Figure 3 and their Appendix), we cannot judge how their findings relate to our own. While our results do not directly contradict previous findings, given that we were addressing related but different questions, the contrast in our conclusions leads us to believe that former statements about the strong impacts of phylogenetic diversity on community biomass may have been partly driven by the strong correlation between phylogenetic diversity and species richness. When we control for the collinearity between species richness and phylogenetic diversity, the residual effects of phylogenetic diversity on community biomass are non-significant.

The recent incorporation of phylogenetic information into biodiversity-ecosystem functioning (BEF) studies, and into community ecology research in general, was motivated at least partially by the relative ease of measuring phylogenetic distances among species compared to measures of their functional differentiation (Cadotte *et al.* 2008, Srivastava *et al.* 2012). With the increased availability of updated phylogenies, some had hoped that phylogenetic diversity metrics would summarize information on ecological traits and thus predict ecosystem function. Our results, showing an absence of effect of phylogenetic diversity on average community biomass and its temporal stability in grassland communities, run counter to this expectation.

The use of phylogenetic diversity as a predictor of ecosystem functioning assumes that evolutionary distance and ecological differentiation are positively related, with close relatives being ecologically more similar than distant relatives (i.e., phylogenetic signal; Losos 2008, Wiens *et al.* 2010). There are currently divergent positions on whether or not the evolutionary relatedness among modern species is a reasonable proxy for ecological similarity (Prinzing *et al.* 2001, Freckleton *et al.* 2002, Johnson & Stinchcombe 2007, Losos 2008, Wiens *et al.* 2010, Cavender-Bares & Reich 2012, Narwani *et al.* 2013, Kelly *et al.* 2014, Venail *et al.* 2014, Münkemüller *et al.* in press). Moreover, in order to positively influence ecosystem functioning, more phylogenetically diverse communities need to somehow maximize resource partitioning (i.e., niche complementarity) or to enhance facilitation, thus leading to greater resource use efficiency compared to less diverse communities. Similarly, to ensure ecosystem functioning in the face of changing conditions (i.e., to increase temporal or spatial stability) phylogenetically diverse communities may generate negative covariances in population dynamics by either increasing competitive interactions (Godoy *et al.* 2014) or by ensuring that species' responses to the environment are independent (Venail *et al.* 2013). Our analyses suggest that the phylogenetic relatedness of species, beyond its covariance with species richness, may not be a good predictor of ecosystem functioning (at least when this is measured as biomass production) with one possible explanation being the lack of phylogenetic signal in traits related to biomass production. This would suggest that, across the suites of species used in these experiments, functional complementarity between species did not increase with increasing phylogenetic distance between them.

More broadly, our result suggest that if standard diversity metrics based on species numbers (e.g., species richness) were to be replaced by alternative metrics based on genetic differentiation (e.g., phylogenetic diversity), caution would be needed when inferring ecosystem functioning because there may be functionally important trait differences among species that are not simply explained in full by phylogenetic relatedness (Kelly *et al.* 2014). While maximizing phylogenetic diversity (Vane-Wright *et al.* 1991, Faith 1992, 1994, Winter *et al.* 2013) might seem to be a promising way to maximize functional diversity and thus ecosystem functioning, management recommendations that suggest conservation of evolutionarily distinct species will lead to higher functional diversity and more stable communities are not well supported by the data explored in this study.

