ECOLOGY LETTERS

Ecology Letters, (2018) **21**: 1604–1619 doi: 10.1111/ele.13145

IDEA AND PERSPECTIVE

Ecosystem context illuminates conflicting roles of plant diversity in carbon storage

E. Carol Adair,[1] David U. Hooper,[2*] Alain Paquette[3] and Bruce A. Hungate[4]

Abstract

Plant diversity can increase biomass production in plot-scale studies, but applying these results to ecosystem carbon (C) storage at larger spatial and temporal scales remains problematic. Other ecosystem controls interact with diversity and plant production, and may influence soil pools differently from plant pools. We integrated diversity with the state-factor framework, which identifies key controls, or 'state factors', over ecosystem properties and services such as C storage. We used this framework to assess the effects of diversity, plant traits and state factors (climate, topography, time) on live tree, standing dead, organic horizon and total C in Québec forests. Four patterns emerged: (1) while state factors were usually the most important model predictors, models with both state and biotic factors (mean plant traits and diversity) better predicted C pools; (2) mean plant traits were better predictors than diversity; (3) diversity increased live tree C but reduced organic horizon C; (4) different C pools responded to different traits and diversity metrics. These results suggest that, where ecosystem properties result from multiple processes, no simple relationship may exist with any one organismal factor. Integrating biodiversity into ecosystem ecology and assessing both traits and diversity improves our mechanistic understanding of biotic effects on ecosystems.

Keywords

biodiversity, community-weighted mean, ecosystem functioning, ecosystem services, functional diversity, functional traits, interactive factors, phylogenetic diversity, species richness, state factors.

Ecology Letters (2018) **21**: 1604–1619

INTRODUCTION

Decades of research have established that biodiversity influences ecosystem properties (Schulze & Mooney 1993; Chapin *et al*. 2000; Tilman *et al*. 2014) and services[1] (Naeem *et al*. 2009; Cardinale *et al*. 2012; Isbell *et al*. 2017b). This research generated its own subdiscipline – biodiversity and ecosystem functioning (BEF). To isolate the influence of diversity from co-varying factors, much BEF research has occurred at small scales in communities where biodiversity is manipulated experimentally (Hooper *et al*. 2005; Cardinale *et al*. 2012). While this work provides strong evidence for isolated effects of altered species richness, genetic diversity and functional diversity, interactions with other drivers may confound extrapolating these effects to landscape scales (Srivastava & Vellend 2005; Wardle *et al*. 2011; Wardle 2016). A more recent approach explores diversity effects across broad geographic gradients, controlling for environmental variation through multiple regression or structural equation modeling. To date, few studies (but see, Díaz *et al*. 2007; Poorter *et al*. 2017) have integrated diversity with all the ecosystem drivers, known as 'state factors' in ecosystem ecology (Vitousek 2004; Chapin *et al*. 2011), typically used to model ecosystem processes at large scales. However, this progress suggests BEF is ripe for more complete integration with ecosystem ecology (Chapin *et al*. 2000).

Integrating plant diversity with other ecosystem drivers will help merge biodiversity research with ecosystem ecology and clarify diversity effects at large spatial scales. Organisms are an integral component of ecosystem ecology's state-factor framework (Jenny 1980; Vitousek 2004; Chapin *et al*. 2011), but it has focused on dominant species or mean plant traits (Grime 1998), rather than diversity (but see Chapin *et al*. 2000). In the state-factor framework, climate, organisms (regional species pool), topography, geological substrate and

[1]Note: following Hooper *et al.* (2005), here we use ecosystem properties to refer to pools and fluxes of energy and matter, equivalent to 'supporting services' in the Millennium Ecosystem Assessment (MEA, 2005). We use the term 'ecosystem services' to refer specifically to benefits to human society derived from ecosystem properties; equivalent to provisioning, regulating and cultural services in the MEA (2005).

[1]*Rubenstein School of Environment and Natural Resources, University of Vermont, Burlington, VT 05405, USA*
[2]*Department of Biology, Western Washington University, Bellingham, WA 98225-9160, USA*
[3]*Centre for Forest Research, Département des sciences biologiques, Université du Québec à Montréal, CP 8888, Succursale Centre-ville, Montréal, QC, Canada H3C 3P8*
[4]*Center for Ecosystem Science and Society and Department of Biological Sciences, Northern Arizona University, Flagstaff, AZ 86011, USA*
**Correspondence: E-mail: david.hooper@wwu.edu*

time since disturbance are independent state factors driving all ecosystem properties (Fig. 1). These distal drivers interact to influence proximal drivers ('interactive factors'): microclimate, resource availability, organismal functional traits and disturbance regimes (Chapin *et al.* 2011). Although the relative strengths of plant functional traits and diversity as biotic drivers affecting ecosystem properties have been compared and contrasted (Grime 1998; Winfree *et al.* 2015; Finerty *et al.* 2016), an alternative approach integrates both, recognising the different axes of influence by which organisms affect ecosystem properties (see also, Díaz *et al.* 2007). In that spirit, we suggest incorporating diversity into the state-factor paradigm by broadening the interactive 'organismal functional traits' factor to 'interactive biotic factors', within which both mean traits and diversity warrant consideration (Fig. 1 and Fig. S1).

Although integrating diversity with environmental drivers has progressed in the past several years, the development and use of a cohesive theoretical framework is needed. The major ecosystem models (e.g. CENTURY, TEM, BIOME-BGC) currently do not incorporate diversity as a driver of carbon and nitrogen cycling, perhaps because of the lack of information on how diversity effects and interactions play out at landscape scales. Recent studies assessing diversity effects in combination with different environmental factors have found significant influences at broad geographic scales, including on algal biomass (Zimmerman & Cardinale 2014), primary production (Paquette & Messier 2011; Cavanaugh *et al.* 2014; Poorter *et al.* 2017), or multifunctionality in ecosystem properties (Maestre *et al.* 2012) or services (Gamfeldt *et al.* 2013; Liang *et al.* 2016; Duffy *et al.* 2017; Oehri *et al.* 2017). One study even structured some analyses on the state-factor framework, but their interpretations instead focused on diversity effects (Jing *et al.* 2015). Most, however, have assessed diversity as the only biotic predictor, ignoring concurrent impacts of important plant traits (but see Poorter *et al.* 2017). Fewer still also structured their analyses to reflect the key relationships among state factors, interactive factors and the processes of interest. A key component of the state-factor framework is whether ecosystem drivers are functionally outside (state factors) or inside (interactive factors) the realm of internal ecosystem feedbacks (Chapin *et al.* 2011). This distinction helps direct the structure of measurements and analyses assessing effects of different ecosystem drivers (Fig. 1 and Fig. S1).

What is the strength of plant traits and diversity, as interactive factors, on landscape-scale ecosystem properties and services after accounting for state factors? Mean functional traits and organismal diversity both have demonstrable mechanisms that should influence ecosystem properties at the landscape

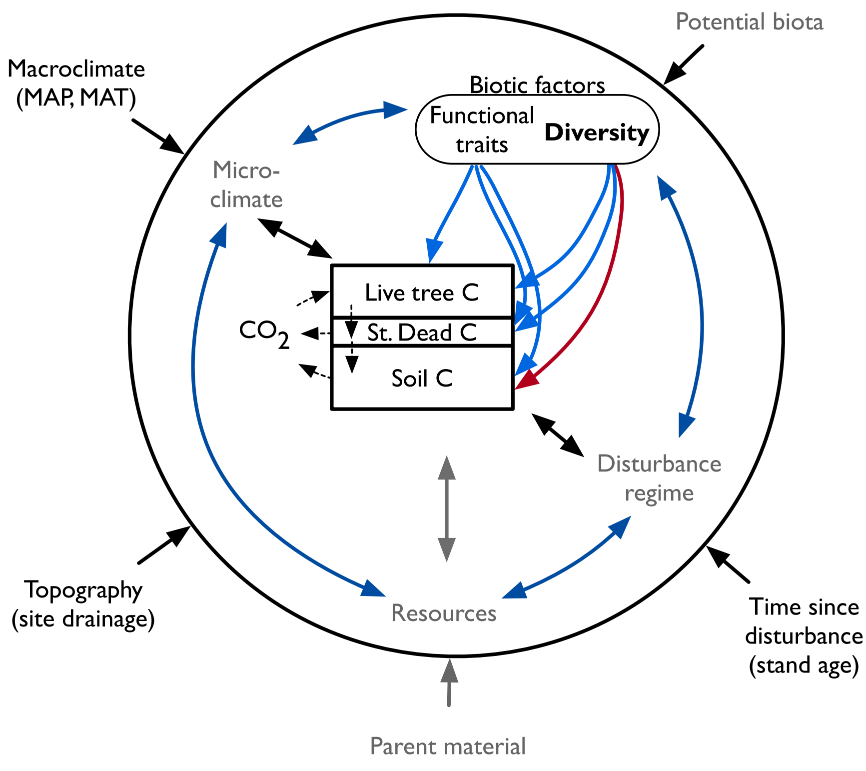

Fig. 1 Relationship among state factors, interactive factors, and ecosystem carbon pools. Diversity, in bold, and the arrows emanating from it, show our proposed modification of the state factor approach. The circle indicates whether ecosystem drivers are functionally outside (state factors; outside the circle) or inside (interactive factors; inside the circle) the realm of internal ecosystem feedbacks. Grey text indicates lack of data in this study to include these factors. We aimed to determine the impact of biotic factors on C pools, so we did not include feedbacks from C pools to biotic factors. Variables used for different state factors are in parentheses (MAT = mean annual temperature, MAP = mean annual precipitation). Dashed arrows indicate fluxes between different ecosystem C pools and atmospheric CO_2 to emphasise that these components of total ecosystem C storage respond differently to the different ecosystem state and interactive factors. Dark blue arrows indicate that interactive factors result from interactions among state factors. Light blue arrows are positive relationships, and red arrows are negative relationships of biotic factors on ecosystem C pools. Adapted from Chapin *et al.* 2011.

scale. For example, mean traits along the leaf economic spectrum affect productivity (Díaz *et al.* 1999; Wright *et al.* 2004; Reich 2014), litter quality, decomposition and plant-soil productivity feedbacks (van der Putten *et al.* 2013; Hobbie 2015). Trait diversity can increase resource capture and productivity via complementarity or facilitation (Vandermeer *et al.* 2002; Cardinale *et al.* 2011; Tilman *et al.* 2014). Diversity may also stabilise services by incorporating environmentally tolerant species or allowing asynchronous responses to environmental fluctuations (Walker *et al.* 1999; Yachi & Loreau 1999; Balvanera *et al.* 2006; Isbell *et al.* 2009; Cardinale *et al.* 2013; Wang & Loreau 2016). We argue that studies should incorporate both mean traits and diversity as biotic drivers to better understand when and where they have strong impacts.

Many ecosystem services are combinations of properties (Cardinale *et al.* 2012) that may respond independently to altered diversity. Thus, understanding mechanisms will require considering the component properties rather than the single aggregated service. For example, total ecosystem C storage reflects imbalances between C inputs from primary production and losses from decomposition (Fig. 1). Decomposition is typically more sensitive to temperature and high soil moisture than production, so that soil C increases at colder, higher latitudes (Díaz *et al.* 2009a) and in waterlogged soils (Schuur *et al.* 2001; Mack *et al.* 2008; Grosse *et al.* 2011). At the same time, production and decomposition are maximised by high nitrogen, low lignin leaves (Hobbie 1992; Wright *et al.* 2004; Reich 2014; Hobbie 2015; Cornelissen *et al.* 1999), and shifts in these traits may have little net effect on soil C.

Production appears to be more sensitive to changes in diversity than decomposition (Srivastava *et al.* 2009; Hooper *et al.* 2012; Jewell *et al.* 2017), but whether this translates into greater ecosystem C storage remains unclear. Generally, diversity increases plant production and plant C (e.g. Cardinale *et al.* 2011; Paquette & Messier 2011; Gamfeldt *et al.* 2013; Ruiz-Benito *et al.* 2014; Liang *et al.* 2016; Duffy *et al.* 2017; Poorter *et al.* 2017), but diversity effects on decomposition and soil C show mixed effects (Hattenschwiler *et al.* 2005; Fornara & Tilman 2008; Srivastava *et al.* 2009; Reid *et al.* 2012; Gamfeldt *et al.* 2013; Hungate *et al.* 2017). Greater plant and detritivore diversity can accelerate decomposition (Nielsen *et al.* 2011; Handa *et al.* 2014), which should decrease C storage. Thus, C storage is an aggregate property of production, decomposition and disturbance – processes that respond independently to environmental and biotic changes, including changes in diversity. The state-factor framework offers structure for evaluating these interacting influences simultaneously.

