

Evolutionary history and the effect of biodiversity on plant productivity

Marc W. Cadotte^{a,b,1}, Bradley J. Cardinale^c, and Todd H. Oakley^c

^aNational Center for Ecological Analysis and Synthesis, University of California, 735 State Street, Santa Barbara, CA 93101; ^bDepartment of Biological Sciences, University of Toronto, 1265 Military Trail, Scarborough, ON, Canada M1C 1A; and ^cDepartment of Ecology, Evolution, and Marine Biology, University of California, Santa Barbara, CA 93106

Edited by Michael J. Donoghue, Yale University, New Haven, CT, and approved September 10, 2008 (received for review June 20, 2008)

Loss of biological diversity because of extinction is one of the most pronounced changes to the global environment. For several decades, researchers have tried to understand how changes in biodiversity might impact biomass production by examining how biomass correlates with a number of biodiversity metrics (especially the number of species and functional groups). This body of research has focused on species with the implicit assumption that they are independent entities. However, functional and ecological similarities are shaped by patterns of common ancestry, such that distantly related species might contribute more to production than close relatives, perhaps by increasing niche breadth. Here, we analyze 2 decades of experiments performed in grassland ecosystems throughout the world and examine whether the evolutionary relationships among the species comprising a community predict how biodiversity impacts plant biomass production. We show 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. Our results reveal how evolutionary history can provide critical information for understanding, predicting, and potentially ameliorating the effects of biodiversity loss and should serve as an impetus for new biodiversity experiments.

community ecology | ecosystem function | phylogenetic diversity | biodiversity experiments | metaanalysis

The modern era has come to be defined as a period of rapid environmental change. One of the most prominent changes taking place globally is a reduction in the number of genes, species, and functional groups of organisms that comprise the biological diversity of natural and managed communities. Widespread loss of biodiversity has prompted scientists from an increasing number of disciplines to begin studying the social, economic, and environmental impacts of diversity change (1–5). For example, seminal experiments in the 1990s suggested that species loss might reduce the amount of biomass produced by plants (6–9), possibly translating to a loss of important ecological services such as the ability of natural habitats to absorb CO₂ from the atmosphere. These experiments stimulated 2 decades of research detailing the functional role of plant diversity in ecosystems. Recent summaries of this body of research have confirmed that systems with fewer species generally produce less biomass than those with more species (10–14).

However, why changes in the number of species cause ecosystems to be less productive is still not entirely clear. Is it because less diverse communities tend to be missing genes, metabolic pathways, or traits that would otherwise allow a more complete utilization of local conditions (4, 15)? To answer this question would require that researchers quantify the biological traits that drive resource use and biomass production. However, because a multitude of traits are potentially associated with the ecological differences among species that drive patterns of resource use, knowing the evolutionary relationships of the members of a community can serve to quantify patterns of trait diversity (16).

Here, we present results from a formal metaanalysis of experiments performed in locations around the world to show that phylogenetic diversity is the single best predictor of how community biomass is altered by changes in species diversity. Our dataset is derived from 29 experiments that manipulated the number of species of terrestrial angiosperms in experimental plots, pots, or garden beds in fields or greenhouses and then measured the impacts of plant species number on the production of plant biomass [for a summary of studies used, see [supporting information \(SI\) Table S1](#)]. For each of the experimental units that contained more than one species for which constituent monocultures were measured, we standardized the diversity “effect size” of the biomass produced in a polyculture to the mean of the constituent species in monoculture, as the log response ratio (LR_{mean} ; see *Materials and Methods*). The pool of species used includes 177 taxa that span all major functional groups of grassland ecosystem plants (C3 and C4 graminoids, legumes, etc.). We calculated not only the number of species in a plot, but also the number of functional groups (for functional group definitions, see *Materials and Methods*) and the amount of phylogenetic diversity in a community (PD_C) in a plot (Fig. 1). PD_C measures the magnitude of the divergences among species that have evolved since a common ancestor, calculated as the sum of phylogenetic branch lengths separating species on a phylogeny. We estimated the phylogenetic relationships among species by using Bayesian inference with Ultrametric rate smoothing for 143 of the 177 species for which nucleotide sequences from 4 genes (*5.8s*, *atpb*, *matk*, and *rbcl*) were available in GenBank [National Center for Biotechnology Information (www.ncbi.nlm.nih.gov); for details, including support metrics, and for comparisons with other phylogenetic methods, see *SI Text* and Figs. S1 and S2]. We were able to estimate PD_C in 78% of all experimental polycultures (i.e., 1,315 experimental units).