References

1. Balvanera, P., Pfisterer, A.B., Buchmann, N., He, J-S., Nakashizuka, T., Raffaelli, D. & Schmid, B. (2006) Quantifying the evidence for biodiversity effects on ecosystem functioning and services. *Ecology Letters*, **9**: 1146-1156.
2. Cadotte, M.W. (2013) Experimental evidence that evolutionary diverse assemblages result in higher productivity. *Proceedings of the National Academy of Sciences USA*, **110**, 8996-9000.
3. Cadotte, M.W. & Strauss, S.Y. (2011) Phylogenetic patterns of colonization and extinction in experimentally assembled plant communities. *Plos One*, **6**, e19363.
4. Cadotte, M.W., Cardinale, B.J. & Oakley, T.H. (2008) Evolutionary history and the effect of biodiversity on plant productivity. *Proceedings of the National Academy of Sciences USA*, **105**, 17012-17017.
5. Cadotte, M. W., Cavender-Bares, J., Tilman, D. & Oakley, T. H. (2009) Using phylogenetic, functional and trait diversity to understand patterns of plant community productivity. *Plos One*, **4**, e5695.
6. Cadotte, M. W., Dinnage, R. & Tilman. D. (2012) Phylogenetic diversity promotes ecosystem stability. *Ecology*, **93**, S223-S233.
7. Campbell, V., Murphy, G. & Romanuk, T. N. (2011) Experimental design and the outcome and interpretation of diversity-stability relations. *Oikos,* **120**, 399-408.
8. Cardinale, B.J., Srivastava, D.S., Duffy, J.E., Wright, J. P., Downing, A.L., Sankaran, M. &. Jouseau, C. (2006) Effects of biodiversity on the functioning of trophic groups and ecosystems. *Nature*, **443**, 989-992.
9. Cardinale, B. J., Matulich, K. L., Hooper, D. U., Byrnes, J. E., Duffy, E., Gamfeldt, L., Balvanera, P., O'Connor M. I. & Gonzalez, A. (2011) The functional role of producer diversity in ecosystems. *American Journal of Botany*, **98**, 1-21.
10. Cavender-Bares, J. & Reich, P.B. (2012) Shocks to the system: Community assembly of the oak savanna in a 40-year fire frequency experiment. *Ecology*, **93**, S52–S69.
11. Diaz, S. & Cabido, M. (2001) Vive la difference: plant functional diversity matters to ecosystem processes. *Trends in Ecology & Evolution*, **16**, 646-655.
12. Edgar, R.C. (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioninformatics*, **5**, 1-19.
13. Faith, D. P. (1992) Conservation evaluation and phylogenetic diversity. *Biological Conservation*, **61**, 1-10.
14. Faith, D. P. (1994) Genetic diversity and taxonomic priorities for conservation. *Biological Conservation*, **68**, 69-74.

15. Faith, D. P. (2002) Quantifying biodiversity: a phylogenetic perspective. *Conservation Biology,* **16**, 248-252.
16. Flynn, D.F.B., Mirotchnick, N., Jain, M., Palmer, M. I. & Naeem, S. (2011) Functional and phylogenetic diversity as predictors of biodiversity-ecosystem-function relationships. *Ecology*, **92**, 1573-1581.
17. Freckleton, R.P., Harvey, P.H., & Pagel, M. (2002) Phylogenetic analysis and comparative data: a test and review of evidence. *The American Naturalist*, **160**, 712-726.
18. Godoy, O., Kraft, N. J. B., & Levine, J. M. (2014) Phylogenetic relatedness and the determinants of competitive outcomes. *Ecology Letters*, **17**, 836-844.
19. Gross, K., Cardinale, B. J., Fox, J., Gonzalez, A., Loreau, M., Polley, H. W., van Ruijven, J. & Reich, P. (2014) Species richness and the temporal stability of biomass production: An analysis of recent biodiversity experiments. *The American Naturalist,* **In press**.
20. Hector, A., Hautier, Y., Saner, P., Wacker, L., Bagchi, R., Joshi, J., Scherer-Lorenzen, M., Spehn, E. M., Bazeley-White, E., Weilenmann, M. *et al.* (2010) General stabilizing effects of plant diversity on grassland productivity through population asynchrony and overyielding. *Ecology,* **91**, 2213-2220.
21. Helmus, M.R., Bland, T.J., Williams, C.K. & Ives, A.R. (2007) Phylogenetic measures of biodiversity*, The American Naturalist*, **169***,* E68-E83.
22. Hooper, D.U. *et al.* (2005) Effects of biodiversity on ecosystem functioning: a consensus of current knowledge. *Ecological Monographs,* **75**, 3-35.
23. Jiang, L. & Pu, Z. (2009) Effects of species diversity on temporal stability in single-trophic and multitrophic communities. *The American Naturalist,* **174**, 651-659.
24. Johnson, M.T.J & Stinchcombe, J.R. (2007) An emerging synthesis between community ecology and evolutionary biology. *Trends in Ecology and Evolution,* **22**, 250-257.
25. Jousset, A., Schmid, B., Scheu, S., Eisenhauer, N. (2011) Genotypic richness and dissimilarity opposingly affect ecosystem functioning. *Ecology Letters,* **14**, 537-545.
26. Kelly, S., Grenyer, R. & Scotland, R.W. (2014) Phylogenetic trees do not reliably predict feature diversity. *Diversity and Distributions*, **20**, 600-612.
27. Kembel, S.W., Cowan, P.D., Helmus, M.R., Cornwell, W.K., Morlon, H., Ackerly, D.D., Blomberg, S.P. & Webb, C.O. (2010) Picante: R tools for integrating phylogenies and ecology. *Bioinformatics,* **26**, 1463-1464.
28. Loreau, M. (2004) Does functional redundancy exist? *Oikos*, **104**, 606-611.
29. Loreau, M., Naeem, S., Inchausti, P., Bengtsson, J., Grime, J. P., Hector, A., Hooper, D.U., Huston, M.A., Raffaelli, D., Schmid, B., Tilman, D. & Wardle, D.A. (2001) Biodiversity