We investigated diversity effects on ecosystem C storage using the state-factor framework with climate, topography and stand age, and the interactive factor of mean organismal traits. We selected variables representative of these state factors based on well-supported ecosystem science (e.g. Chapin *et al.* 2011) and recent studies in our focal ecosystem (Paquette & Messier 2011). Our goal was to illustrate how integrating biodiversity with other ecosystem controls leads to mechanistic insights that may be lost when this framework is not used, when key drivers are not accounted for, or when emphasis is solely on whether biodiversity is significant or not. We used Québec Forest Survey data to evaluate these effects on C in live trees, standing dead trees, the soil organic horizon, and their sum, total ecosystem C. We asked: 1. How does geographic variation in interactive biotic factors (plant traits and diversity) affect C storage when accounting for variation in state factors across broad geographic regions? What are the relative strengths of mean traits and diversity effects? We used structural equation modeling (SEM) to reflect direct effects of state factors (those outside the realm of internal ecosystem feedbacks) on C storage, as well as indirect effects via interactive biotic factors (mean plant traits and diversity), consistent with the state-factor framework (Fig. 1 and Fig. S1). This structure also allowed us to compare relative effects of mean traits and diversity. Debate continues about which functional traits and diversity components (phylogenetic diversity, functional diversity, or species richness) have greatest effects on ecosystem properties (e.g. Cadotte 2015; Cardinale *et al.* 2015; Venail *et al.* 2015; Naeem *et al.* 2016). We therefore evaluated a variety of traits and diversity metrics to better understand which provide the most explanatory power when combined with state factors in this framework. 2. Is there evidence that different biotic mechanisms control various C pools? Ecosystem theory and evidence suggest this likelihood (Chapin *et al.* 2011), with production traits driving live tree C and decomposition traits driving organic horizon C. We hypothesised that explaining variation in landscape C storage requires mechanistic linkages among C pools and different functional traits or diversity metrics. Because we expected that state and interactive biotic factors would affect C pools with varying strengths (Díaz *et al.* 2009b), we assessed whether models of total C would perform as well as those investigating individual pools: is it sufficient to lump contributing properties into a single ecosystem service, or should they be separate? Although we focused on ecosystem C, we anticipate this approach will clarify organismal effects on other ecosystem services as well.

METHODS

Datasets and carbon pools

We used three state factors – climate, topography and time since disturbance – plus mean plant traits and diversity as facets of the interactive biotic factor (Table 1 and Table S1; Fig. 1 and Fig. S1; Appendix 2) to assess ecosystem controls on C pools in temperate and boreal forests of Québec (Fig. 2; Paquette & Messier 2011): live tree, standing dead, soil organic horizon, and total C (their sum). Data at each site were collected in 400 m^2 circular plots as part of the Québec Forest Survey (MRNFQ 2006; summarised by Paquette & Messier 2011, Appendix 1; Table S1). Sites were sampled every decade; we selected the most recent measurement for sites with a recorded age. Data were insufficient to estimate coarse woody debris C and did not include mineral soil C or measurements of nutrient availability.

We expected effects of topography (drainage) on C pools would overwhelm other factors in fast (dry) and slow (wet) draining sites, so we performed analyses on full (n = 2624) and moderately drained datasets (drainage rating 2–4,

Table 1 Biotic measures of functional traits and diversity. Definitions and units of functional traits are given in Table S2

Functional traits:			
Community Weighted Mean		Diversity metrics	
GR	Growth rate	SR	Tree species richness
LL	Leaf longevity	PSV	Phylogenetic species variability
LMA	Leaf mass per area	FDis	Functional Dispersion for
LS	Leaf size	*Single traits*	Drought tolerance
MaxH	Average maximum height		Leaf mass per area
N	N mass per leaf mass		Average maximum height
SM	Seed Mass		N mass per leaf mass
TolD	Drought tolerance		Seed Mass
TolS	Shade tolerance		Shade tolerance
TolW	Waterlogging tolerance		Waterlogging tolerance
Veg	Vegetative reproduction		Wood density
WDR	Wood decay resistance	*Three traits*	
WD	Wood density	3a	Wood density, seed mass, max height
		3b	Shade tol, seed mass, max height
		3c	Shade tol, drought tol, max height
		Five traits	
		5	Max height, shade tol, drought tol, waterlogging tol, N mass

n = 2323), where we expected biotic factors would play a stronger role in C accumulation. Temperate forests were restricted to a relatively limited climatic range that may not represent the breadth of environmental conditions in this biome, challenging our ability to test for distinct relationships in temperate vs. boreal systems; we therefore grouped these forests together for analyses.

At each site, all trees > 9.1 cm DBH were measured, identified, and counted. Live tree biomass was the sum of stem wood, bark, branches, leaves and roots using species-specific allometric equations for aboveground pools (Lambert *et al.* 2005) and coarser deciduous and conifer allometries for roots (Li *et al.* 2003). Dead tree biomass was the sum of dead wood, bark, branches and roots. Live and standing dead biomass were converted to carbon (C) content using C = 0.5 * biomass. Although the C content of conifer and hardwood temperate trees ranges from 43 to 55% C (Lamlom & Savidge 2003; Thomas & Martin 2012), we chose 50% as a mid-point value because we did not have data on all species encountered. We estimated organic horizon C using site measurements of soil group and organic layer depth, combined with soil great group bulk density and C content (Shaw *et al.* 2008): Organic horizon C = bulk density * organic layer depth * %C. Total C was the sum of live, standing dead and organic horizon C. All values are Mg C ha^{-1}.

Plant traits and diversity metrics

Site-level species richness was assessed for each 400 m^2 site. We assembled trait data for the species in the dataset from published sources (Tables 1, S1, S2) (Paquette & Messier 2011), focusing on traits we hypothesised would affect C pool sizes either via trait means or diversity. To represent the traditional ecosystem approach, we calculated community-weighted means (CWM) for each functional trait at each site, per Lavorel *et al.* (2008), weighted by species basal area. For ordered traits, CWM calculations return the value of the dominant class (Tables S1, S2). We investigated three types of diversity metrics: species richness, functional diversity and phylogenetic diversity. We used species richness (SR) to represent the dominant BEF approach. We computed functional dispersion (FDis), which provides the average multivariate distance of individual species from the centroid of all species in functional trait space, weighted by basal area (Laliberté & Legendre 2010; Paquette & Messier 2011) (Table 1). We calculated twelve FDis metrics, using either single or multiple traits, for traits we hypothesised could influence C pools via diversity (Tables 1 and S1). While many studies using FDis include a wide variety of traits, we were concerned that, in such an approach, variation in less relevant traits could obscure effects of more relevant traits (Bernhardt-Römermann *et al.* 2008). Our limited number of multiple trait FDis metrics were intended to capture a variety of explicit mechanisms by which species might contribute to complementary resource use or tolerance to abiotic conditions (Table 1). We chose FDis as a measure of functional diversity because it computes for plots with only two species (FDis = 0 when SR = 1) and is not strongly correlated with SR for theoretical communities (Laliberté & Legendre 2010). In practice, SR and FDis metrics were moderately correlated across our datasets (Pearson's r = 0.39–0.70), but were never used together to predict C pools. For phylogenetic diversity, we calculated Phylogenetic Species Variability (PSV), using a molecular phylogeny of our species based on chloroplast genes (Paquette & Messier 2011). PSV quantifies how phylogenetic relatedness decreases variance of a hypothetical trait shared by species in a community and is mathematically independent of SR (Helmus *et al.* 2007). Although PSV was moderately correlated with SR across the dataset (r = 0.51), they were never used together to predict C pools. We computed CWM and FDis using the 'FD' package (Laliberté & Shipley 2011) and PSV using the 'picante' package for R (Kembel *et al.* 2010; R Core Team 2011).

All metrics were calculated using *a priori* species traits rather than plot-based measurements. While this certainly missed some within-species variability resulting from genetic effects and phenotypic plasticity, it is a common approach in

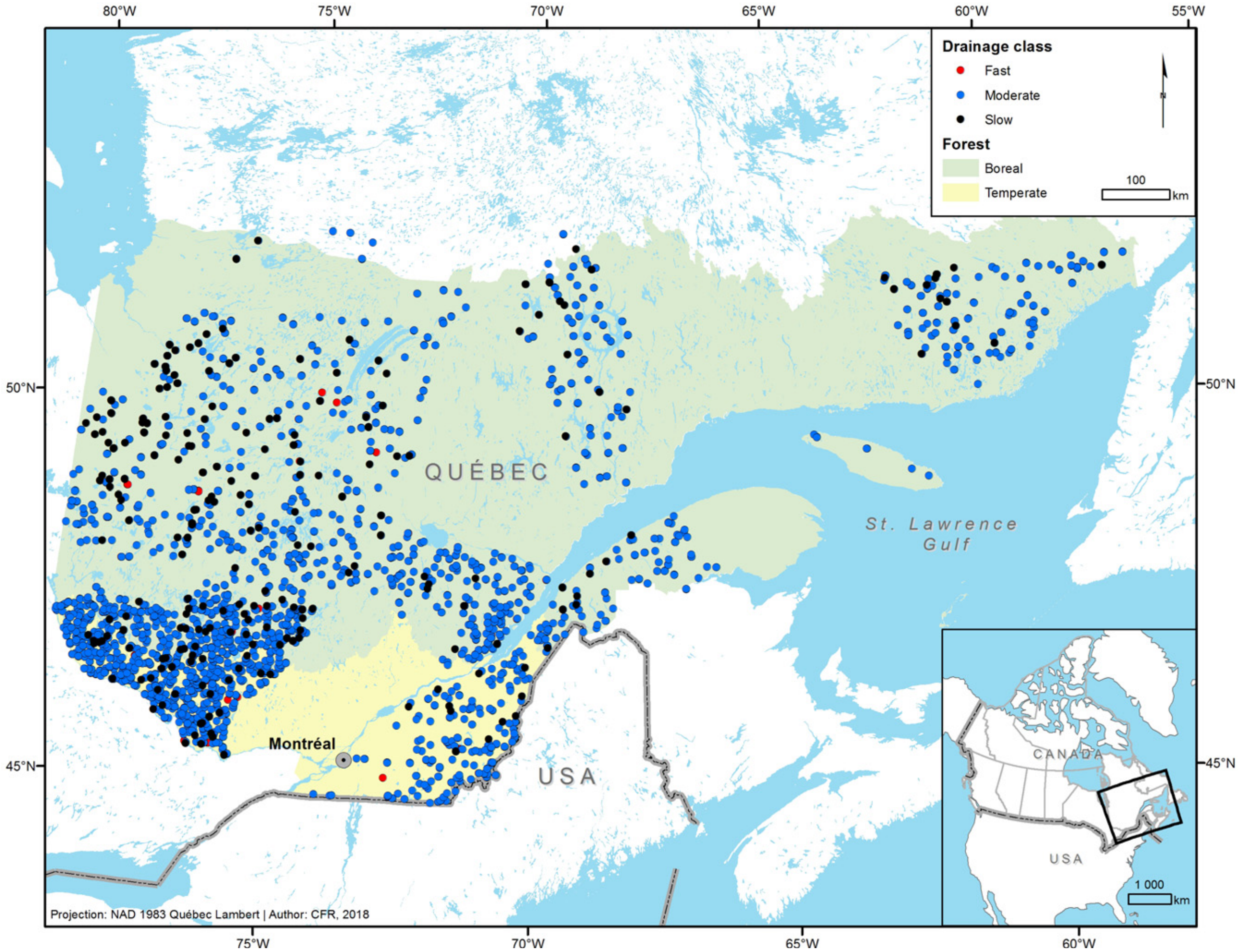

Fig. 2 Distribution of sites across temperate and boreal forest biomes of Québec. Full dataset: n = 2624; moderately-drained dataset: n = 2323; temperate forests: n = 948; boreal forests: n = 1676. Fast draining sites are drainage classes 0–1.5, moderate are 2–4, and slow are 4.5–6. Geographic gaps in coverage likely resulted from using only sites with measurements of stand age. However, distribution of all datasets cover the E-W and N-S breadth of both biomes in Québec.

broad-scale geographic studies, was essential given our large number of sites, and avoided circular logic for some functional traits. For example, maximum tree height was an *a priori* species trait taken from the literature and was not also used to calculate live tree C in individual plots. Thus, correlation between maximum height and plot-based measurements such as basal area was very low (r = 0.05).

