Why does grassland productivity increase with species richness? Disentangling species richness and composition with tests for overyielding and superyielding in biodiversity experiments

John M. Drake

Department of Biological Sciences, University of Notre Dame, Notre Dame, IN 46556, USA (drake.4@nd.edu)

The causal relationship between the biodiversity of natural and modified environments and their net primary production has been a topic of significant scientific controversy and scrutiny. Early theoretical and empirical results indicated that production was sometimes significantly correlated with species richness when species richness was directly manipulated in experimental systems. Possible mechanisms for this phenomenon include statistical sampling effects, complementary resource use and mutualistic interactions. However, the interpretation of experimental results has sometimes confounded species richness with species composition, and disentangling the effects of species diversity from species identity has proved a formidable challenge. Here, I present a statistical method that is based on simple probability models and does not rely on the species composition of individual plots to distinguish among three phenomena that occur in biodiversity-production experiments: underyielding, overyielding and (a new concept) superyielding. In some cases, distinguishing these phenomena will provide evidence for underlying mechanisms. As a proof-of-concept, I first applied this technique to a simulated dataset, indicating the strengths of the method with both clear and ambiguous cases. I then analysed data from the BIODEPTH experimental biodiversity manipulations. No evidence of either overyielding or superyielding was detected in the BIODEPTH experiment.

Keywords: complementarity; overyielding; primary production; sampling effect; species richness; superyielding

1. INTRODUCTION

Several recent experiments tested the hypothesis that increasing the number of grassland plant species causes an increase in total net primary production (Tilman *et al.* 1996; Hector *et al.* 1999). This hypothesis is a special case of the more general hypothesis that changes in species richness have measurable effects on ecosystem processes (Mooney *et al.* 1993; Naeem *et al.* 1995; Kinzig *et al.* 2001). Much of the current debate about biodiversity and ecosystem processes (or functions) is about the mechanisms responsible for these hypothesized effects. Complementarity explanations suppose that different species use available resources in different ways so that more diverse assemblages more fully exploit available resources than less diverse assemblages (Loreau 1998; Tilman 1999). Alternative explanations, based on sampling and selection effects, hypothesize that more speciose assemblages, which are drawn from a pool of species that differ in their metabolic and photosynthetic rates, are more likely than species-poor assemblages to contain species capable of high rates of production (Aarsen 1997; Huston 1997; Tilman *et al.* 1997).

2. A CONCEPTUAL ANALYSIS OF YIELD MEASURES

Standard theory (e.g. Kinzig *et al.* 2001) characterizes two types of primary production phenomena. 'Overyielding' is a phenomenological concept borrowed from the theory of agricultural intercropping (Vandermeer 1989). It occurs when the total production of a mixed plot of two or more species exceeds the production that would have been obtained by growing the species in monoculture. In other words, overyielding occurs when the addition of species to a plot results in an increase in total net primary production. 'Underyielding' occurs in assemblages that do not exhibit an increase in total primary production with additional species. In addition to these phenomenological concepts, current theory supposes that the primary mechanisms that determine the total amount of production are (i) sampling and selection effects, and (ii) interspecific complementarity. Sampling effects occur because species-rich plots are more likely than species-poor plots to contain highly productive species owing to greater sampling of the regional species pool. Complementarity occurs when resource use has been optimized by niche differentiation and/or facilitation. An important goal of theory is to provide tests to distinguish these mechanisms and the way that they produce the observed phenomena.

Loreau & Hector (2001) argue that because every plot in an experiment is a sample from the regional species pool, the sampling process is a characteristic of both classes of mechanism. According to this version of the theory, the substantive component of the sampling effect hypothesis is not chance selection, but selection for dominant species, which they properly point out is a degenerate null hypothesis: plots need not contain a single dominant species, nor must numerical dominance correlate with production. Thus, they believe that the challenge is to

develop a model to partition the effects of selection and complementarity, assuming simple additivity (Loreau & Hector 2001). However, according to this model, over-yielding (positive deviation from the expected yield) can be attributed to selection effects, complementarity or both. By contrast, I suggest that overyielding can only result from niche differentiation (complementarity). Hence, in my view the sampling effect is a meaningful hypothesis in its own right, consistent with the original derivations by Aarsen (1997), Huston (1997) and Tilman *et al.* (1997). It follows from this definition that overyield-ing cannot be accomplished by the sampling effect alone because plots that exhibit overyielding exceed the pro-duction of the individual species. Further, according to this theory, evidence of overyielding constitutes *prima facie* evidence of complementarity. (Note that this does not imply the contrary proposition that niche differentiation has not occurred unless the plot exhibits overyielding.)