Results and Discussion

Similar to prior summaries (6, 12, 17), our analyses confirm that both the number of species and the number of plant functional groups in an experiment are significant predictors of plant biomass production (Table 1, Model A, and Fig. 2*A* and *B*). The finding of our analysis is that phylogenetic diversity is also a highly significant predictor of biomass production (Table 1, Model A, Fig. 2*C*). Given that we have data on the number of plant species, the number of plant functional groups, and the phylogenetic diversity in an experimental unit, it is possible to ask which of these metrics of biological diversity best explains variation in biomass production among experimental commu-

Author contributions: M.W.C., B.J.C., and T.H.O. designed research; M.W.C. performed research; M.W.C. and T.H.O. analyzed data; and M.W.C., B.J.C., and T.H.O. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

¹To whom correspondence should be addressed. E-mail: cadotte@nceas.ucsb.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0805962105/DCSupplemental.

© 2008 by The National Academy of Sciences of the USA



Fig. 1. Summary of phylogenetic results. (A) Illustrated is the majority rule consensus from a Bayesian MCMC search, with several clades collapsed because of space considerations. Triangles represent collapsed clades, with numbers of species per clade indicated. Two clades (Fabaceae and Asteraceae) representing different functional groups are highlighted for illustrating different biodiversity metrics. Details of phylogenetic analyses and results, including support values, are presented in the *Materials and Methods* and the *SI Text*. (B) Biodiversity can be measured in a number of ways, including the number of species in a community (N), the number of trait-based functional groups (F), or the total phylogenetic diversity (PD_C) representing the members of a community. PD_C is calculated as the sum of branch lengths leading to all members of a community, and two examples are shown using orange and blue branches. The number of species and functional groups often give little information about the underlying phylogeny describing the evolutionary history of a group of species. In 2 hypothetical communities, C1 and C2, the number of species is the same, but variation exists in the number of functional groups and the calculated PD_C.

nities. Such information would potentially be useful for conservation and management where it is often beneficial to predict how community composition affects ecologically important processes. When we compared single-variable models that included the different measures of diversity as explanatory variables, PD_C was a significantly better predictor of variation in plant biomass than either the number of species or functional groups (Table 1, Model A).

When productivity was modeled as a function of species number (the form of diversity directly manipulated in these studies), PD_C, and their two-way interaction, we found that PD_C was not only a significant predictor of plant biomass, but that there was a highly significant interaction between species number and PD_C (Table 1, Model B). Thus, the impact of plant species number on biomass in past experiments is at least partially explained by the amount of phylogenetic diversity represented in an experimental unit. Further, because PD_C and species number are positively correlated (Fig. 2D), we examined the impacts of PD_C on production within species number treatments. We found that the effects of PD_C on production were most pronounced at lower numbers of species where studies typically included many different combinations of species (2 species: $F_{1,439} = 4.813$, $P = 0.029$; 4 species: $F_{1,277} = 10.435$, $P = 0.001$), whereas the impacts of PD_C became less pronounced at higher levels of diversity (6 species: $F_{1,106} = 0.026$, $P = 0.873$; 8 species: $F_{1,90} = 3.077$, $P = 0.083$), probably because researchers tended to use far fewer combinations of species, resulting in less variation in PD_C.

The number of functional groups of plants was also significantly related to PD_C among species within experimental units ($F_{1,1262} = 1016.61$, $P < 0.0001$, also see Fig. S3). This relationship raises the possibility that the explanatory value of PD_C might simply be a result of its correlation to the number of functional groups. However, when we ran a mixed effects model that included a hierarchical effect of the number of functional groups used in an experiment and the effects PD_C, species number and the species number–PD_C interaction, PD_C and the interaction were both significant predictors of plant community biomass (Table 1, Model C). Thus, given that controlling for the effect of the number of functional groups does not diminish the productivity–PD_C relationship, we conclude that existing definitions of functional groups are too coarse and do not correspond well to true functional trait differences in communities (also see ref. 18).

Although the R^2_{pseudo} (see *Materials and Methods*) values from the mixed effects models ranged from 0.181 to 0.211 (Table 1), individual experiments were highly variable in the amount of variation explained. The within-experiment R^2 for the effect of species number on biomass ranged from <0.001 to 0.62, and ranged from 0.01 to 0.69 for PD_C and 0.03 to 0.78 for the full model, including species number, PD_C, and their interaction. Although there are strong overall effects of PD_C and species number on community biomass production, environmental or experimental context seems to be a major factor determining the magnitude of these effects within individual experiments. However, the explanatory power of PD_C appeared unchanged even after we accounted for the fact that a small number of species are

lengths better. We also chose to use the Ultrametric version of the Bayesian tree because present-day taxa are assumed to be equally divergent from the shared ancestor.