Statistical analyses

Strategy and conceptual framework

Our primary goal was to build on the wealth of the knowledge from ecosystem and community ecology about drivers of C storage and biotic composition to develop ecologically relevant SEMs, based on the state-factor framework (Fig. 1), that would evaluate the role of both plant functional traits and diversity at landscape extents. To this end, we developed full and simplified meta-models (Grace *et al.* 2016) to guide our analysis (Fig. S1). However, given the large range of theoretically and empirically supported potential relationships among state and interactive factors and of potential relationships between these factors and different C pools, we first aimed to find the best descriptors for each state and interactive factor using multi-model inference and *a priori* ecological knowledge. The goal for this step was to minimise the likelihood of dismissing conceptually important explanatory variables solely because we had not used an appropriate statistical relationship or had arbitrarily used a particular metric that strongly covaried with another metric. Based on our simplified meta-model (Fig. S1), we then selected variables to build SEMs. We describe these steps in more detail below and in Appendix 2.

Multi-model inference: state and biotic interactive factor selection

We first identified the best variables or function shape for effects on different C pools for each state factor: climate, topography and time since disturbance. Because previous studies found strong correlations among various temperature and precipitation predictors at these sites, we used the simplest estimators: mean annual temperature (MAT) and mean annual precipitation (MAP) (Paquette & Messier 2011). Based on those results and abundant previous research on ecosystem C controls (e.g. Chapin *et al*. 2011; Taylor *et al*. 2017), we expected MAT, MAP, and their interaction would affect all pools, so we included them in all state factor models. Plot coordinates and elevation were used to compute MAT and MAP using interpolation of 30-year normals from all available weather stations (as in Paquette & Messier 2011).

Drainage represented the topography state factor, and ranged from 0 to 6 in 0.5 increments: 0-1.5 are rapidly draining, dry soils; 2–4 are moderately draining soils; 4.5–6 are slowly draining, wet soils. We tested for linear, exponential, or unimodal relationships between C pools and drainage. Stand age, representing time since disturbance, was estimated by coring from the five trees with the largest DBH at each site (MRNFQ 2006). We tested for linear, exponential, unimodal or lognormal (saturating) relationships between C pools and age. We used multi-model inference (MMI), including model averaged parameters, sums of weights (SW; Burnham & Anderson 2002), and standardised effect sizes (Galipaud *et al*. 2014; Galipaud *et al*. 2017; Cade 2015), to select the best relationships between each variable and C pool combination (Table 2; see Appendix 2 and Tables S3 and S4 for detailed methods and results).

We then selected variables for the interactive biotic factors using MMI to identify the trait mean and diversity metric that added the most explanatory power to the selected state factor models for each C pool. We did so by including only a single trait mean or diversity metric at a time, or one trait mean plus a diversity metric, in addition to the state factors. Selection of biotic factors was almost fully balanced across the candidate models; correcting for potential bias caused by the slight imbalance did not alter results (Appendix 2; Fig. S2). We chose this approach, rather than examining the effects of biotic factors on C pools in a fully factorial model set with abiotic factors, because the latter would allow models consisting of only biotic factors - not ecologically realistic given known drivers of C pools (Chapin *et al*. 2011; Liang *et al*. 2016; Oehri *et al*. 2017). Even so, we repeated the above analysis including models that contained only biotic factors (one trait, one diversity, or one trait + one diversity), but selected metrics were identical (Appendix 2). For live tree C, MMI clearly selected potential maximum height and PSV as the best predictors for CWM and diversity metrics in both datasets (Table 2, Fig. S3). Leaf mass per area was the clear selection for the CWM variable for organic horizon C. Diversity of LMA was the best predictor of organic horizon C, though more strongly in the moderately drained than full dataset. MaxH and FDis of wood density were the strongest biotic predictors for total C, though with less clear predominance than in the individual C pools (Table 2, Fig. S3). See Appendix 2 for detailed methods and Tables S5 and S6 for detailed results. Before constructing the SEMs, we used MMI to identify the shape of the relationships among state factors and the selected trait mean and diversity metrics for each C pool in each dataset, using the same process described above for the selection of state and interactive factors (Appendix 2; Tables S7-S8).

Table 2 Selected state factor relationships for each dataset and C pool. Full = all sites (n = 2624). Mod. Drain = Moderately drained sites, drainage classes 2–4 (n = 2323)

		State factor		Interactive factors	
Dataset	C pool	Topography	Time	Trait means	Diversity
Full	LTC	Unimodal Drain	Unimodal Age	MaxH	PSV
	SDC	exp (Drain)	Unimodal Age	WDR	FDis.TolS
	OHC	exp (Drain)	Age	LMA	FDis.LMA
	TC	exp (Drain)	Unimodal Age	MaxH	FDis.WD
Mod. Drain.	LTC	Unimodal Drain	Unimodal Age	MaxH	PSV
	SDC	Drain	Unimodal Age	WDR	FDis.TolS
	OHC	Unimodal Drain	Age	LMA	FDis.LMA
	TC	Unimodal Drain	Unimodal Age	MaxH	FDis.WD

LTC = Live Tree Carbon. SDC = Standing Dead Carbon. OHC = Organic Horizon Carbon. TC = Total Carbon. Trait mean and diversity metrics are as defined in Tables 1 and S1.

Structural equation model

For each C pool within each dataset, we constructed SEMs based on our simplified meta-model (Fig. S1, Appendix 2) using the selected state and interactive biotic factors for that pool (package 'lavaan' in R) (Rosseel 2012). Each SEM accounted for the direct state and interactive biotic factor effects on a given C pool and indirect state factor effects via diversity and mean plant traits. Each also allowed for potential correlations among state factors (not shown) and between mean trait and diversity metrics. Because the impact of diversity on ecosystem properties often saturates at high diversities, we used AIC_c to compare SEMs with untransformed and ln-transformed diversity. If there was no best model, the simpler linear model was selected. To understand the predictive impact of including plant traits and diversity with state factors, we compared the C pool R^2 from the complete SEM to the R^2s from an SEM without plant traits (state and diversity factors only) and one with only state factors.

RESULTS

Pool size distributions

Across all sites, the distribution of total C was skewed. On average, sites contained 150 Mg C ha^{-1} (median = 130 Mg C ha^{-1}), but the distribution tail reached 766 Mg C ha^{-1} (Fig. 3, Table S2). The tail was largely driven by organic horizon C (OHC), which increased exponentially with drainage (Fig. 3c); OHC as a percentage of total C increased with drainage from fast (< 2, 15–30%) to slow (>4, 70–90%; Fig. 3g). Live tree C, as a percent of total C, followed the opposite pattern (Fig. 3e). While standing dead C reached 25–35% of total C in some sites, in most it was <5% (median = 0.73%; Fig. 3b and f).

Structural equation models: integrating state and interactive biotic factors

For all C pools, including biotic interactive factors with state factors in the SEMs better predicted forest C pools,

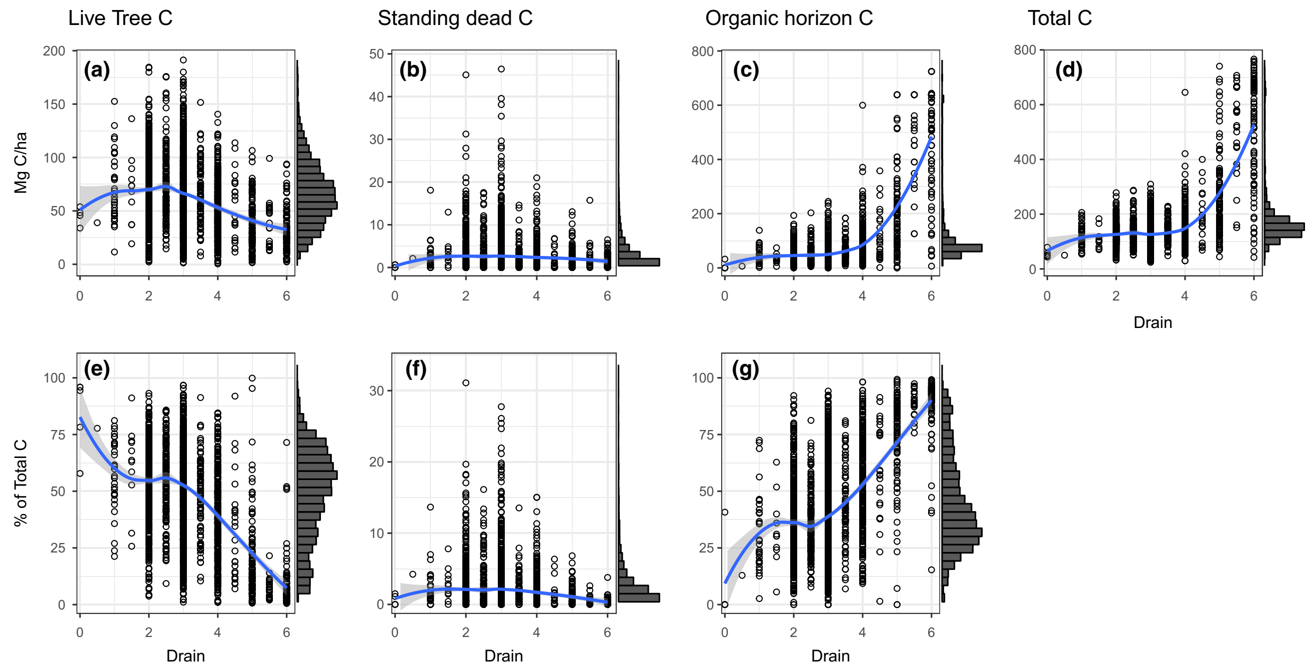

Fig. 3 Carbon pool sizes (a–d) and percent of total C (e–g) for live tree C (a and e), standing dead C (b and f), organic horizon C (c and g), and total C (d) for the full dataset. N = 2624 sites, with 149 sites poorly drained (D>4). Lines are loess smoothers with 95% confidence intervals (shaded areas).

particularly for live tree C (Table 3). In most models, abiotic state factors were the most important factors with the largest impacts on C pools (Table 4). The exception was for live tree C, where interactive biotic variables had equally large impacts compared to state factors (Table 4; Fig. 4–6). For all pools except organic horizon and total C in the full dataset, diversity effects did not saturate; the best SEMs used untransformed diversity (Figs 4–6; Tables 3, S9).

Temperature and age had the greatest effects of individual state factors on live tree C, with similar effects in both datasets (Fig. 4). Across both datasets, the strongest individual direct effect on live tree C was from a biotic factor, the CWM of maximum potential tree height (MaxH; Fig. 4; Table 4). However, variation in state factors drove up to 40% of the variation in MaxH (Fig. 4), indicating substantial species turnover responding to environmental conditions. After accounting for indirect and direct state factor effects, the effect of MaxH on live tree C was comparable to effects of temperature and age. Increasing temperature and MaxH from their minimum to maximum value across each dataset increased live tree C by 80 Mg C ha^{-1} in both datasets. Increasing age, holding other factors constant, initially increased live tree C to a maximum at *c.* 150 years, but decreased it by 15–19 Mg C ha^{-1} in the oldest sites (unimodal relationship; Fig. 4; Table 4). PSV responded less strongly to environmental variation and had smaller impacts on live tree

Table 3 Comparison of the full SEM containing state factors, best community-weighted mean plant trait (CWM) and best diversity (Div) metric, with the state-factor-only and state factor plus diversity SEMs

							Full SEM R^2				
C pool	Dataset	CWM	Diversity	Best Div	dAICc (between best and next best full SEM)	Full SEM CFI	CWM	Div	C pool	State factor SEM R^2, C pool	State Factor + Diversity SEM R^2, C pool
LTC	Full	MaxH	PSV	Linear	16.9	1.000	0.3969	0.1968	0.4419	0.3370	0.3503
	Mod Drain	MaxH	PSV	Linear	4.7	1.000	0.3766	0.1440	0.4139	0.3034	0.3160
SDC	Full	WDR	FDis.TolS	Linear	26.8	0.993	0.0667	0.1458	0.0249	0.0163	0.0194
	Mod Drain	WDR	FDis.TolS	Linear	42.0	1.000	0.0614	0.1045	0.0224	0.0126	0.0153
OHC	Full	LMA	FDis.LMA	Nonlinear	5.6	0.989	0.5027	0.1608	0.6712	0.6634	0.6647
	Mod Drain	LMA	FDis.LMA	Linear	24.5	1.000	0.4650	0.0313	0.2354	0.1922	0.1956
TC	Full	MaxH	lnFDis.WD	Nonlinear	2.8	0.993	0.3969	0.2667	0.6011	0.5989	0.5996
	Mod Drain	MaxH	FDis.WD	Linear	3.0	1.000	0.3766	0.2236	0.1177	0.1009	0.1055

CWM and diversity metric abbreviations as in Table 1. Full = all sites (n = 2624). Mod. Drain = Moderately drained sites, drainage classes 2–4 (n = 2323). LTC = Live Tree Carbon. SDC = Standing Dead Carbon. OHC = Organic Horizon Carbon. TC = Total Carbon. A comparative fit index (CFI) > 0.9 indicates a good fit. dAIC$_c$ was used to compare SEMs with linear and nonlinear diversity.