Possibly, part of the reason for this conceptual ambi-guity is that overyielding and underyielding are phenom-enological concepts, whereas the sampling effect hypo-thesis and complementarity pertain to the mechanisms of production. Furthermore, the relationship between the phenomena and the mechanisms remains vague because overyielding is explicitly defined by comparison with monocultures and there are several possible ways in which these comparisons could be made (see Hector *et al.* 2002). Currently, two concepts of overyielding that have been derived from the intercropping literature are in common use (Fridley 2001; Hector *et al.* 2002). These should be considered as possible tests for complementarity.

(i) Non-transgressive overyielding occurs when the yield of a mixture is greater than expected based on a weighted average of the monoculture yields of the component species.

(ii) Transgressive overyielding occurs when a mixture yields more than any monoculture of the compo-nent species.

Definition (i) is not intuitively satisfying because even if no niche differentiation has occurred (an appropriate null hypothesis), fully half of experimental plots would be expected to exceed the weighted average of monoculture yields based on random variation and sampling error alone, irrespective of the underlying mechanism. By con-trast, definition (ii) seems unduly restrictive because dominant species are not always the most productive indi-vidually: the species that yields most in monoculture may be unimportant in mixed assemblages (Loreau & Hector 2001). A further objection is that both of these concepts are defined with reference to the specific composition of the plot in question, a fact that has important conse-quences for the interpretation of experimental results: an experiment that seeks to relate species richness to over-yielding according to one of these definitions is con-founded by the effect of species composition. An alternative test that does not depend on plot composition, such as I provide in this paper, would be preferable.

To develop such a test, first consider the intuitions motivating these concepts. Fridley (2001) noted that the sampling effect and complementarity mechanisms are hierarchically related, although not mutually exclusive:

failing to reject the sampling effect hypothesis does not mean that complementary resource use has not occurred, only that niche differentiation is not required to explain the observed production. Obviously, there is an upper limit to the production explainable by the sampling effect (i.e. the production of the most productive species in the plot), and any production that exceeds this exhibits over-yielding by definition. This fact will be exploited to develop a test for overyielding. First, however, I take this rationale a step further to suggest that there is also an upper limit to production obtainable by additive niche dif-ferentiation. This upper limit is equivalent to the simple addition of the potential productions of each species grown in monoculture. This is an extreme upper limit and could only occur when niche differentiation is complete, i.e. when there is no overlap in resource use or when facili-tation is so great between some species that it masks effects of competition among others. I call any production that exceeds this level 'superyielding'. Because the over-yielding upper limit is based on the simple addition of the potential production of each species, evidence of super-yielding is *prima facie* evidence of synergistic, positive non-linear interactions among species, namely mutualism.

Thus, define three phenomenological categories, such that $Y_0 < Y_1 < Y_2$: Y_0, underyielding; Y_1, overyielding; and Y_2, superyielding. Jointly, these categories exhaust the possible measures of production and correspond to three hierarchically nested hypotheses such that $H_0 \subset H_1 \subset H_2$. H_0, observed production is explainable by the sampling effect alone; H_1, observed production is not explainable by the sampling effect; the plot exhibits complementarity and/or mutualism; and H_2, observed production is not explainable by complementarity; the plot exhibits mutual-ism.

According to this conceptual framework, the relation-ship between mutualism and complementarity is anal-ogous to the relationship between complementarity and the sampling effect. This paper suggests a test for dis-tinguishing these phenomena (Y_0, Y_1, Y_2) in data from experimental and natural systems. Distinguishing these phenomena allows us to make some inferences about the underlying causal mechanisms (H_0, H_1, H_2). All three mechanisms are sufficient to explain underyielding (Y_0). Complementarity (H_1) is the minimally sufficient mech-anism accounting for overyielding (Y_1). Mutualism (H_2) is the minimally sufficient mechanism accounting for superyielding (Y_2). Finally, the test proposed here is not a test for pairwise complementarity or mutualism among species. Rather, this test proposes to detect different assemblage-level mechanisms that have been adduced to explain increasing primary production with species rich-ness in biodiversity manipulation experiments.