Assigning Functional Groups. To assign species to plant functional groups we used the classifications provided by individual researchers for their own experiments (17, 34, 36, 37), which generally organized plants into nitrogen fixers (Fabaceae), woody species, C3 graminoids, C4 graminoids, and nonnitrogen-fixing forbs. To standardize the classification of graminoids, we used a single source (53) to determine C3–C4 status. We then enumerated the number of functional groups within experimental plots.

Statistical Analysis. Our analyses focused on comparing the relative importance of seven potential predictors of LR_{mean} , including species number, the number of plant functional groups, and the five PD_C estimates from different phylogenetic methods. For each predictor we ran single variable mixed effects models of the general form:

$$y_i = \beta_o + \beta_1 x_1 + b_i + \varepsilon_i, \quad [1]$$

where y_i is the LR_{mean} value in a plot in experiment i , β_o is the intercept, β_1 is the coefficient associated with the fixed effect variable x_i (either species number, PD_C , or functional group number), b_i is the coefficient of the random effect (experiment), and the error term, ε_i , is the remaining variation. Parameters in all mixed-effects models were estimated by using restricted likelihood estimation (54). Individual models were compared by using log-likelihood values, Akaike Information Criterion (AIC), and the Akaike weight, which gives the likelihood that model i explains the most variation in an observed data given a set of candidate models (55). We also calculated pseudo- R^2 values by regressing model-fitted response values to the observed response variable. Twenty-one statistical outliers (of 1,315) were excluded from our analyses. These were identified from Bonferroni 2-sided tests on Studentized residuals.

Species number, PD_C , and their interaction were then combined into the single mixed-effects model:

$$y_i = \beta_o + \beta_{sp} x_{sp} + \beta_{PD} x_{PD} + \beta_{sp \times PD} x_{sp} x_{PD} + b_i + \varepsilon_i. \quad [2]$$

With this model we tested the assumption that the predictor variables were fixed effects, as in a standard mixed-effects model, vs. whether heterogeneous results necessitate allowing β estimates to vary among experiments. To test this assumption, we fit models with random effects for both the intercept and the slope estimates for either species number or PD_C nested within experiment and compared these models with the corresponding one where

the intercept and slope were modeled as fixed effects (54). Fixed independent-variable models (i.e., single β) were most parsimonious for both richness and PD_C (likelihood ratios: 2.995, $P = 0.2236$; 4.091, $P = 0.1293$, respectively).

Although we know that functional groups are predictors of LR_{mean} , we also know that functional groupings generally have a phylogenetic signal (28). Also, from the data it is apparent that the number of functional groups in a plot is correlated with the average PD_C in that plot (see Fig. S3). Therefore, because not all experiments explicitly manipulated the number of plant functional groups, we included the number of functional groups, j , as a random effect nested hierarchically within experiment. The model is:

$$y_{ij} = \beta_o + \beta_{sp} x_{sp} + \beta_{PD} x_{PD} + \beta_{sp \times PD} x_{sp} x_{PD} + b_i + b_{ij} + \varepsilon_{ij}, \quad [3]$$

where y_{ij} is the LR_{mean} value in a plot in experiment i with j functional groups, β_o is the intercept, β_{sp} is the slope of the effect of species number, β_{PD} is the slope of the PD_C effect, and $\beta_{sp \times PD}$ is the slope of the effect of the interaction between species number and PD_C . b_i is the coefficient of the random effect of experiment i , b_{ij} is the coefficient of the random effect of j number of functional groups nested with experiment i , and the error term, ε_{ij} , is the remaining variation. Because of the apparent colinearity between predictors, we performed a ridge regression (e.g., 56) on the full model and found no significant change in parameter estimation with an estimated Hoerl–Kennard–Baldwin parameter of 2.315.

Finally, species varied in the number of plots in which they occurred (see Fig. S5), and commonly used species could have a disproportionate effect on the results. Therefore, we created n data subsets corresponding to the n species found in >10% of plots. In each data subset, all plots containing common species n_i were removed. We reran the statistical analyses and compared these results with the full dataset to determine whether species n_i had a disproportionate effect on the results (see SI Text). All analyses were run by using R 2.6.2 (R Development Core Team, www.R-project.org).