Table 4 Standardised (Std) total (direct + indirect) and partial (direct only) effect sizes, unstandardised (Unstd) total effect sizes, and net C gain/loss (in Mg C ha^{-1}) across the range of values for each state or interactive factor for live tree C (LTC), organic horizon C (OHC) and total C (TC; standing dead values in Table S10). Factor effects are ranked by standardised total effect magnitude. Unstandardised partial estimates are shown in Figs 4–6. Ranges of values for each state and interactive factor are shown in Table S2

Full					Moderate drainage				
Effect	Std Total (SE)	Std Partial (SE)	Unstd Total (SE)	C gain	Effect	Total (SE)	Partial (SE)	Unstd Total (SE)	C gain
T → LTC	0.47 (0.015)	0.23 (0.019)	9.00 (0.322)	81	T → LTC	0.49 (0.015)	0.24 (0.02)	9.24 (0.345)	83
CWM MaxH → LTC	0.39 (0.018)	0.39 (0.018)	3.31 (0.160)	78	CWM MaxH → LTC	0.40 (0.019)	0.40 (0.019)	3.33 (0.169)	79
Age → LTC	0.36 (0.025)	0.29 (0.024)	0.26 (0.019)	71	Age → LTC	0.35 (0.026)	0.27 (0.024)	0.26 (0.020)	70
Age2 → LTC	−0.27 (0.025)	−0.20 (0.023)	−1.99 (0.183)	−85	Age2 → LTC	−0.26 (0.026)	−0.18 (0.024)	−2.07 (0.208)	−89
Drain2 → LTC	−0.15 (0.019)	−0.10 (0.018)	−2.55 (0.330)	−24	PSV → LTC	0.14 (0.017)	0.14 (0.017)	16.0 (2.011)	16
PSV → LTC	0.15 (0.016)	0.15 (0.016)	16.64 (1.852)	17	Drain → LTC	−0.11 (0.019)	−0.03 (0.018)	−5.75 (0.959)	−11
Drain → LTC	−0.14 (0.019)	−0.05 (0.018)	−4.35 (0.618)	−26	Drain2 → LTC	−0.09 (0.018)	−0.04 (0.017)	−5.96 (1.191)	−7
Ppt → LTC	0.06 (0.017)	0.01 (0.016)	0.02 (0.004)	13	T x PPt → LTC	0.06 (0.018)	0.07 (0.017)	0.01 (0.003)	18
T x PPt → LTC	0.06 (0.017)	0.06 (0.016)	0.01 (0.003)	16	Ppt → LTC	0.05 (0.018)	0.00 (0.017)	0.01 (0.005)	11
Total Age → LTC				−15	Total Age → LTC				−19
Total Drain → LTC				−50	Total Drain → LTC				−18
eDrain → OHC	0.77 (0.008)	0.77 (0.008)	1.08 (0.018)	436	CWM LMA → OHC	0.27 (0.024)	0.27 (0.024)	0.13 (0.012)	34
CWM LMA → OHC	0.10 (0.016)	0.10 (0.016)	0.13 (0.021)	34	Drain → OHC	0.25 (0.020)	0.21 (0.020)	16.11 (1.292)	32
T → OHC	−0.05 (0.012)	0.02 (0.015)	−3.12 (0.773)	−28	T → OHC	−0.23 (0.019)	−0.07 (0.024)	−5.50 (0.464)	−50
ln(FDis LMA) → OHC	−0.03 (0.012)	−0.03 (0.012)	−1.29 (0.519)	−12	Drain2 → OHC	0.16 (0.020)	0.13 (0.019)	13.27 (1.602)	15
Ppt → OHC	−0.03 (0.012)	−0.02 (0.012)	−0.02 (0.010)	−20	T x Ppt → OHC	0.11 (0.020)	0.09 (0.020)	0.02 (0.004)	40
T x Ppt → OHC	0.02 (0.012)	0.01 (0.012)	0.01 (0.006)	16	Ppt → OHC	−0.1 (0.019)	−0.07 (0.019)	−0.03 (0.006)	−24
Age → OHC	0.02 (0.012)	0.02 (0.012)	0.04 (0.030)	11	FDis LMA → OHC	−0.08 (0.019)	−0.08 (0.019)	−8.68 (1.942)	−13
					Age → OHC	0.08 (0.020)	0.09 (0.019)	0.08 (0.019)	21
					Total Drain → OHC				47
eDrain → TC	0.77 (0.009)	0.77 (0.009)	1.04 (0.018)	420	Age → TC	0.35 (0.029)	0.32 (0.030)	0.36 (0.031)	98
Age → TC	0.14 (0.020)	0.13 (0.020)	0.33 (0.047)	91	Age2 → TC	−0.19 (0.030)	−0.17 (0.030)	−2.14 (0.330)	−92
Age2 → TC	−0.10 (0.020)	−0.09 (0.020)	−2.27 (0.468)	−97	Drain → TC	0.15 (0.021)	0.17 (0.021)	10.54 (1.518)	21
T → TC	0.09 (0.013)	0.07 (0.016)	5.80 (0.812)	52	CWM Max H → TC	0.14 (0.025)	0.14 (0.025)	1.64 (0.289)	39
CWM Max H → TC	0.06 (0.016)	0.06 (0.016)	1.63 (0.442)	39	T → TC	0.14 (0.021)	0.10 (0.026)	3.63 (0.547)	33
ln(FDis WD) → TC	−0.03 (0.014)	−0.03 (0.014)	−1.24 (0.578)	−12	T x Ppt → TC	0.12 (0.021)	0.11 (0.021)	0.03 (0.004)	48
T x Ppt → TC	0.03 (0.013)	0.02 (0.013)	0.01 (0.007)	24	Drain2 → TC	0.09 (0.021)	0.10 (0.021)	8.02 (1.886)	9
Ppt → TC	0.00 (0.013)	−0.01 (0.013)	0.00 (0.011)	−1	FDis WD → TC	−0.08 (0.022)	−0.08 (0.022)	−11.35 (3.329)	−15
Total Age → TC				−6	Ppt → TC	−0.03 (0.020)	−0.04 (0.021)	−0.01 (0.007)	−8
					Total Age → TC				6
					Total Drain → TC				30

C than mean traits: increasing from minimum to maximum diversity in both datasets increased live tree C by 16 Mg C ha^{-1} (Fig. 4; Table 4).

Neither state nor interactive factors explained much variation in standing dead C. The state factor SEMs explained ≤ 2.5% of the variation in standing dead C and adding mean trait and diversity indices increased R^2 by < 1% (Tables 3, S10; Fig. S4).

Drainage dominated direct effects on organic horizon C in the full dataset, with organic horizon C increasing exponentially as sites became poorly drained (Fig. 3; Table 4). Moving from fast to slow drainage increased that C pool by 435 Mg C ha^{-1} (Fig. 5; Table 4). In the moderately drained dataset, the restricted range of site drainages reduced the impact of state factors (Tables 3 and 4). Drainage also indirectly increased organic horizon C by increasing mean LMA in both datasets (Fig. 5). Climate variables had no or only weak *direct* effects on organic horizon C in the full and moderately drained datasets respectively. After accounting for indirect temperature and precipitation effects on organic horizon C via interactive biotic factors, the temperature by precipitation interaction remained significant only in the moderately drained sites (Fig. 5). For the full set, direct climate effects disappeared; they were mediated through biotic factors, with strong negative temperature effects on mean LMA. Similarly, direct age effects on organic horizon C were minor in the moderately drained or insignificant in the full datasets (Table 3, Fig. 5). Again, age effects were moderated by interactive biotic factors, resulting in positive indirect age effects in both datasets (Fig. 5).

Biotic factors had significant but smaller effects on organic horizon C than live tree C. Increasing LMA across the range of values in both datasets increased organic horizon C by 30-35 Mg C ha^{-1}, but in the moderate drainage dataset, this was

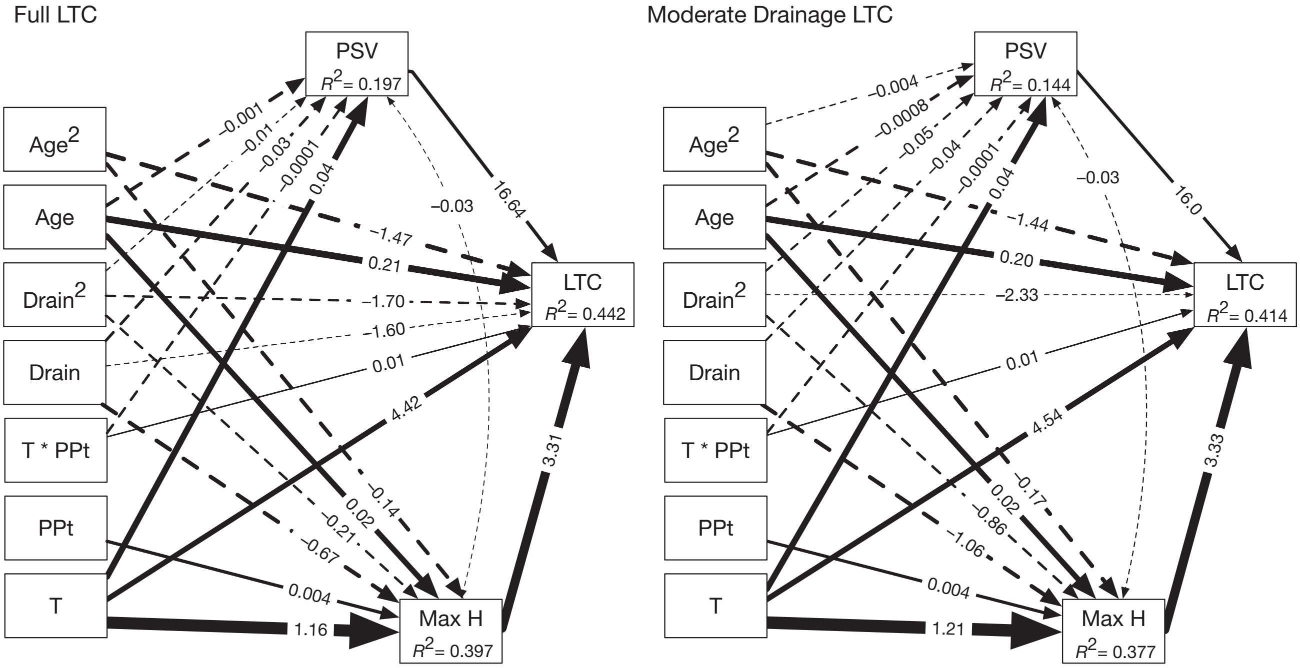

Fig. 4 Full structural equation models (SEMs) for live tree C (LTC) in each dataset. Arrows are scaled by standardised path coefficient values. Unstandardised path coefficients are shown. Dashed lines indicate negative and solid lines indicate positive path coefficients. Only significant pathways are shown ($P < 0.05$). Age = time since disturbance; Drain = drainage class; T = mean annual temperature; Ppt = mean annual precipitation; T*Ppt = interaction of mean annual temperature and mean annual precipitation. Diversity and functional trait abbreviations as in Table 1.

one of the dominant effects, on par with the magnitude of temperature and drainage impacts (Table 4). Again, plant trait effects were stronger than those of diversity. Although *mean* LMA increased organic horizon C, FDis of LMA *decreased* organic horizon C by a maximum of 12–13 Mg C ha^{-1} (Table 4).

Drivers of total C reflected the most important variables affecting live tree and organic horizon C. As with organic horizon C, drainage had the largest effect in the full dataset: reducing drainage from fast to slow increased total C by 420 Mg C ha^{-1} (Fig. 6; Table 4). It had a smaller impact in the moderately drained sites, increasing total C by only 30 Mg C ha^{-1}. Climate, age and mean plant trait (MaxH) effects on total C were similar in direction to live tree C. However, path coefficients were muted compared to live tree C (Figs 4–6): increasing maximum tree height from its minimum to maximum value increased total C by 39 Mg C ha^{-1} in both datasets, roughly half of the increase for live tree C. Diversity effects – both variables and direction – differed between live tree C and total C. In full and moderately drained sites, increasing the FDis of wood density across the full range of values *decreased* total C by 12 and 15 Mg C ha^{-1} respectively (Fig. 6), similar to negative diversity effects on organic horizon C.