3. TEST STATISTICS

The goal of this paper is to develop a technique that tests for overyielding and superyielding in a way in which these concepts are not explicitly dependent on the specific assemblage of species in a plot. Therefore, consider a set of observed sampling plots, which may be experimental units or arbitrarily chosen plots in a natural system, for which only two data are known: the species richness of the plot and an estimate of total production for the year

(e.g. biomass). First, these data must be related to the possible production of each species individually for all species in the experiment or in the regional species pool. These possible individual productions are the monoculture productions of the set of species: call the probability density function $f(x)$ of this set, the 'potential production distribution'. Experimentally, the potential production distribution is estimated from the frequency histogram of observed measures of productivity or from sample statistics for parametric distributions. Treating production probabilistically ensures that the effects of species richness and the specific composition of the plot can be separated, distinguishing the technique proposed here from other procedures. Integrating $f(x)$ obtains the cumulative distribution function (CDF) $F(x)$, which gives the probability that an observed measure of production (yield) is less than that expected by chance sampling from the species pool represented by the potential production distribution. This distribution will be crucial to detecting overyielding and superyielding in biodiversity experiments. To use this distribution in the familiar statistical way, each species must have an equal probability of arriving at any given plot, an assumption met by the randomized experiments for which this test is designed. (Where this assumption is not met, modifications can be made: see § 6.) Given a plot containing k species and an observed measure X units of production, rejecting the null hypothesis that the observed biomass X can be explained by the sampling effect (at the α level) is equivalent to saying that the probability that X is exceeded by any one of k independent and identically distributed (IID) random variables is less than α. More formally,

$$P\{X < \text{at least one of } k \text{ IID random variables}\}$$
$$= 1 - P\{X \text{ is less than } x \text{ in each of } k \text{ draws}\}$$
$$= 1 - [P\{X_1 < x\}\&P\{X_2 < x\}\&...P\{X_k < x\}]$$
$$= 1 - [P\{X_1 < x\}P\{X_2 < x\}...P\{X_k < x\}]$$
$$= 1 - [P\{X < x\}]^k \text{ by independence.} \quad (3.1)$$

Hence, reject the null hypothesis H_0 that the sampling effect can explain the observed production X if $[P\{X < x\}]^k > (1 - \alpha)$. This completes the first step. If H_0 is rejected, then the sampling effect cannot account for the observed productivity.

The second step in the testing process is to adopt complementarity as our null hypothesis and attempt to reject overyielding in favour of superyielding. Here, we exploit the fact that superyielding is defined to occur where X exceeds the sum of k IID random variables drawn from the potential production distribution, i.e. the 'k-fold convolution distribution', denoted $F_k(x)$ (Olkin *et al.* 1994). Hence, the probability that the observed biomass X can be explained by overyielding is

$$P\{X \text{ is less than the sum of each set of } k \text{ draws}\}$$
$$= [P\{X < y\}], \quad (3.2)$$

where y is the quantile of the $F_k(x)$ distribution corresponding to probability $(1 - \alpha)$. Thus, rejecting the null hypothesis that complementarity can explain the observed production X amounts to demonstrating that $[P\{X < y\}] > (1 - \alpha)$. Summarizing, the relationship of these three hypotheses can be envisaged as a series of overlapping test distributions (figure 1) where distribution (*a*) is the CDF of the potential production distribution $f(x)$,
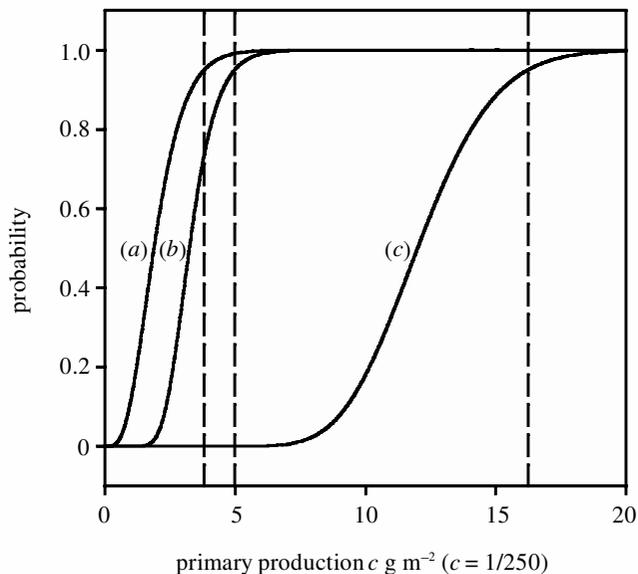