and ecosystem functioning: Current knowledge and future challenges. *Science*, **294**, 804-808.

30. Loreau, M., & de Mazancourt, C. (2008) Species synchrony and its drivers: Neutral and nonneutral community dynamics in fluctuating environments. *The American Naturalist,* **172**:E48-E66.
31. Loreau, M., & de Mazancourt, C. (2013) Biodiversity and ecosystem stability: a synthesis of underlying mechanisms. *Ecology Letters,* **16,** 106-115.
32. Losos, J.B. (2008) Phylogenetic niche conservatism, phylogenetic signal and the relationship between phylogenetic relatedness and ecological similarity among species. *Ecology Letters,* **11,** 995-1003.
33. Maherali, H., & Klironomos J.N. (2007) Influence of phylogeny on fungal community assembly and ecosystem functioning. *Science*, **316**, 1746-1748.
34. Munkemuller, T., Boucher, F., Thuiller, W. & Lavergne, S. (in press) Phylogenetic niche conservatism – common pitfalls and ways forward. *Functional Ecology*.
35. Narwani, A., Alexandrou, M.A., Oakley, T.H., Carroll, T. & Cardinale, B.J. (2013) Experimental evidence that evolutionary relatedness does not affect the ecological mechanisms of coexistence in freshwater green algae. *Ecology Letters*, **16**, 1373-1381.
36. Oakley, T.H., Alexandrou, M.A., Ngo, R., Pankey, M.S., Churchill, C.K., Chen, W. & Lopker, K.B. (2014) Osiris: accessible and reproducible phylogenetic and phylogenomic analyses within the Galaxy workflow management system. *BMC bioinformatics*, **15**: 230.
37. Paquette, A. & Messier, C. (2011) The effect of biodiversity on tree productivity: from temperate to boreal forests. *Global Ecology & Biogeography,* **20**, 170-180.
38. Petchey, O.L. & Gaston, K.J (2006) Functional diversity: back to the basics and looking forward. *Ecology Letters,* **9**, 741-758.
39. Prinzing, A., Durka, W., Klotz, S. & Brandl, R. (2001) The niche of higher plants: evidence for phylogenetic conservatism. *Proceedings of the Royal Society of London Series: B*, **268**, 2383-2398.
40. Pu, Z., Daya, P., Tan J. & Jiang, L. (2014) Phylogenetic diversity stabilizes community biomass. *Journal of Plant Ecology* (in press).
41. Reich, P.B., Tilman, D., Isbell, F., Mueller, K., Hobbie, S.E., Flynn, D.F.B. & Eisenhauer N. (2012) Impacts of biodiversity loss escalate through time as redundancy fades. *Science,* **336**, 589-592.
42. Schulze, E.D. & Mooney, H.A. (1993) Biodiversity and Ecosystem Function, Springer.