DISCUSSION

Overview

The state-factor paradigm provided a clear conceptual framework to integrate BEF with ecosystem ecology. Our intent was to integrate diversity with known controls on ecosystem properties by asking: what does including diversity as a control add to our knowledge of ecosystem C storage? Our results suggest that diversity effects on ecosystem processes are an important second dimension of organismal traits as an interactive factor (Fig. 1) (Díaz *et al.* 2007; Chapin *et al.* 2011). Diversity and mean traits both responded to external state factors and directly influenced C pools in SEMs. That said, diversity *per se* (richness, phylogenetic, or functional) generally had substantially less impact on C pools than mean traits, and state factors together had a similar or stronger impact than interactive biotic factors. Our conclusions differ somewhat from recent studies (Liang *et al.* 2016; Duffy *et al.* 2017; Oehri *et al.* 2017): while those studies emphasised significant diversity effects and similar ranking to some individual abiotic variables in landscape-scale studies, we found that diversity had significant, but small, impacts on forest C storage. Our approach is conceptually consistent with recent studies of diversity effects on ecosystem properties at landscape scales (Grace *et al.* 2007; Grace *et al.* 2016; Cavanaugh *et al.* 2014; Poorter *et al.* 2017), but explicitly differentiated between state and interactive factors, included mean traits and diversity as drivers, and included both live and detrital C storage. Our findings are also consistent with studies that found diversity to have relatively small impacts on ecosystem processes across broad geographic gradients in dryland (Maestre 2012), tropical (Poorter *et al.* 2017), and grassland (Grace *et al.* 2007) ecosystems. However, we do not propose that this will always hold, either across ecosystems or across processes, depending both on ecological differences among ecosystems (see, for example, Paquette & Messier 2011; Liang *et al.* 2016) and the range of variability in key drivers within any given study (as we saw in comparing our two datasets). Within

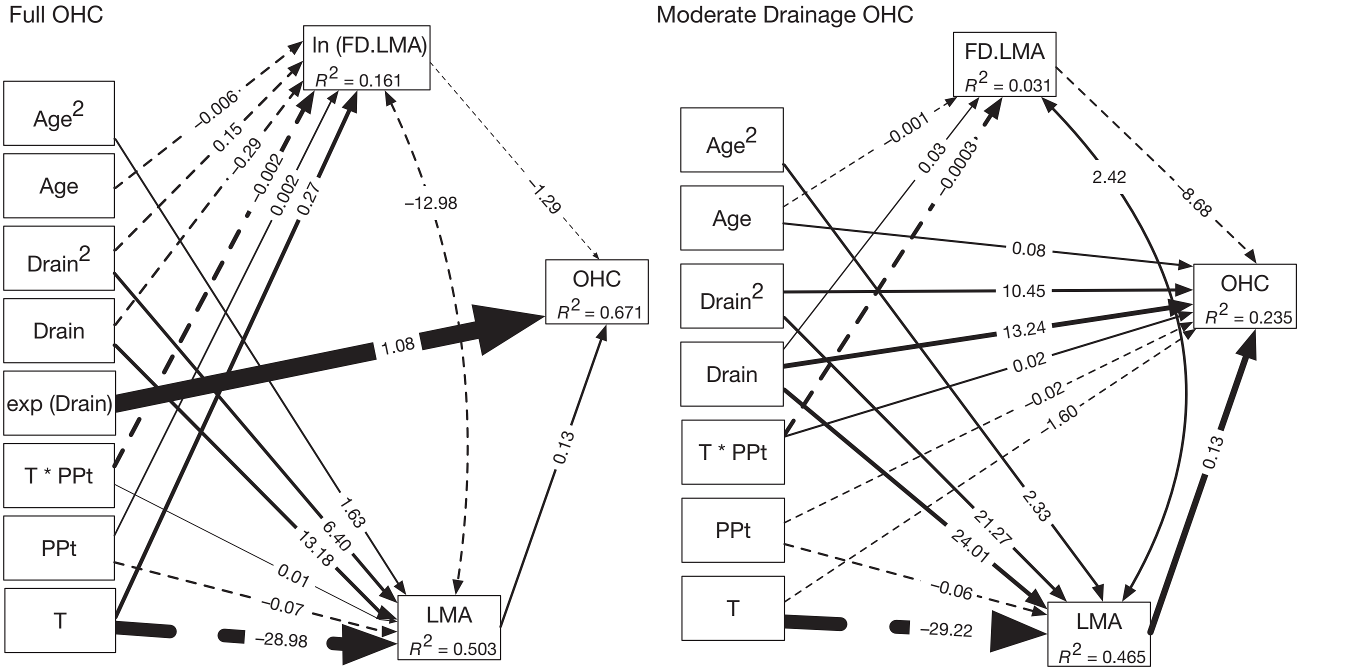

Fig. 5 Full structural equation models (SEMs) for organic horizon C (OHC) in each dataset. Arrows are scaled by standardised path coefficient values. Unstandardised path coefficients are shown. Dashed lines indicate negative and solid lines indicate positive path coefficients. Only significant pathways are shown ($P < 0.05$). Abbreviations as in Fig. 4.

tropical forests, some studies show stronger effects of diversity than mean traits and environmental variables (e.g. Cavanaugh *et al.* 2014) and others show the opposite (Poorter *et al.* 2017). More importantly, we argue that evaluating drivers together using common metrics (e.g. actual changes in ecosystem services) and integrating them within the state-factor framework will paint a more complete picture of mechanisms affecting C storage, or other ecosystem services, than opportunistic subsets of drivers alone. The approach we advocate will also guide consistent data collection for future studies aiming to refine and test within- and cross-system comparisons.

Our results emphasise the importance of modeling processes independently when assessing complex ecosystem services. When we assessed total C, rather than independent pools, biotic effects were masked, obscuring underlying mechanisms. While some complex processes (Caliman *et al.* 2013) or suites of positively correlated processes (Maestre *et al.* 2012) may show stronger diversity effects, the pattern we saw is likely whenever contributing processes exhibit different responses to environmental drivers. Furthermore, quantifying responses of individual processes to diversity allowed us to compare positive and negative diversity effects on a similar scale (i.e. Mg C ha^{-1}) and to evaluate the net impact on the ecosystem service of C storage. This has not been possible with previous multifunctionality studies (Zavaleta *et al.* 2010; Maestre *et al.* 2012; Byrnes *et al.* 2014). Understanding how complex 'bundles' of ecosystem services will respond to changing environmental conditions and management (Raudsepp-Hearne *et al.* 2010; Balvanera *et al.* 2014; Cavender-Bares *et al.* 2015) will require understanding the relationships among state and interactive factors – both abiotic and biotic – and the processes that contribute to specific ecosystem services.

Relative impacts of mean traits and diversity on C storage at landscape extents

Across C pools, interactive biotic factors had the strongest effect on live tree C. Increasing from minimum to maximum diversity and plant trait values added 16 and 78 Mg live tree C ha^{-1} respectively, an amount comparable to the direct increase in live tree C across the full range of temperatures. The consistently positive diversity effects on live tree C agree with other studies that assessed diversity effects on live tree biomass and productivity (e.g. Paquette & Messier 2011; Gamfeldt *et al.* 2013; Liang *et al.* 2016; Oehri *et al.* 2017). Our results further suggest that mean plant traits and diversity influenced C pools independently. Plot experiments frequently show diversity effects on plant biomass via complementarity, facilitation and insurance effects (Cardinale *et al.* 2011; Tilman *et al.* 2014), in addition to the effects of particular species (sampling effects) (Loreau & Hector 2001). Similarly, the consistent significance of both mean traits and diversity in our models suggests that some diversity effects occur orthogonally to mean traits (Cavanaugh *et al.* 2014; Ruiz-Benito *et al.* 2014; Poorter *et al.* 2017).

For organic horizon C, drainage was the dominant driver. Despite this, total biotic effects across the range of values in these sites were still half the size of their effects on live tree C. Mean functional traits increased organic horizon C by a maximum of 34 Mg C ha^{-1} and diversity *decreased* it by a maximum of 12–13 Mg C ha^{-1}. Negative effects of FDis of LMA on organic horizon C stocks in our dataset align with recent

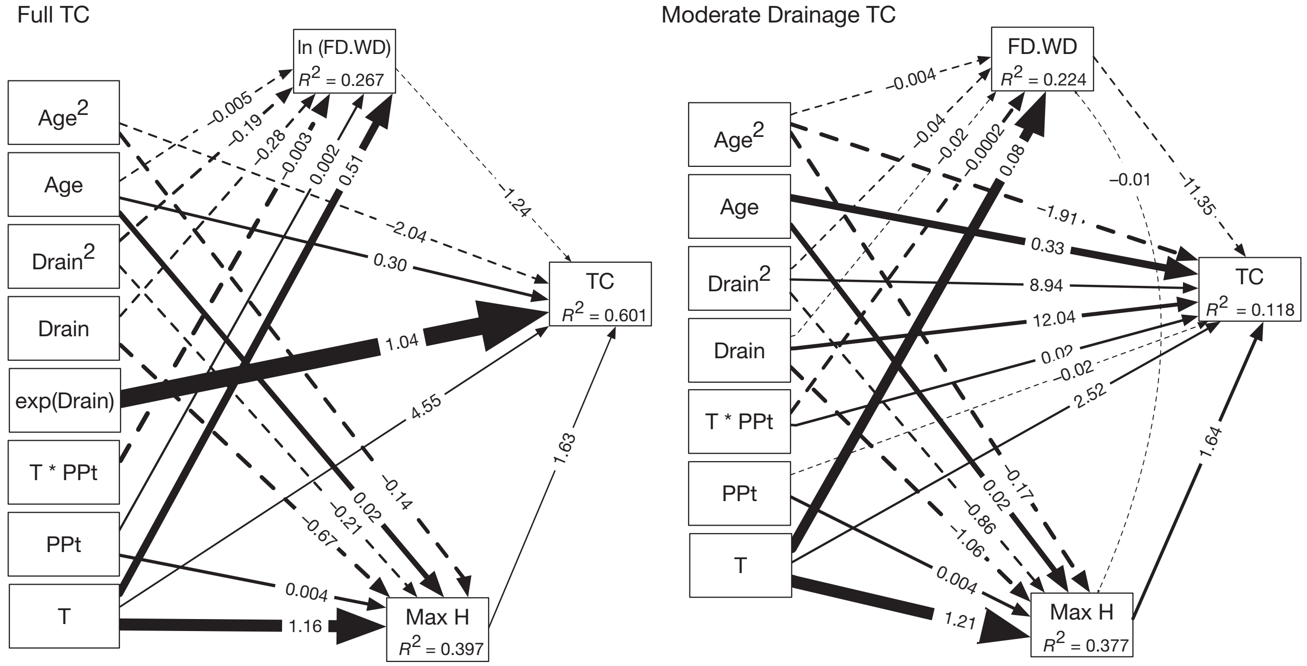

Fig. 6 Full structural equation models (SEMs) for total C (TC) in each dataset. Arrows are scaled by standardised path coefficient values. Unstandardised path coefficients are shown. Dashed lines indicate negative and solid lines indicate positive path coefficients. Only significant pathways are shown ($P < 0.05$). Abbreviations as in Fig. 4.

estimates of positive effects of litter mixing on decomposition rates (e.g. Handa *et al.* 2014), which leads to smaller soil C pools. As discussed below, however, among-study variability indicates that we still have much to learn about the mechanisms driving such responses (Gessner *et al.* 2010).

Despite the importance and independence of diversity in our results, it ranked well below mean functional traits as a biotic predictor of C storage (Table 4, Figs 4–6). Díaz *et al.* (2007) also found stronger mean trait than diversity influences for various ecosystem services. In our study, diversity always had low standardised effect sizes (< 0.15) and accounted for C gains or losses of ≤ 16 Mg C ha^{-1}. These effects were lowest for non-living pools, where C stored often greatly exceeded that in live trees. These findings contrast with recent analyses that found diversity effects to be as strong as the most important abiotic variables across landscapes (Liang *et al.* 2016; Duffy *et al.* 2017; Oehri *et al.* 2017). Few other studies have assessed plant diversity effects on non-living pools at these scales, but those that have also found weak (although positive) effects of tree richness on soil C pools (Gamfeldt *et al.* 2013) and other belowground properties (Maestre *et al.* 2012). Additional analyses that explicitly incorporate the state-factor framework will help assess whether our results about the relative effect sizes of diversity on C storage are general.