Figure 1. Test distributions illustrate the hierarchical relationship among hypotheses for a specific plot from the BIODEPTH experiment, which contained six species. Distribution (*a*) is the CDF of the potential production distribution ($x \sim \Gamma$:0.45,4.5); (*b*) is the test distribution for overyielding, $F(x)^6$; and (*c*) is the test distribution for superyielding, $F_6(x)$. Dashed lines are the $(1 - \alpha)$ test values.

distribution (*b*) is the test distribution for overyielding $F(x)^k$, and distribution (*c*) is the test distribution for superyielding $F_k(x)$.

These statistics can be calculated as follows. First, the potential production distribution $f(x)$ must be specified. Theoretically, as these models are quite general, any distribution that fits the observed frequency of monoculture productions can be used. In practice, the convolution of the distribution will not always be known and numerical methods may be required. (A cumulative frequency histogram can be used in lieu of the potential production distribution and an appropriate convolution distribution can be numerically simulated.) After $F(x)$ and $F_k(x)$ have been specified (Olkin *et al.* 1994), calculating equations (3.1) and (3.2) is straightforward.

4. METHODS

(a) *Hypothetical example*

As an example, hypothetical data were simulated to demonstrate the way in which this approach distinguishes different cases. The example is intended to mimic a marsh ecosystem in which a subset of species ('facilitators') is important for stabilizing the substrate, and species vary widely in their potential productions. Here, the regional species pool consisted of five facilitator species and 25 dependent species. The potential productions of these species were independently drawn from a gamma distribution with shape parameter $a = 2$ and scale parameter $b = 1.5$. Three hundred random assemblages were constructed, ranging in richness from 1 to 30 species. The example is built up by considering four scenarios.

(i) *Scenario 1*

All species are independent and no facilitation occurs (facilitator species were excluded). Here, observed yield is the simple addition of the potential productions of all species p_i:

$$Y_{obs} = Y_{add} = \sum_i p_i. \tag{4.1}$$

(ii) *Scenario 2*

Species compete for a limiting resource so that maximum yield (Y_{max}) is approached asymptotically as species richness increases. In our hypothetical example, observed yield is given by

$$Y_{obs} = \gamma Y_{add}, \tag{4.2}$$

where γ is a coefficient describing the average effect of competition, defined by

$$\gamma = Y_{max}/(Y_{max} + Y_{add}). \tag{4.3}$$

Although maximum yield will depend on the total available resources and the way they are partitioned among species in an actual assemblage, in my example I assign Y_{max} to the maximum potential production of any species in the regional species pool ($Y_{max} = 7.6$).

(iii) *Scenario 3*

No competition occurs among species and facilitation occurs in proportion to the number of facilitators present. Specifically,

$$Y_{obs} = \delta^n Y_{add}, \tag{4.4}$$

where δ is a parameter that describes the relative strength of the average facultative relationship and n is the number of facilitators present. In the hypothetical example $\delta = 1.8$ and n varies from 0 to 5.

(iv) *Scenario 4*

The outcome of each of the previous simulations is easily anticipated. Scenario 1 will usually exhibit overyielding, but not superyielding; scenario 2 will not exhibit either; and scenario 3 will usually exhibit superyielding. A fourth scenario, however, is ambiguous. In this case, yield is determined by both facilitation and competition among species. Here, observed yield is given by

$$Y_{obs} = \gamma \delta^n Y_{add}, \tag{4.5}$$

with γ and δ as above. Note that in this case the total yield of the assemblage is not limited by the maximum monoculture yield (as in scenario 2), as species cultivated in monoculture are not facilitated by definition.

