Supporting Information

Bradford et al. 10.1073/pnas.1413707111

SI Materials and Methods

We used three complementary approaches to measure multifunctionality: weighted average (1, 2), single threshold (3), and multiple threshold (4). To calculate the weighted average, values for each separate process were first standardized by their variance so all processes were comparable in magnitude. The average of all of the processes is then the weighted average metric. For the single-threshold approach, we first calculated the maximum observed value for each process by taking the average of the $n + 1$ highest observed values, where n is the smallest sample size of a single treatment (e.g., low complexity, control). We then selected thresholds of this maximum value at which to assess the effect of the treatments. For each single threshold (20 and 80%), we calculated the number of processes that had values greater than or equal to the chosen threshold of the maximum value. For example, with five processes with maximum values of 20, 10, 30, 70, and 100, our response variable at the 20% threshold would be the number of processes that, for a given replicate microcosm, exceed 4, 2, 6, 14, and 20, respectively. If the values for a replicate were then 5, 7, 5, 13, and 11 for the five processes, respectively, our response variable would equal 2 (of 5 maximum). Given that the determination of these thresholds is not standardized and can be arbitrary (3), we also used a multiple thresholds approach. For this approach, we calculated all of the thresholds between 5 and 99% at 1% intervals (4). We then modeled the relationship between the complexity treatment and the multifunctional response for a continuous range of thresholds. This multiple-threshold approach allowed us to evaluate the thresholds at which the treatments began to have a significant impact on the response (4).

To model the relationship between functional complexity, fertilization, and their interaction on multifunctionality, we used linear mixed models (LMMs). We also used LMMs to assess the treatment responses of the five ecosystem processes used to calculate multifunctionality. The sign of net ecosystem productivity was reversed so that positive values reflected greater carbon storage. We then considered greater values of all variables to be optimal, as they would indicate greatest rates of plant production, carbon storage, and nutrient turnover. All response variables fit a Gaussian error distribution. Plot was used as a random effect to account for the repeated measurement (four times) of each ecosystem process following fertilization (5). We included moisture (wetter or drier) and measurement time (cut 1–4) as main effects in the models because we recognized that they could explain variance in our data that might otherwise be included in the community complexity and nitrogen effects. Moisture and time effects were not, however, explored beyond model inclusion because they were not the main focus of the study and we did not have enough degrees of freedom to test how they interacted with complexity or fertilization. Our models contained a mix of categorical and continuous factors, which have different variances, and so we centered categorical data by subtracting the mean and standardized continuous variables by subtracting the mean and dividing by two SDs (6). An advantage of this analytic approach is that centering our independent variables makes main effects interpretable even when involved in interactions (7, 8).

The LMMs were fit using the lme4 package for the R statistical program (9). The F-statistic is not considered accurate for the lme4 package (10), so we used a Markov Chain Monte Carlo approach to estimate coefficients and P values. All reported P values are quasi-Bayesian, rather than the classical frequentist P values, and, like the coefficients, retain the same interpretation. We considered coefficients with $P < 0.05$ significant and coefficients with $P < 0.10$ marginally significant (11). Supporting metadata (e.g., root densities) were also analyzed using LMMs but without time because these metrics were assessed only at the end of the experiment.

1. Maestre FT, et al. (2012) Plant species richness and ecosystem multifunctionality in global drylands. *Science* 335(6065):214–218.
2. Wagg C, Bender SF, Widmer F, van der Heijden MGA (2014) Soil biodiversity and soil community composition determine ecosystem multifunctionality. *Proc Natl Acad Sci USA* 111(14):5266–5270.
3. Zavaleta ES, Pasari JR, Hulvey KB, Tilman GD (2010) Sustaining multiple ecosystem functions in grassland communities requires higher biodiversity. *Proc Natl Acad Sci USA* 107(4):1443–1446.
4. Byrnes JEK, et al. (2014) Investigating the relationship between biodiversity and ecosystem multifunctionality: Challenges and solutions. *Methods Ecol Evol* 5(2): 111–124.
5. Bolker BM, et al. (2009) Generalized linear mixed models: A practical guide for ecology and evolution. *Trends Ecol Evol* 24(3):127–135.
6. Gelman A (2008) Scaling regression inputs by dividing by two standard deviations. *Stat Med* 27(15):2865–2873.
7. Schielzeth H (2010) Simple means to improve the interpretability of regression coefficients. *Methods Ecol Evol* 1(2):103–113.
8. Schielzeth H, Nakagawa S (2013) Nested by design: Model fitting and interpretation in a mixed model era. *Methods Ecol Evol* 4(1):14–24.
9. R Core Team (2012) *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna).
10. Baayen RH, Davidson DJ, Bates DM (2008) Mixed-effects modeling with crossed random effects for subjects and items. *J Mem Lang* 59:390–412.
11. Hurlbert SH, Lomabardi CM (2009) Final collapse of the Neyman-Pearson decision theoretic framework and rise of the neoFisherian. *Ann Zool Fenn* 46(5):311–349.