Impacts of different state and interactive biotic factor controls on C pools

Direct and indirect state factor controls

The strength and direction of state factors and interactive biotic drivers varied by C pool. State factor SEMs explained 30–34% of variation in live tree C, 19–66% in organic horizon C, and 10–60% in total C, but very little variation in standing dead C (≤ 2%). Live tree C responses to state factors were consistent with known biogeochemical and successional drivers of plant production and biomass, such as positive interactions between temperature and precipitation (Fig. S5), unimodal effects of drainage, and unimodal relationships with age (Díaz *et al.* 2009a; Chapin *et al.* 2011). Similarly, responses for organic horizon C were consistent with the positive effects of temperature and precipitation on decomposition rates (Fig. S5), negative effects of anaerobic conditions on decomposition, and accumulation of C in soil and litter layers over successional time (Díaz *et al.* 2009a; Chapin *et al.* 2011).

However, substantial effects of abiotic state factors on C pools also occurred as indirect effects via interactive biotic factors. For example, MaxH increased with stand age, temperature, and precipitation, and diversity increased with temperature, while also having positive direct effects on live tree C. Such interactive effects are well known in ecosystem ecology, when indirect abiotic effects via plant traits can outweigh direct abiotic effects on decomposition (e.g. Van Cleve *et al.* 1991; Vitousek *et al.* 1994) and production (e.g. Lauenroth & Sala 1992; Hooper & Johnson 1999). In our datasets, changes in mean functional traits across sites resulted from changing relative abundances and species turnover due to changing abiotic conditions. The strength of these effects, particularly for live tree C, suggests that species turnover across sites (beta diversity) may be more important in driving ecosystem properties across broad geographic extents than alpha diversity within sites (e.g. Winfree *et al.* 2015).

Different interactive biotic factor controls on component C pools

Mean functional traits differentially affected the various pools, with strongest effects on live tree C. MaxH had the strongest direct effects on live tree C and was correlated with several other traits (positive with leaf N, wood density, and leaf size; negative with leaf longevity and LMA; Table S11). These traits are consistent with late successional forest growth, light-competitive life history traits (Grime 2001), and faster growth on the leaf economic spectrum (Hobbie 1992; Wright *et al.* 2004; Reich 2014), which predominate in locations with high resource availability and long disturbance intervals. Therefore, the strong effects of maximum height may reflect species' natural associations with soil fertility (i.e. larger species in high fertility sites), for which we did not have a direct measure.

Mean functional traits had weaker effects on organic horizon C than live tree C. The principal trait affecting this pool reflected previous understanding from ecosystem ecology: positive correlation of LMA with organic horizon C, consistent with its negative effects on decomposition rates (Table 4, Fig. 5; Díaz *et al.* 2009a). LMA was also negatively correlated with leaf N (Table S11), which could further explain the impact of LMA on organic horizon C, as declining N content reduces decomposition rates (Hobbie 2015) and increases landscape-scale soil C (Díaz *et al.* 2007; Gamfeldt *et al.* 2013).

Diversity had its strongest impact on live tree C via PSV. The independence of this effect from trait means suggests diversity increased resource uptake or environmental tolerances of the entire community, increasing productivity and live tree C. The importance of PSV over functional diversity metrics suggests that the diversity of unmeasured traits affected this pool, though which and how requires further study (Paquette *et al.* 2015). On the other hand, where diversity was significant for organic horizon C, its effects were negative. Increasing FDis of LMA *decreased* organic horizon C. Mechanisms driving this relationship remain unclear, as decomposition rates show variable responses to aboveground diversity: litter mixtures can show increases, decreases, or no effects on decomposition compared to single-species litter (Srivastava *et al.* 2009; Handa *et al.* 2014; Jewell *et al.* 2015, 2017) and sparse literature shows inconsistent diversity effects on soil and litter pools (Díaz *et al.* 2007; Fornara & Tilman 2008; Gamfeldt *et al.* 2013; Lange *et al.* 2015; Isbell *et al.* 2017a). However, observations from forests generally support increased decomposition with increasing diversity (Nadrowski *et al.* 2010), potentially via priming effects or decomposer composition and diversity (Srivastava *et al.* 2009; Gessner *et al.* 2010; Hooper *et al.* 2012). Manipulative experiments and assessing the generality of our results across other ecosystems will provide tests of the diversity mechanisms hypothesised here.

Our results raise two related questions about diversity metrics. First, which is best to use: species richness, phylogenetic diversity, or functional diversity (e.g. Cadotte *et al.* 2009; Cadotte 2015; Cardinale *et al.* 2015; Venail *et al.* 2015) - or all of them (Villéger *et al.* 2008; Naeem *et al.* 2016)? Second, if functional diversity, which traits should be included? Because we did not use multiple diversity metrics in the same model, our results best answer the question, 'Which single diversity metric was the strongest driver of C pools, within the state-factor framework?' Of metrics tested, richness was never foremost at predicting C pools. PSV had the strongest positive effects on live tree C; FDis of LMA had weaker effects on organic horizon C. Although we expected functional trait diversity to outperform other metrics because of its more direct links to processes than species richness or phylogenetic diversity (Villéger *et al.* 2008), this varied by C pool. When FDis was a strong driver in our models, it typically used single traits related to known ecosystem mechanisms, consistent with diversity effects in plot-scale studies, such as complementarity, facilitation, and decomposer interactions (Gessner *et al.* 2010; Cardinale *et al.* 2011). We found comparatively poor performance of FDis combinations of multiple traits. Our results suggest that combining many functional traits into one metric may obscure mechanisms affecting C pools (Bernhardt-Römermann *et al.* 2008). In addition, while we have confidence in the traits selected by our statistical procedures, selecting *a priori* among many potential traits for use in functional diversity metrics remains a vexing problem (Petchey & Gaston 2006; Laliberté & Legendre 2010; Mouchet *et al.* 2010).

Assessing total C pools: relationship to complex ecosystem services

Combining C pools reduced the predictive power of state and interactive biotic factors. Responses of total C to drainage followed relationships found for soil organic horizon C, but the remaining abiotic state factors followed patterns found in live tree C. Combining component pools into total C greatly reduced the importance of biotic variables. For example, maximum tree height was among the strongest effects on live tree C, and diversity (PSV) also made significant contributions. However, when assessing total C, mean trait and diversity effects greatly decreased. This suggests that combining pools into total C obscured the importance of biotic variables and the underlying mechanisms driving C accumulation (see also Gamfeldt *et al.* 2013).

Diversity did not simultaneously maximise all C pools. Indeed, assessing diversity effects only on live tree C would have missed contrasting effects on organic horizon C, one of the largest ecosystem carbon pools. Similar cancelling effects muted responses of forest C dynamics to diversity in the tropics: tree species richness increased both growth of surviving trees and biomass loss by mortality, with no significant effect on net change in aboveground biomass (Poorter *et al.* 2017). Multifunctionality studies indicate that such tradeoffs are common (Zavaleta *et al.* 2010; Byrnes *et al.* 2014), such that conditions that maximise one ecosystem service may not maximise others (e.g. the well-known tradeoffs between production and regulation services; Cavender-Bares *et al.* 2015). On the other hand, some ecosystem processes or services may be positively associated with one another, or may depend on different traits of different species, leading, for example, to greater effects of diversity on multiple or more complex functions (e.g. Hooper & Vitousek 1998; Bracken & Stachowicz 2006; Isbell *et al.* 2011; Maestre *et al.* 2012; Caliman *et al.* 2013; Barnes *et al.* 2018). In our C storage example, however, contrasting effects of diversity and effects of different functional traits on different C pools meant that explicitly modeling the drivers of each property provided the clearest path toward understanding the underlying responses of C storage (Gamfeldt *et al*. 2013; Ricketts *et al*. 2016).

Conclusions: The merger of BEF and ecosystem ecology?

Using the well-established state-factor framework of ecosystem ecology provided insights into abiotic state factor and interactive biotic controls of C storage and helped reveal underlying mechanisms. We used SEMs to test this framework, but they are not in themselves a panacea for merging BEF and ecosystem ecology. First, SEMs can only assess effects of biotic variables if they have unshared variance with abiotic drivers. If biotic factors are so highly correlated with abiotic state factors that their effects are indistinguishable (e.g. MAP, taxon richness, and plant and soil nutrient content across precipitation gradients in dryland ecosystems; Jing *et al*. 2015), only experimental approaches can tease apart interactive effects (e.g. Van Cleve *et al*. 1983; Lauenroth & Sala 1992; Hooper & Johnson 1999). However, the state-factor framework can and should inform interpretation of covariance among abiotic factors, diversity, functional traits, and ecosystem properties, rather than attributing the effects of shared variance to diversity alone (e.g. Jing *et al*. 2015). Thus, there is a need for complementary experimental approaches for better understanding how ecosystem properties might change with species gain or loss at multiple sites across environmental gradients (Wardle *et al*. 2011; Hooper *et al*. 2012; Craven *et al*. 2016).

Because of strong collinearity among potential biotic predictors, we restricted our work to identifying the best single predictors among mean trait and diversity metrics. Greater predictability may be achieved using multiple traits or diversity metrics (Díaz *et al*. 2007; Villéger *et al*. 2008; Naeem *et al*. 2016), incorporating individual species or functional type effects (Nadrowski *et al*. 2010; Gamfeldt *et al*. 2013), or assessing interactions and feedbacks among traits, diversity and abiotic variables (Grace *et al*. 2016). For example, we did not include non-recursive SEM feedbacks of C storage (biomass) on diversity, as in Grace *et al*. (2016). While a feedback from biomass to diversity is plausible in temperate and boreal forests, its influence is likely small, as diversity typically responds to changes in broad climatic and topographic gradients and successional status (Mittelbach *et al*. 2001), as captured in our analyses. While our results provide clues about C storage mechanisms, they should be tested directly and mechanisms driving the strong relationship with PSV determined.

Our approach emphasises a shift in perspective when applying BEF results at the landscape scale for ecosystem management. The catchall term 'biodiversity' continues to cause confusion (Hooper *et al*. 2005). Often, 'biodiversity effects' are cited when what is meant is 'the effects of traits, or presence/absence of particular species,' rather than effects of diversity *per se*, as clarified by Cardinale *et al*. (2012). It makes more sense – and would cause less confusion – to refer to diversity as a subset of interactive biotic factors. Doing so more explicitly integrates BEF studies into the powerful state-factor paradigm of ecosystem ecology. Rather than merely assessing diversity effects to assert that 'diversity matters,' this perspective emphasises mechanistic understanding of when, where, why and how much it matters, in combination with other known state and interactive (e.g. mean plant traits) factors. These questions are much more relevant to ecosystem services management than such vague (and common) statements as 'biodiversity increases ecosystem functioning'. The ecosystem perspective argues that, rather than existing as its own subdiscipline, BEF becomes more relevant by explicitly integrating diversity with other controls on ecosystem properties in the state-factor framework.

ACKNOWLEDGEMENTS

Initiation of this work was supported by the Working Group on 'Biodiversity and the Functioning of Ecosystems: Translating Model Experiments into Functional Reality' at the National Center for Ecological Analysis and Synthesis, a Center funded by NSF (Grant no. EF-0553768), the University of California, Santa Barbara, and the State of California. We thank the members of this working group for many stimulating conversations at the outset of this project. We also thank Jarrett Byrnes and Jim Regetz for statistical consultation and aid with data management. We thank the *Ministère des Forêts, de la Faune et des Parcs du Québec* (Canada) for sharing the data. Aimée Classen, Brad Cardinale, and four anonymous reviewers provided useful comments on earlier drafts of this manuscript. Funding for E.C. Adair was provided by Vermont EPSCoR with funds from National Science Foundation Grant EPS-1101317.

AUTHOR CONTRIBUTIONS

ECA and DUH contributed equally to the writing of this manuscript. AP, with help from all authors, developed the working dataset from the original data provided by the *Ministère des Forêts, de la Faune et des Parcs du Québec* (Canada), including synthesis of plant trait data, and calculating diversity metrics, community-weighted mean plant traits, and carbon pool sizes. All authors contributed to developing the candidate model sets. ECA performed the model comparisons, model averaging analyses, and the structural equation modeling. DUH performed principle component analyses. All authors contributed substantially to interpretation of results and manuscript revisions.

DATA ACCESSIBILITY STATEMENT

The working data are archived on Dryad, https://doi.org/10.5061/dryad.4hj00fr. The original dataset is publically available from the *Ministère des Forêts, de la Faune et des Parcs du Québec* (Canada): http://mffp.gouv.qc.ca/le-ministere/acces-aux-donnees-gratuites/

REFERENCES

Balvanera, P., Pfisterer, A.B., Buchmann, N., He, J.S., Nakashizuka, T., Raffaelli, D. *et al*. (2006). Quantifying the evidence for biodiversity effects on ecosystem functioning and services. *Ecol. Lett*., 9, 1146–1156.