(b) *BIODEPTH experiment*

To demonstrate my method with a real example, data were obtained from the Biodiversity and Ecological Processes in Terrestrial Herbaceous Ecosystems (BIODEPTH) experimental manipulations (A. Hector, personal communication), which were previously reported to exhibit complementarity (Hector *et al.* 1999), but the interpretation of which has been challenged (Huston *et al.* 2000). These experiments were conducted in eight different European locations: Bayreuth, Germany; Riverstick, Ireland; Silwood, UK; Sheffield, UK; Lupsingen, Switzerland; Lezirias, Portugal; Ulmea, Sweden; and Mytilini, Greece. The entire experiment consisted of 480 plots containing 200 different plant assemblages. Primary production was estimated from annual above-ground biomass and ranged from 0.25 g m^{-2} to 1435.94 g m^{-2}. Plot species richness ranged from 1 to 32 species. Plots containing species for which monoculture productions are unknown were excluded from this analysis. The set of plots meeting this requirement is the same dataset analysed by

Loreau & Hector (2001). The range of biomass for the restricted dataset is 3.51 g m^{-2} to 1401.58 g m^{-2}; the range of species richnesses is 1 to 14. Eighty-five monoculture plots were used to estimate the primary production distributions and 201 mixed species plots were tested for overyielding and superyielding.

Because yield differed by location (one-way ANOVA $F_{7,278} = 17.59$, $p < 0.0001$), all subsequent analyses were conducted independently for each site. After linear transformation (to reduce computational complexity) gamma distributions were fitted to the primary production histograms of all sites from sample means and variances (Hastings & Peacock 1975). The k-fold convolution of the gamma distribution with shape parameter a and scale parameter b is given by the gamma distribution with shape parameter ka and scale parameter b (Olkin *et al.* 1994). These estimated distributions were used to test for overyielding and superyielding precisely as defined by equations (3.1) and (3.2) with $\alpha = 0.05$.

Additionally, the original investigators further classified their plots as containing one, two or three functional groups depending on the combination of grasses, legumes and other herbs present. Therefore, I also investigated whether or not functional group richness was associated with the fraction of plots that exhibited overyielding.

5. RESULTS

(a) *Hypothetical example*

Primary production increased with species richness in all four hypothetical scenarios and the relationship among scenarios was consistent with the anticipated result: scenario 3 was generally the most productive, scenario 2 was usually the least productive, and scenarios 1 and 4 fell in between (figure 2). It is not possible, however, to determine whether or not competition and facilitation have occurred simply by comparing these results with Y_{max}. Therefore, supposing that the potential production distribution was unknown, a gamma distribution was fitted to the simulated data and the tests for overyielding and superyielding were applied accordingly. Consistent with the predicted results, scenario 1 often exhibited overyielding, but rarely exhibited superyielding; scenario 2 never exhibited overyielding (thus not superyielding either); and scenario 3 usually exhibited overyielding and superyielding. The ambiguous case, scenario 4, usually exhibited overyielding and less commonly exhibited superyielding (table 1).