43. Srivastava, D. S., Cadotte, M. W., MacDonald, A. M., Marushia, R. G., Mirotchnick, N. (2012) Phylogenetic diversity and the functioning of ecosystems. *Ecology Letters,* **15**, 637-648.
44. Stamatakis, A. & Ott, M. (2008) Efficient computation of the phylogenetic likelihood function on multigene alignments and multi-core architectures. *Philos. Trans. Roy. Soc.,* **363**, 3977-3984.
45. Tilman, D, & Downing, J.A. (1994) Biodiversity and stability in grasslands. *Nature,* **367**, 363-365.
46. Tilman, D. (1999) The ecological consequences of changes in biodiversity: a search for general principles. *Ecology*, **80**, 1455-1474.
47. Tilman, D., Lehman, D. & Thompson, K. (1997) Plant diversity and ecosystem productivity: Theoretical considerations. *Proceedings of the National Academy of Sciences USA*, **94**,1857-1861.
48. Vane-Wright, R.I., Humphries, C.J. & Williams, P.H. (1991) What to protect?-systematics amd the agony of choice. *Biological Conservation*, **55**, 235-254.
49. Venail, P. & Vives, M.J. (2013) Phylogenetic distance and species richness interactively affect the productivity of bacterial communities. *Ecology*, **94**, 2529-2536.
50. Venail P., Alexandrou, M.A., Oakley, T.H. & Cardinale, B. J. (2013) Shared ancestry influences community stability by altering competitive interactions: evidence from an experiment using freshwater green algae. *Proceedings of the Royal Society: B,* **280**, 20131548.
51. Venail P., Narwani, A, Fritschie, K., Alexandrou, M.A., Oakley, T.H. & Cardinale, B. J. (2014) The influence of phylogenetic relatedness on species interactions among freshwater green algae. *Journal of Ecology*, **102**, 1288-1299.
52. Verdu, M., Gomez-Aparicio, L. & Valiente-Banuet, A. (2012) Phylogenetic relatedness as a tool in restoration ecology: a meta-analysis. *Proceedings of the Royal Society of London Series B: Biological Sciences*, **279**, 1761-1767.
53. Violle, C., Nemergut, D.R., Pu, Z. & Jiang, L. (2011) Phylogenetic limiting similarity and competitive exclusion. *Ecology Letters*, **14**, 782-787.
54. Webb, C. O., Ackerly, D.D., McPeek, M.A. & Donoghue, M. J. (2002) Phylogenies and community ecology. *Annual Review of Ecology and Systematics,* **33**, 475-505.
55. Webb, C.O, Ackerly, D.D. & Kembel, S.W. (2008) Phylocom, software for the analysis of phylogenetic community structure and trait evolution. *Bioinformatics*, **24**, 2098-2010.

56. Wiens, J.J., Ackerly, D.D., Allen, A.P., Anacker, B.L., Buckley, L.B., Cornell, H.V., Damschen, E.I., Davies, T.J., Grytnes, J.A., Harrison, S.P., Hawkins, B.A., Holt, R.D., McCain, C.M. & Stephens, P.R. (2010) Niche conservatism as an emerging principle in ecology and conservation biology. *Ecology Letters*, **13,** 1310-1324.
57. Winter, M., Devictor, V. & Schweiger, O. (2013) Phylogenetic diversity and nature conservation: where are we? *Trends in Ecology and Evolution*, **28**, 199-204.
58. Zanne, A.E., Tank, D.C., Cornwell, W.K., Eastman, J.M., Smith, S.A. *et al.* (2014) Three keys to the radiation of angiosperms into freezing environments. *Nature*, **506**, 89-92.