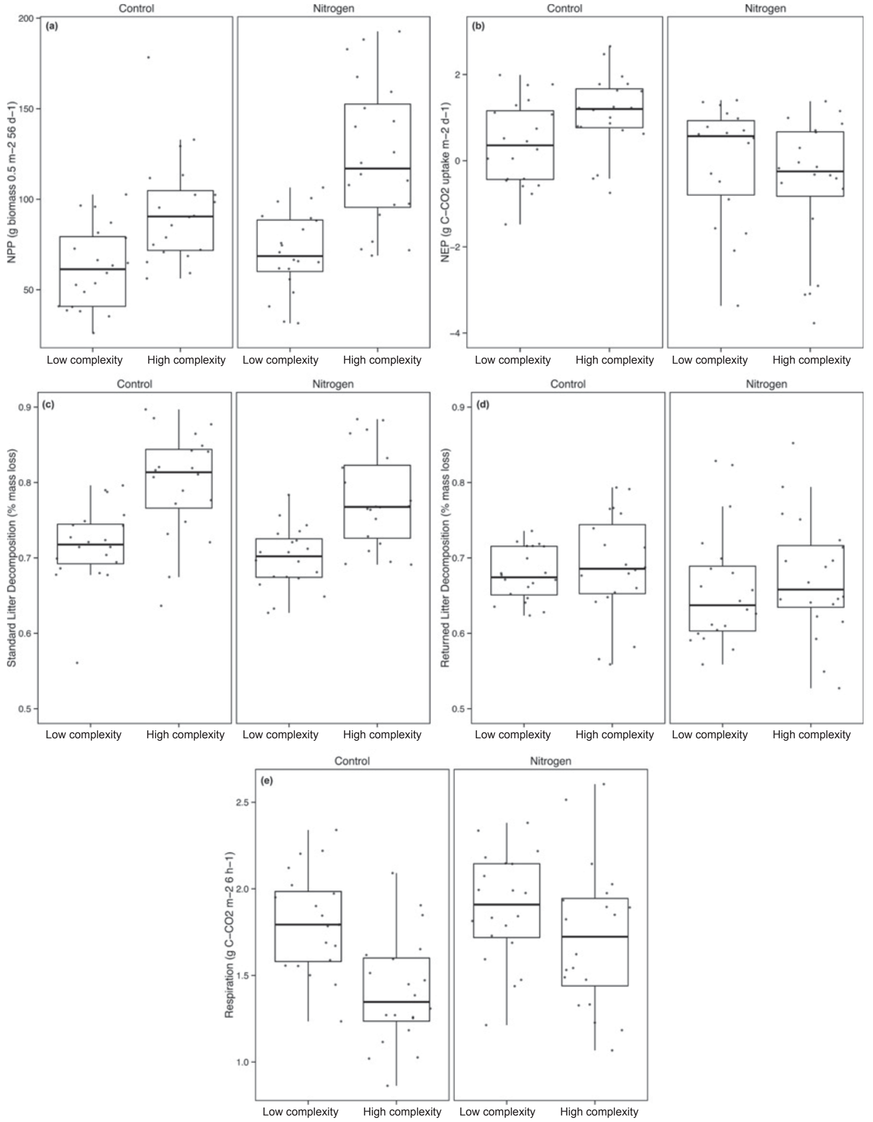

Fig. S1. Effects of soil community functional complexity loss and nitrogen fertilization on the rates of the five biogeochemical processes evaluated: (*A*) net primary productivity (NPP); (*B*) net ecosystem productivity (NEP); (*C*) decomposition of a standard litter; (*D*) decomposition of litter returned within each replicate community; and (*E*) total community respiration. The horizontal line in each box plot shows the median, the box indicates the 25th and 75th percentiles of the data, and the extent of the whiskers 1.5 times the interquartile range. Data points are jittered to visualize vertical spread and represent replicate values for each of the four time points that measures were taken. The statistical significance of the community complexity by nitrogen treatments is given in Table 1.

Table S1. Soil community functional complexity and nitrogen fertilization effects on ecosystem and community properties of the experimental grasslands

	Control		Nitrogen added		Coefficients and significance	
Variables	Low complexity	High complexity	Low complexity	High complexity	Complexity	Nitrogen
Legume biomass (% community)	38.7 ± 10.02	0.1 ± 0.05	33.6 ± 10.23	0.1 ± 0.05	−36*** (51, −20)	NS
Grass nitrogen (% mass)	2.46 ± 0.21	1.98 ± 0.16	2.47 ± 0.24	2.22 ± 0.15	−0.36† (−0.81, 0.05)	NS
Roots—organic (g·m^{-3})	4,679 ± 972	8,282 ± 2196	3,873 ± 1019	3,520 ± 326	NS	−2,467† (−5241, 253)
Roots—mineral (g·m^{-3})	1,058 ± 209	1,828 ± 265	1,103 ± 134	1,257 ± 255	0.34* (0.01, 0.68)	NS
Soil NH_4^+ (μg N g·soil^{-1})	17.5 ± 1.45	15.2 ± 3.04	19.4 ± 2.44	28.0 ± 6.38	NS	7.96* (0.58,16.42)
Soil NO_3^- (μg N g·soil^{-1})	5.0 ± 0.64	4.0 ± 0.96	10.1 ± 2.75	8.6 ± 1.96	NS	0.61** (0.21,1.01)

Shown are attributes of the aboveground plant community (legume and grass), root densities, and estimates of plant-available soil nitrogen. Values are means ± SE ($n = 5$) at the end of the experiment, and the final two columns show the coefficients with statistical significance from the linear models with the lower and upper bounds shown in parentheses. Interactions between the community complexity and nitrogen fertilization effects are not shown because none were significant. $^{\dagger}P < 0.1$, $*P < 0.05$, $**P < 0.01$, $***P < 0.001$; NS, not significant.