(b) *BIODEPTH experiment*

No evidence of superyielding was detected in the BIODEPTH experiment. For experimental richnesses between 2 and 14 species, overyielding occurred in 9 out of 201 plots (4.5%) (figure 3*a*). As species richness increased, the number of plots exhibiting evidence of overyielding decreased (note the small sample size where species richness is three). This is reasonable because more speciose communities are expected to be increasingly saturated, and additional species are therefore more likely to be functionally redundant. Further, only those plots that contained two or three (not one) functional groups exhibit any evidence of possible overyielding (figure 3*b*). Thus, although there is evidence of overyielding in some plots independently, for $\alpha = 0.05$ we fail to reject the null hypothesis of overyielding in the experiment as a whole.
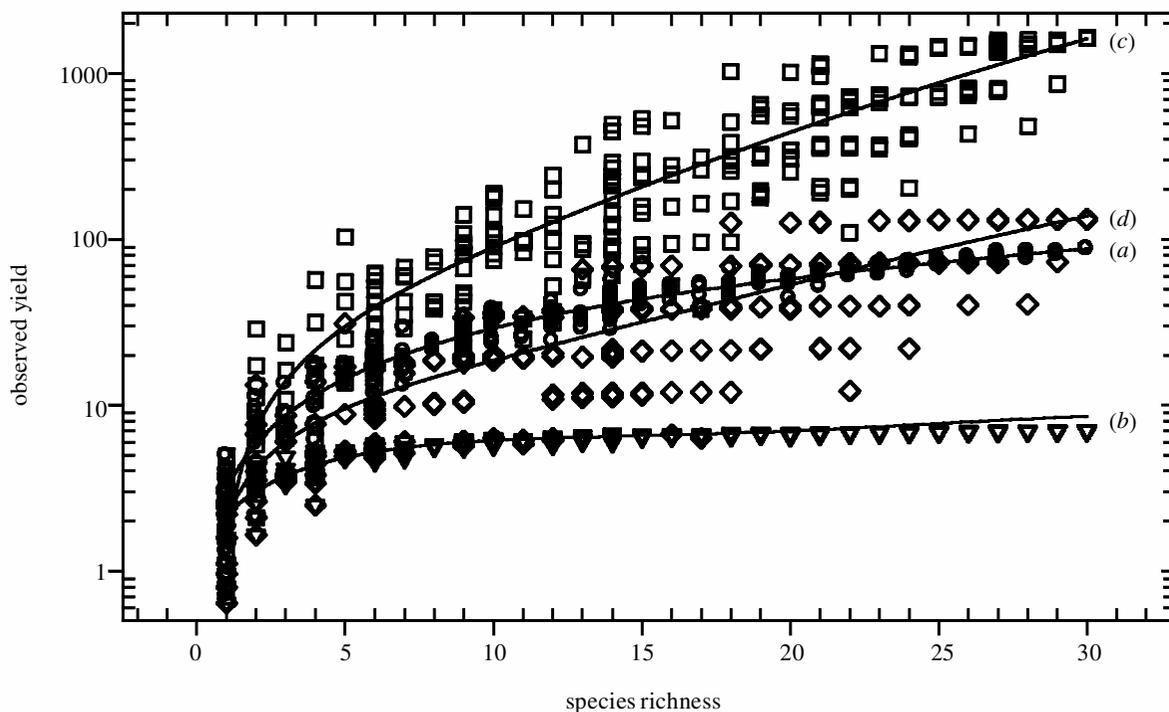
Figure 2. Primary production compared with species richness for simulated datasets consisting of a species drawn from a pool of: (*a*) independent species (circles) line; (*b*) competitive species (triangles) line; (*c*) mutualistic species (squares) line; or (*d*) mixed competitive and mutualistic species (diamonds) line.

Table 1. The fraction of plots exhibiting overyielding and superyielding in a hypothetical example where yield is (i) independent, (ii) competitive, (iii) mutualistic or (iv) ambiguous.

	overyielding (%)	superyielding (%)
(i) scenario 1	91.7	2.0
(ii) scenario 2	0	0
(iii) scenario 3	92.0	82.0
(iv) scenario 4	80.0	21.7

6. DISCUSSION

The above result, which comes from a reasonably liberal test for complementarity, confirms the conclusion of Huston *et al.* (2000); that there is no evidence for overyielding in the BIODEPTH experiments at the $\alpha = 0.05$ level. Given the independence of species composition and the test for overyielding in this analysis, results of the BIODEPTH experiment suggest that species richness *per se* may be relatively unimportant to production. Rather, the result supports the alternative speculation that plot composition, in contrast to merely richness, is crucial for understanding observed yields. Because the assumption of this test is that overyielding should be understood as an effect of strictly independent potential productions, the upshot of this is that the method presented here constitutes the most liberal test possible that does not confound species richness with composition, in the sense that any test that is more liberal (rejects the null hypothesis in fewer cases) must posit a null hypothesis in which the yields of different species are known not to be independent. The assumption that individual species' productions are not independent is justified only when the community is

known to behave in a certain way, behaviour that is determined by that community's particular composition (i.e. whether it consists primarily of competitors, mutualists, etc.). Other currently available tests for complementarity, for example the concepts of transgressive overyielding and non-transgressive overyielding defined above, cannot detect effects of richness *per se*. Note, however, this test is not intended to replace the model of Loreau & Hector (2001), which partitions the determinants of production in particular plots, but complements it by providing a test suitable to the experiment as a whole and can reject the null hypothesis that production results from chance sampling (sampling effects).