Balvanera, P., Siddique, I., Dee, L., Paquette, A., Isbell, F., Gonzalez, A., *et al.* (2014). Linking biodiversity and ecosystem services: current uncertainties and the necessary next steps. *BioSci*, 64, 49–57.

Barnes, A.D., Jochum, M., Lefcheck, J.S., Eisenhauer, N., Scherber, C., O'Connor, M.I., *et al.* (2018). Energy flux: The link between multitrophic biodiversity and ecosystem functioning. *Trends Ecol. Evol.*, 33, 186–197.

Bernhardt-Römermann, M., Römermann, C., Nuske, R., Parth, A., Klotz, S., Schmidt, W. *et al.* (2008). On the identification of the most suitable traits for plant functional trait analyses. *Oikos*, 117, 1533–1541.

Bracken, M.E.S. & Stachowicz, J.J. (2006). Seaweed diversity enhances nitrogen uptake via complementary use of nitrate and ammonium. *Ecology*, 87, 2397–2403.

Burnham, K.P. & Anderson, D.R. (2002). *Model Selection and Multi-Model Inference: A Practical Information Theoretic Approach*, 2nd edn. Springer-Verlag, New York.

Byrnes, J.E.K., Gamfeldt, L., Isbell, F., Lefcheck, J.S., Griffin, J.N., Hector, A., *et al.* (2014). Investigating the relationship between biodiversity and ecosystem multifunctionality: challenges and solutions. *Methods Ecol. Evol.*, 5, 111–124.

Cade, B.S. (2015). Model averaging and muddled multimodel inferences. *Ecology*, 96, 2370–2382.

Cadotte, M.W. (2015). Phylogenetic diversity and productivity: gauging interpretations from experiments that do not manipulate phylogenetic diversity. *Funct. Ecol.*, 29, 1603–1606.

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.

Caliman, A., Carneiro, L.S., Leal, J.J.F., Farjalla, V.F., Bozelli, R.L. & Esteves, F.A. (2013). Biodiversity effects of ecosystem engineers are stronger on more complex ecosystem processes. *Ecology*, 94, 1977–1985.

Cardinale, B.J., Matulich, K.L., Hooper, D.U., Byrnes, J.E., Duffy, E., Gamfeldt, L., *et al.* (2011). The functional role of producer diversity in ecosystems. *Am. J. Bot.*, 98, 572–592.

Cardinale, B.J., Duffy, J.E., Gonzalez, A., Hooper, D.U., Perrings, C., Venail, P., *et al.* (2012). Biodiversity loss and its impact on humanity. *Nature*, 486, 59–67.

Cardinale, B.J., Gross, K., Fritschie, K., Flombaum, P., Fox, J.W., Rixen, C., *et al.* (2013). Biodiversity simultaneously enhances the production and stability of community biomass, but the effects are independent. *Ecology*, 94, 1697–1707.

Cardinale, B.J., Venail, P., Gross, K., Oakley, T.H., Narwani, A., Allan, E., *et al.* (2015). Further re-analyses looking for effects of phylogenetic diversity on community biomass and stability. *Funct. Ecol.*, 29, 1607–1610.

Cavanaugh, K.C., Gosnell, J.S., Davis, S.L., Ahumada, J., Boundja, P., Clark, D.B., *et al.* (2014). Carbon storage in tropical forests correlates with taxonomic diversity and functional dominance on a global scale. *Glob. Ecol. Biogeogr.*, 23, 563–573.

Cavender-Bares, J., Polasky, S., King, E. & Balvanera, P. (2015). A sustainability framework for assessing trade-offs in ecosystem services. *Ecol. Soc.*, 20, 317–333.

Chapin, F.S. III, Zavaleta, E.S., Eviner, V.T., Naylor, R.L., Vitousek, P.M., Reynolds, H.L., *et al.* (2000). Consequences of changing biodiversity. *Nature*, 405, 234–242.

Chapin, F.S. III, Matson, P.A. & Vitousek, P. (2011). *Principles of Terrestrial Ecosystem Ecology*, 2nd edn.. Springer-Verlag, New York, NY.

Cornelissen, J.H.C., Pérez-Harguindeguy, N., Díaz, S., Grime, J.P., Marzano, B., Cabido, M., *et al.* (1999). Leaf structure and defence control litter decomposition rate across species, life forms and continents. *New Phytol.*, 143, 191–200.

Craven, D., Isbell, F., Manning, P., Connolly, J., Bruelheide, H., Ebeling, A., *et al.* (2016). Plant diversity effects on grassland productivity are robust to both nutrient enrichment and drought. *Philos Trans R Soc London [Biol]*, 371, pii: 20150277. https://doi.org/10.1098/rstb.2015.0277.

Díaz, S., Cabido, M. & Casanoves, F. (1999). Functional implications of trait-environment linkages in plant communities. In *Ecological Assembly Rules: Perspectives, Advances, Retreats* (eds Weiher, E., Keddy, P.). Cambridge University Press Cambridge, UK, pp. 338–362.

Díaz, S., Lavorel, S., de Bello, F., Quetier, F., Grigulis, K. & Robson, M. (2007). Incorporating plant functional diversity effects in ecosystem service assessments. *Proc. Natl Acad. Sci. USA*, 104, 20684–20689.

Díaz, S., Hector, A. & Wardle, D.A. (2009a). Biodiversity in forest carbon sequestration initiatives: not just a side benefit. *Curr. Opin. Environ. Sustain.*, 1, 55–60.

Díaz, S., Wardle, D.A. & Hector, A. (2009b). Incorporating Biodiversity in Climate Change Mitigation Initiatives. In *Biodiversity, Ecosystem Functioning, and Human Wellbeing: An Ecological and Economic Perspective* (eds Naeem, S., Bunker, D.E., Hector, A., Loreau, M., Perrings, C.). Oxford University Press, Oxford, UK, pp. 149–166.

Duffy, J.E., Godwin, C.M. & Cardinale, B.J. (2017). Biodiversity effects in the wild are common and as strong as key drivers of productivity. *Nature*, 549, 261–264.

Finerty, G.E., de Bello, F., Bila, K., Berg, M.P., Dias, A.T.C., Pezzatti, G.B. *et al.* (2016). Exotic or not, leaf trait dissimilarity modulates the effect of dominant species on mixed litter decomposition. *J. Ecol.*, 104, 1400–1409.

Fornara, D.A. & Tilman, D. (2008). Plant functional composition influences rates of soil carbon and nitrogen accumulation. *J. Ecol.*, 96, 314–322.

Galipaud, M., Gillingham, M.A.F., David, M. & Dechaume-Moncharmont, F.X. (2014). Ecologists overestimate the importance of predictor variables in model averaging: a plea for cautious interpretations. *Methods Ecol. Evol.*, 5, 983–991.

Galipaud, M., Gillingham, M.A.F. & Dechaume-Moncharmont, F.X. (2017). A farewell to the sum of Akaike weights: The benefits of alternative metrics for variable importance estimations in model selection. *Methods Ecol. Evol.*, 8, 1668–1678.

Gamfeldt, L., Snall, T., Bagchi, R., Jonsson, M., Gustafsson, L., Kjellander, P., *et al.* (2013). Higher levels of multiple ecosystem services are found in forests with more tree species. *Nat. Commun.*, 4, 1340.

Gessner, M.O., Swan, C.M., Dang, C.K., McKie, B.G., Bardgett, R.D., Wall, D.H. *et al.* (2010). Diversity meets decomposition. *Trends Ecol. Evol.*, 25, 372–380.

Grace, J.B., Anderson, T.M., Smith, M.D., Seabloom, E., Andelman, S.J., Meche, G., *et al.* (2007). Does species diversity limit productivity in natural grassland communities? *Ecol. Lett.*, 10, 680–689.

Grace, J.B., Anderson, T.M., Seabloom, E.W., Borer, E.T., Adler, P.B., Harpole, W.S., *et al.* (2016). Integrative modelling reveals mechanisms linking productivity and plant species richness. *Nature*, 529, 390–393.

Grime, J.P. (1998). Benefits of plant diversity to ecosystems: immediate, filter and founder effects. *J. Ecol.*, 86, 902–910.

Grime, J.P. (2001). *Plant Strategies, Vegetation Processes and Ecosystem Properties*. Wiley, Chichester.

Grosse, G., Harden, J., Turetsky, M., McGuire, A.D., Camill, P., Tarnocai, C., *et al.* (2011). Vulnerability of high-latitude soil organic carbon in North America to disturbance. *J. Geophys. Res. Biogeosci.*, 116, G00K06 https://doi.org/10.1029/2010jg001507.

Handa, I.T., Aerts, R., Berendse, F., Berg, M.P., Bruder, A., Butenschoen, O., *et al.* (2014). Consequences of biodiversity loss for litter decomposition across biomes. *Nature*, 509, 218–221.

Hattenschwiler, S., Tiunov, A.V. & Scheu, S. (2005). Biodiversity and litter decomposition in terrestrial ecosystems. *Annu. Rev. Ecol. Evol. Syst.*, 36, 191–218.

Helmus, M.R., Bland, T.J., Williams, C.K. & Ives, A.R. (2007). Phylogenetic measures of biodiversity. *Am. Nat.*, 169, E68–E83.

Hobbie, S.E. (1992). Effects of plant species on nutrient cycling. *Trends Ecol. Evol.*, 7, 336–339.

Hobbie, S.E. (2015). Plant species effects on nutrient cycling: revisiting litter feedbacks. *Trends Ecol. Evol.*, 30, 357–363.

Hooper, D.U. & Johnson, L. (1999). Nitrogen limitation in dryland ecosystems: responses to geographical and temporal variation in precipitation. *Biogeochemistry*, 46, 247–293.

Hooper, D.U. & Vitousek, P.M. (1998). Effects of plant composition and diversity on nutrient cycling. *Ecol. Monogr.*, 68, 121–149.

Hooper, D.U., Chapin, F.S. III, Ewel, J.J., Hector, A., Inchausti, P., Lavorel, S., *et al.* (2005). Effects of biodiversity on ecosystem functioning: a consensus of current knowledge. *Ecol. Monogr.*, 75, 3–35.

Hooper, D.U., Adair, E.C., Cardinale, B.J., Byrnes, J.E.K., Hungate, B.A., Matulich, K.L., *et al.* (2012). A global synthesis reveals biodiversity loss as a major driver of ecosystem change. *Nature*, 486, 105–108.

Hungate, B.A., Barbier, E.B., Ando, A.W., Marks, S.P., Reich, P.B., van Gestel, N., *et al.* (2017). The economic value of grassland species for carbon storage. *Sci Adv*, 3, e1601880.

Isbell, F.I., Polley, H.W. & Wilsey, B.J. (2009). Biodiversity, productivity and the temporal stability of productivity: patterns and processes. *Ecol. Lett.*, 12, 443–451.

Isbell, F., Calcagno, V., Hector, A., Connolly, J., Harpole, W.S., Reich, P.B., *et al.* (2011). High plant diversity is needed to maintain ecosystem services. *Nature*, 477, 199–202.

Isbell, F., Adler, P.R., Eisenhauer, N., Fornara, D., Kimmel, K., Kremen, C., *et al.* (2017a). Benefits of increasing plant diversity in sustainable agroecosystems. *J. Ecol.*, 105, 871–879.

Isbell, F., Gonzalez, A., Loreau, M., Cowles, J., Díaz, S., Hector, A., *et al.* (2017b). Linking the influence and dependence of people on biodiversity across scales. *Nature*, 546, 65–72.

Jenny, H. (1980). *Soil Genesis with Ecological Perspectives*. Springer-Verlag, New York, NY.

Jewell, M.D., Shipley, B., Paquette, A., Messier, C. & Reich, P.B. (2015). A traits-based test of the home-field advantage in mixed-species tree litter decomposition. *Ann. Bot.*, 116, 781–788.

Jewell, M.D., Shipley, B., Low-Decarie, E., Tobner, C.M., Paquette, A., Messier, C. *et al.* (2017). Partitioning the effect of composition and diversity of tree communities on leaf litter decomposition and soil respiration. *Oikos*, 126, 959–971.

Jing, X., Sanders, N.J., Shi, Y., Chu, H., Classen, A.T., Zhao, K., *et al.* (2015). The links between ecosystem multifunctionality and above- and belowground biodiversity are mediated by climate. *Nat. Commun.*, 6, 8159.