Figure legends

Figure 1: The relationship between species richness (SR) and phylogenetic distance (PD) or (MPD) for the grassland studies used in this data synthesis. a) For the full dataset with 824 data points (plots) from 16 independent studies (experiments), PD and SR are very highly correlated (r = 0.90; plots with SR = 60 are not shown in the graph). This leads to problems of multi-collinearity that make it difficult to separate the effects of PD on community stability from those of SR in any multivariate analyses. Because of this, we performed four complementary types of analyses. For Type 1 analyses, we analyzed the impacts of PD on stability within levels of SR, (i.e., to analyze effects of PD whilst holding SR constant). In the Type 2 analysis, we did the opposite and identified 1417 contrasts where plots within a study had very similar values of PD, but differed in SR. While it was not possible to hold PD statistically 'constant', these contrasts offered the closest approximation. b) In the Type 3 analyses, we used five of the 16 studies where PD and SR had the lowest correlations ($r \leq 0.80$; studies 5, 6, 9, 13 and 14 from Table 2), which allowed us to perform more traditional multivariate analyses on this subset of data while accounting for the covariance among explanatory variables. c) In the Type 4, we used an alternative metric of phylogenetic diversity (mean pairwise distance, MPD), which is independent of SR (r = -0.013, plots with SR = 60 are not shown), allowing us to include the full dataset (824 plots). See text for further explanation.

Figure 2: The effect of phylogenetic diversity (PD) on stability and its different components, whilst holding species richness (SR) constant. a) Coefficients of correlation relating phylogenetic diversity (PD) to stability (diamonds), as well as the two components

contributing to stability: average biomass (circles) and standard deviation (squares); and to sum of species variances (sum.var., triangles) and synchrony (crosses). Each data point represents the correlation for one individual study. Results are presented for each species richness level (SR, vertical axis) so that conclusions can be drawn about the influence of PD, without confounding changes in SR. Filled data points and plus signs represent studies where correlation coefficient values were significant ($p < 0{,}05$). b) Overall weighted and normalized average coefficients of correlation (Weighted Z_r, see text for details) between phylogenetic diversity (PD) and each component of temporal community stability including all the species richness levels. The sign of overall Z_r represents the overall shape of the relationship between PD and each component (either positive, neutral or negative). Horizontal bars represent 95% confidence intervals. Filled symbols represent overall Z_r values that are significantly different from zero.

Figure 3: The effect of species richness (SR) on stability and its different components, while holding phylogenetic diversity (PD) constant. For the 1417 contrasts used in type 2 analysis, we further calculated the log ratios for community stability, average community biomass, standard deviation of biomass (S.D.) and sum of variances of individual species' biomass (sum.var.) in plots through time for higher vs. lower species richness. For a clearer interpretation of the data the x-axis is presented in a non-log scale. Values higher than one for stability and average biomass indicate that more speciose communities are more stable and produce more biomass than less speciose ones. A value lower than one for S.D. indicates that the biomass of more speciose communities has lower temporal variation than the biomass of less speciose communities. Data points are the mean and 95% confidence

intervals. Note than synchrony is not represented because it is not possible to estimate log ratios on negative values.

Figure 4: Results of a structural equations model (SEM) showing the joint effects of species richness (SR) and phylogenetic diversity (PD) on stability, The SEM that used data from 5 studies (*n* = 222 data points, χ^2 = 1.19, *d.f.* = 2, *P* = 0.55) where the correlation between SR and PD was ≤ 0.8. The reduced correlation of the sub dataset allowed us to explicitly model the covariance between SR and PD and then examine the partial regression coefficients (showed as values above the paths) relating both explanatory factors to community biomass (biomass) and the S.D. of biomass through time (SD). Lines with single headed arrows represent causal pathways whereas lines with double headed arrows represent co-varying variables. Community biomass and the S.D. of biomass through time are the two components of stability. Significance is indicated by asterisks: * for $p < 0.05$, ** for $p < 0.01$, *ns* for non-significant. See also Table S3 for more details.