ACKNOWLEDGMENTS. We thank T. Jonathan Davies for providing R code that calculates PD_C and for discussions; D. D. Ackerly, T. J. Davies, M. J. Donoghue, E. J. Edwards, A. Hector, S. Kembel, S. Mazer, D. Tilman, and an anonymous reviewer for providing comments on an earlier version of the manuscript; and researchers for supplying raw datasets. This work was supported by National Science Foundation Grant DEB-0614428 (to B.J.C.). A portion of this work was conducted while M.W.C. was a Postdoctoral Associate at the National Center for Ecological Analysis and Synthesis, a Center funded by National Science Foundation Grant DEB-0553768, the University of California, Santa Barbara, and the State of California.

- Sala OE, et al. (2000) Biodiversity: Global biodiversity scenarios for the year 2100. *Science* 287:1770–1774.
- Vitousek PM, Mooney HA, Lubchenco J, Melillo JM (1997) Human domination of Earth's ecosystems. *Science* 277:494–499.
- Chapin SI, et al. (1998) Ecosystem consequences of changing biodiversity. *Bioscience* 48:45–52.
- Hooper DU, et al. (2005) Effects of biodiversity on ecosystem functioning: A consensus of current knowledge. *Ecol Monogr* 75:3–35.
- Tilman D, Downing JA (1994) Biodiversity and stability in grasslands. *Nature* 367:363–365.
- Tilman D, et al. (2001) Diversity and productivity in a long-term grassland experiment. *Science* 294:843–845.
- Tilman D, Wedin D, Knops J (1996) Productivity and sustainability influenced by biodiversity in grassland ecosystems. *Nature* 379:718–720.
- Naeem S, Thompson LJ, Lawler SP, Lawton JH, Woodfin RM (1994) Declining biodiversity can alter the performance of ecosystems. *Nature* 368:734–737.
- Hector A, et al. (1999) Plant diversity and productivity experiments in European grasslands. *Science* 286:1123–1127.
- Cardinale BJ, et al. (2007) Impacts of plant diversity on biomass production increase through time due to complementary resource use: A meta-analysis. *Proc Natl Acad Sci USA* 104:18123–18128.
- Stachowicz J, Bruno JF, Duffy JE (2007) Understanding the effects of marine biodiversity on communities and ecosystems. *Annu Rev Ecol Evol Syst* 38:739–766.
- Cardinale BJ, et al. (2006) Effects of biodiversity on the functioning of trophic groups and ecosystems. *Nature* 443:989–992.
- Balvanera P, et al. (2006) Quantifying the evidence for biodiversity effects on ecosystem functioning and services. *Ecol Lett* 9:1146–1156.
- Schmid B, et al. (2002) The design and analysis of biodiversity experiments. *Biodiversity and Ecosystem Functioning: Synthesis and Perspectives*, eds Loreau M, Naeem S, Inchausti P (Oxford Univ Press, Oxford), pp xii, 294.
- Petchey OL, Hector A, Gaston KJ (2004) How do different measures of functional diversity perform? *Ecology* 85:847–857.
- Felsenstein J (1985) Phylogenies and the comparative method. *Am Nat* 125:1–15.
- Lanta V, Leps J (2006) Effect of functional group richness and species richness in manipulated productivity–diversity studies: A glasshouse pot experiment. *Acta Oecolog Int J Ecol* 29:85–96.
- Wright JP, et al. (2006) Conventional functional classification schemes underestimate the relationship with ecosystem functioning. *Ecol Lett* 9:111–120.
- Venail PA, et al. (2008) Diversity and productivity peak at intermediate dispersal rate in evolving metacommunities. *Nature* 452:210–257.
- Cavender-Bares J, Wilczek A (2003) Integrating micro- and macroevolutionary processes in community ecology. *Ecology* 84:592–597.
- Maherali H, Klironomos JN (2007) Influence of phylogeny on fungal community assembly and ecosystem functioning. *Science* 316:1746–1748.
- Silvertown J, Dodd M, Gowing D, Lawson C, McConway K (2006) Phylogeny and the hierarchical organization of plant diversity. *Ecology* 87:539–549.
- Webb CO (2000) Exploring the phylogenetic structure of ecological communities: An example for rain forest trees. *Am Nat* 156:145–155.
- Valiente-Banuet A, Verdu M (2007) Facilitation can increase the phylogenetic diversity of plant communities. *Ecol Lett* 10:1029–1036.
- Cahill JF, Kembel SW, Lamb EG, Keddy PA (2008) Does phylogenetic relatedness influence the strength of competition among vascular plants? *Perspect Plant Ecol Evol Syst* 10:41–50.
- Stephens PR, Wiens JJ (2004) Convergence, divergence, and homogenization in the ecological structure of emydid turtle communities: The effects of phylogeny and dispersal. *Am Nat* 164:244–254.
- Heemsbergen DA, et al. (2004) Biodiversity effects on soil processes explained by interspecific functional dissimilarity. *Science* 306:1019–1020.
- Edwards EJ, Still CJ, Donoghue MJ (2007) The relevance of phylogeny to studies of global change. *Trends Ecol Evol* 22:243–249.
- Dimitrakopoulos PG, Schmid B (2004) Biodiversity effects increase linearly with biotope space. *Ecol Lett* 7:574–583.
- Dukes JS (2001) Productivity and complementarity in grassland microcosms of varying diversity. *Oikos* 94:468–480.
- Fridley JD (2002) Resource availability dominates and alters the relationship between species diversity and ecosystem productivity in experimental plant communities. *Oecologia* 132:271–277.