Kembel, S.W., Cowan, P.D., Helmus, M.R., Cornwell, W.K., Morlon, H., Ackerly, D.D., *et al.* (2010). Picante: R tools for integrating phylogenies and ecology. *Bioinformatics*, 26, 1463–1464.

Laliberté, E. & Legendre, P. (2010). A distance-based framework for measuring functional diversity from multiple traits. *Ecology*, 91, 299–305.

Laliberté, E. & Shipley, B. (2011). FD: measuring functional diversity from multiple traits, and other tools for functional ecology. R package version 1.0-11.

Lambert, M.C., Ung, C.H. & Raulier, F. (2005). Canadian national tree aboveground biomass equations. *Can. J. For. Res.*, 35, 1996–2018.

Lamlom, S.H. & Savidge, R.A. (2003). A reassessment of carbon content in wood: variation within and between 41 North American species. *Biomass Bioenerg.*, 25, 381–388.

Lange, M., Eisenhauer, N., Sierra, C.A., Bessler, H., Engels, C., Griffiths, R.I., *et al.* (2015). Plant diversity increases soil microbial activity and soil carbon storage. *Nat. Commun.*, 6, 6707.

Lauenroth, W.K. & Sala, O.E. (1992). Long-term forage production of North American shortgrass steppe. *Ecol. Appl.*, 2, 397–403.

Lavorel, S., Grigulis, K., McIntyre, S., Williams, N.S.G., Garden, D., Dorrough, J., *et al.* (2008). Assessing functional diversity in the field - methodology matters!. *Funct. Ecol.*, 22, 134–147.

Li, Z., Kurz, W.A., Apps, M.J. & Beukema, S.J. (2003). Belowground biomass dynamics in the Carbon Budget Model of the Canadian Forest Sector: recent improvements and implications for the estimation of NPP and NEP. *Can. J. For. Res.*, 33, 126–136.

Liang, J., Crowther, T.W., Picard, N., Wiser, S., Zhou, M., Alberti, G., *et al.* (2016). Positive biodiversity-productivity relationship predominant in global forests. *Science*, 354, pii: aaf8957.

Loreau, M. & Hector, A. (2001). Partitioning selection and complementarity in biodiversity experiments. *Nature*, 412, 72–76.

Mack, M.C., Treseder, K.K., Manies, K.L., Harden, J.W., Schuur, E.A.G., Vogel, J.G., *et al.* (2008). Recovery of aboveground plant biomass and productivity after fire in mesic and dry black spruce forests of interior alaska. *Ecosystems*, 11, 209–225.

Maestre, F.T., Quero, J.L., Gotelli, N.J., Escudero, A., Ochoa, V., Delgado-Baquerizo, M., *et al.* (2012). Plant species richness and ecosystem multifunctionality in global drylands. *Science*, 335, 214–218.

Millennium Ecosystem Assessment (2005). *Ecosystems and Human Well-being: Biodiversity Synthesis*. World Resources Institute, Washington, DC.

Mittelbach, G.G., Steiner, C.F., Scheiner, S.M., Gross, K.L., Reynolds, H.L., Waide, R.B., *et al.* (2001). What is the observed relationship between species richness and productivity? *Ecology*, 82, 2381–2396.

Mouchet, M.A., Villéger, S., Mason, N.W.H. & Mouillot, D. (2010). Functional diversity measures: an overview of their redundancy and their ability to discriminate community assembly rules. *Funct. Ecol.*, 24, 867–876.

MRNFQ (2006). Normes d'inventaire forestier: placettes-échantillons permanentes, 2006 edn. Direction des inventaires forestiers. Ministère des Ressources Naturelles et de la Faune Québec, Québec, Canada.

Nadrowski, K., Wirth, C. & Scherer-Lorenzen, M. (2010). Is forest diversity driving ecosystem function and service? *Curr. Opin. Environ. Sustain.*, 2, 75–79.

Naeem, S., Bunker, D.E., Hector, A., Loreau, M. & Perrings, C. (eds.) (2009). *Biodiversity, Ecosystem Functioning, and Human Wellbeing: An ecological and economic perspective*. Oxford University Press, Oxford, UK.

Naeem, S., Prager, C., Weeks, B., Varga, A., Flynn, D.F., Griffin, K., *et al.* (2016). Biodiversity as a multidimensional construct: a review, framework and case study of herbivory's impact on plant biodiversity. *Proc. R. Soc. Lond. B Biol. Sci.*, 283, 20153005.

Nielsen, U.N., Ayres, E., Wall, D.H. & Bardgett, R.D. (2011). Soil biodiversity and carbon cycling: a review and synthesis of studies examining diversity-function relationships. *Eur. J. Soil Sci.*, 62, 105–116.

Oehri, J., Schmid, B., Schaepman-Strub, G. & Niklaus, P.A. (2017). Biodiversity promotes primary productivity and growing season lengthening at the landscape scale. *Proc. Natl Acad. Sci. USA*, 114, 10160–10165.

Paquette, A. & Messier, C. (2011). The effect of biodiversity on tree productivity: from temperate to boreal forests. *Glob. Ecol. Biogeogr.*, 20, 170–180.

Paquette, A., Joly, S. & Messier, C. (2015). Explaining forest productivity using tree functional traits and phylogenetic information: two sides of the same coin over evolutionary scale? *Ecol. Evol.*, 5, 1774–1783.

Petchey, O.L. & Gaston, K.J. (2006). Functional diversity: back to basics and looking forward. *Ecol. Lett.*, 9, 741–758.

Poorter, L., van der Sande, M.T., Arets, E.J.M.M., Ascarrunz, N., Enquist, B.J., Finegan, B., *et al.* (2017). Biodiversity and climate determine the functioning of Neotropical forests. *Glob. Ecol. Biogeogr.*, 26, 1423–1434.

van der Putten, W.H., Bardgett, R.D., Bever, J.D., Bezemer, T.M., Casper, B.B., Fukami, T., *et al.* (2013). Plant-soil feedbacks: the past, the present and future challenges. *J. Ecol.*, 101, 265–276.

R Core Team (2011). R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Raudsepp-Hearne, C., Peterson, G.D. & Bennett, E.M. (2010). Ecosystem service bundles for analyzing tradeoffs in diverse landscapes. *Proc. Natl Acad. Sci. USA*, 107, 5242–5247.

Reich, P.B. (2014). The world-wide 'fast–slow' plant economics spectrum: a traits manifesto. *J. Ecol.*, 102, 275–301.

Reid, J.P., Adair, E.C., Hobbie, S.E. & Reich, P.B. (2012). Biodiversity, nitrogen deposition, and CO_2 affect grassland soil carbon cycling but not storage. *Ecosystems*, 15, 580–590.

Ricketts, T.H., Watson, K.B., Koh, I., Ellis, A.M., Nicholson, C.C., Posner, S., *et al.* (2016). Disaggregating the evidence linking biodiversity and ecosystem services. *Nat. Commun.*, 7, 13106.

Rosseel, Y. (2012). lavaan: an R package for structural equation modeling. *J. Stat. Softw*., 48, 1–36.

Ruiz-Benito, P., Gómez-Aparicio, L., Paquette, A., Messier, C., Kattge, J. & Zavala, M.A. (2014). Diversity increases carbon storage and tree productivity in Spanish forests. *Glob. Ecol. Biogeogr*., 23, 311–322.

Schulze, E.D. & Mooney, H.A. (eds.) (1993). *Biodiversity and Ecosystem Function*. Springer-Verlag, Berlin, Germany.

Schuur, E.A.G., Chadwick, O.A. & Matson, P.A. (2001). Carbon cycling and soil carbon storage in mesic to wet Hawaiian montane forests. *Ecology*, 82, 3182–3196.

Shaw, C.H., Banfield, E. & Kurz, W.A. (2008). Stratifying soils into pedogenically similar categories for modeling forest soil carbon. *Can. J. Soil Sci*., 88, 501–516.

Srivastava, D.S. & Vellend, M. (2005). Biodiversity-ecosystem function research: is it relevant to conservation? *Annu. Rev. Ecol. Evol. Syst*., 36, 267–294.

Srivastava, D.S., Cardinale, B.J., Downing, A.L., Duffy, J.E., Jouseau, C., Sankaran, M. *et al*. (2009). Diversity has stronger top-down than bottom-up effects on decomposition. *Ecology*, 90, 1073–1083.

Taylor, P.G., Cleveland, C.C., Wieder, W.R., Sullivan, B.W., Doughty, C.E., Dobrowski, S.Z. *et al*. (2017). Temperature and rainfall interact to control carbon cycling in tropical forests. *Ecol. Lett*., 20, 779–788.

Thomas, S.C. & Martin, A.R. (2012). Carbon content of tree tissues: a synthesis. *Forests*, 3, 332.

Tilman, D., Isbell, F. & Cowles, J.M. (2014). Biodiversity and ecosystem functioning. *Annu. Rev. Ecol. Evol. Syst*., 45, 471–493.

Van Cleve, K., Oliver, L., Schlentner, R., Viereck, L.A. & Dyrness, C.T. (1983). Productivity and nutrient cycling in taiga forest ecosystems. *Can. J. For. Res*., 13, 747–767.

Van Cleve, K., Chapin, F.S. III, Dryness, C.T. & Vireck, L.A. (1991). Element cycling in taiga forest: State-factor control. *BioSci*, 41, 78–88.

Vandermeer, J.H., Lawrence, D., Symstad, A. & Hobbie, S. (2002). Effect of biodiversity on ecosystem functioning in managed ecosystems. In *Biodiversity and Ecosystem Functioning: Synthesis and Perspectives* (eds Loreau, M., Naeem, S., Inchausti, P.). Oxford University Press, Oxford, UK, pp. 221–233.

Venail, P., Gross, K., Oakley, T.H., Narwani, A., Allan, E., Flombaum, P., *et al*. (2015). Species richness, but not phylogenetic diversity, influences community biomass production and temporal stability in a re-examination of 16 grassland biodiversity studies. *Funct. Ecol*., 29, 615–626.

Villéger, S., Mason, N.W.H. & Mouillot, D. (2008). New multidimensional functional diversity indices for a multifaceted framework in functional ecology. *Ecology*, 89, 2290–2301.

Vitousek, P.M. (2004). *Nutrient cycling and limitation: Hawai'i as a model system*. Princeton University Press, Princeton, NJ.

Vitousek, P.M., Turner, D.R., Parton, W.J. & Sanford, R.L. (1994). Litter decomposition on the Mauna Loa environmental matrix, Hawai'i: Patterns, mechanisms, and models. *Ecology*, 75, 418–429.

Walker, B., Kinzig, A. & Langridge, J. (1999). Plant attribute diversity, resilience, and ecosystem function: The nature and significance of dominant and minor species. *Ecosystems*, 2, 95–113.

Wang, S. & Loreau, M. (2016). Biodiversity and ecosystem stability across scales in metacommunities. *Ecol. Lett*., 19, 510–518.

Wardle, D.A. (2016). Do experiments exploring plant diversity-ecosystem functioning relationships inform how biodiversity loss impacts natural ecosystems? *J. Veg. Sci*., 27, 646–653.

Wardle, D.A., Bardgett, R.D., Callaway, R.M. & Van der Putten, W.H. (2011). Terrestrial ecosystem responses to species gains and losses. *Science*, 332, 1273–1277.

Winfree, R., Fox, J.W., Williams, N.M., Reilly, J.R. & Cariveau, D.P. (2015). Abundance of common species, not species richness, drives delivery of a real-world ecosystem service. *Ecol. Lett*., 18, 626–635.

Wright, I.J., Reich, P.B., Westoby, M., Ackerly, D.D., Baruch, Z., Bongers, F., *et al*. (2004). The worldwide leaf economics spectrum. *Nature*, 428, 821–827.

Yachi, S. & Loreau, M. (1999). Biodiversity and ecosystem productivity in a fluctuating environment: the insurance hypothesis. *Proc. Natl Acad. Sci. USA*, 96, 1463–1468.

Zavaleta, E.S., Pasari, J.R., Hulvey, K.B. & Tilman, G.D. (2010). Sustaining multiple ecosystem functions in grassland communities requires higher biodiversity. *Proc. Natl Acad. Sci. USA*, 107, 1443–1446.

Zimmerman, E.K. & Cardinale, B.J. (2014). Is the relationship between algal diversity and biomass in North American lakes consistent with biodiversity experiments? *Oikos*, 123, 267–278.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Editor, Tadashi Fukami
Manuscript received 29 January 2018
First decision made 8 March 2018
Second decision made 20 June 2018
Manuscript accepted 24 July 2018