32. Fridley JD (2003) Diversity effects on production in different light and fertility environments: An experiment with communities of annual plants. *J Ecol* 91:396–406.
33. Naeem S, Hakansson K, Lawton JH, Crawley MJ, Thompson LJ (1996) Biodiversity and plant productivity in a model assemblage of plant species. *Oikos* 76:259–264.
34. Naeem S, Tjossem SF, Byers D, Bristow C, Li SB (1999) Plant neighborhood diversity and production. *Ecoscience* 6:355–365.
35. Reich PB, et al. (2001) Plant diversity enhances ecosystem responses to elevated CO₂ and nitrogen deposition. *Nature* 410:809–812.
36. Spehn EM, et al. (2005) Ecosystem effects of biodiversity manipulations in European grasslands. *Ecol Monogr* 75:37–63.
37. Tilman D, et al. (1997) The influence of functional diversity and composition on ecosystem processes. *Science* 277:1300–1302.
38. Hector A, Bazeley-White E, Loreau M, Otway S, Schmid B (2002) Overyielding in grassland communities: Testing the sampling effect hypothesis with replicated biodiversity experiments. *Ecol Lett* 5:502–511.
39. Davies TJ, et al. (2004) Darwin's abominable mystery: Insights from a supertree of the angiosperms. *Proc Natl Acad Sci USA* 101:1904–1909.
40. Webb CO, Donoghue MJ (2005) Phylomatic: Tree assembly for applied phylogenetics. *Mol Ecol Notes* 5:181–183.
41. Webb CO, Ackerly DD, Kembel SW (2008) Phylocom: Software for the analysis of phylogenetic community structure and trait evolution. *Bioinformatics* 24:2098–2100.
42. Wikstrom N, Savolainen V, Chase MW (2001) Evolution of the angiosperms: Calibrating the family tree. *Proc R Soc London Ser B* 268:2211–2220.
43. Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Wheeler DL (2005) *GenBank Nucleic Acids Res* 33:D34–D38.
44. Edgar RC (2004) MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792–1797.
45. Posada D, Crandall KA (1998) MODELTEST: Testing the model of DNA substitution. *Bioinformatics* 14:817–818.
46. Posada D, Crandall KA (2001) Selecting the best-fit model of nucleotide substitution. *Syst Biol* 50:580–601.
47. Anisimova M, Gascuel O (2006) Approximate likelihood ratio test for branches: A fast accurate and powerful alternative. *Syst Biol* 55:539–552.
48. Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* 52:696–704.
49. Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogeny. *Bioinformatics* 17:754–755.
50. Ronquist F, Huelsenbeck JP (2003) MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572–1574.
51. Sanderson MJ (1997) A nonparametric approach to estimating divergence times in the absence of rate constancy. *Mol Biol Evol* 14:1218–1231.
52. Faith DP (1992) Conservation evaluation and phylogenetic diversity. *Biol Conserv* 61:1–10.
53. Waller SS, Lewis JK (1979) Occurrence of C3 and C4 photosynthetic pathways in North American grasses. *J Range Man* 32:12–28.
54. Pinheiro JC, Bates DM (2004) *Mixed-Effects Models in S and S-PLUS*. (Springer, New York) p 528.
55. Johnson JB, Omland KS (2004) Model selection in ecology and evolution. *Trends Ecol Evol* 19:101–108.
56. Byrtek M, O'Sullivan F, Muzi M, Spence AM (2005) An adaptation of ridge regression for improved estimation of kinetic model parameters from PET studies. *IEEE Trans Nucl Sci* 52:63–68.