Modification of the theory is possible where propagule dispersal is an important factor for explaining total production. In this case, the assumption that each species is equally likely to arrive at a given plot is violated. (In experimental systems this problem is overcome by distributing equal numbers of colonizing propagules for each species.) In unmanipulated environments, the solution to this problem is to modify $f(x)$ by the distribution of propagule pressures $g(x)$ so that $h(x) = f(x)g(x)$ is actual potential production and not just the frequency distribution of monoculture production, because actual potential production is the product of the potential production and the relative probability of propagule arrival. One such situation would occur when species that produce more propagules are more likely to colonize any given site. In the special case where the propagule pressure is directly proportional to potential production the actual potential production $h(x) = (f(x))^2$. Alternatively, one could use percentage cover by species or some other factor as a surrogate for relative propagule pressure, though experiments or observational data would be required to justify using such a procedure.
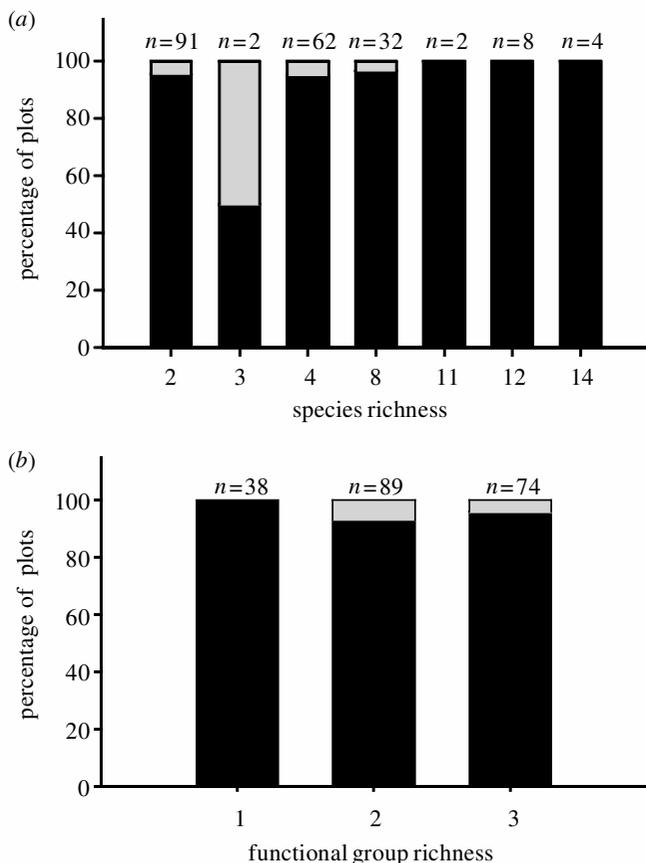
Figure 3. Tests for overyielding in the BIODEPTH experiment exhibit possible overyielding in a small minority of cases by (*a*) species richness and (*b*) functional group. There was no evidence of superyielding. Black bars, no evidence of overyielding; grey bars, evidence of overyielding.

Though demanding, in that the distribution of monoculture productions must be known, the method presented here is extremely general and may be applied to a variety of ecological problems including the theory of intercropping in agroecosystems, where the lack of statistical techniques for addressing complex combinations of multiple crops retards research progress (Vandermeer 1989; Federer 1993, 1998). Furthermore, as noted by Loreau & Hector (2001), any variable that can be integrated over a community can replace primary production, including most measures of important ecosystem process rates (e.g. decomposition, respiration, nutrient cycling). Additionally, although most studies to date have relied on randomized manipulations of species diversity to detect relationships between diversity and ecosystem processes, this is not required here. Indeed, even observational studies of natural systems could plausibly be analysed in this manner, provided that physical conditions are relatively homogeneous (i.e. variation in production among plots is not due to spatial heterogeneity among environmental variables) and either natural or cultivated monocultures are available for determining the potential production distribution.

Finally, a far-reaching application of the test for superyielding is to detect mutualisms by replacing primary production with a measure of population fitness (e.g. intrinsic rate of increase). This is particularly relevant to the current debate about biodiversity and ecosystem processes as

the study of mutualisms intersects the analysis of species diversity and production (Cardinale *et al.* 2002). Specifically, mutualisms that cause the structural modification of environments would be expected to create conditions conducive to superyielding if a significant proportion of production is contributed by epiphytic species (e.g. epiphyton and invertebrates in aquatic macrophyte beds and coral reef communities) or by species that make use of structural modifications for habitat (e.g. canopy animals in forest ecosystems and interstitial invertebrates). These situations are most likely to occur in physiologically stressful environments (Callaway *et al.* 2002) or where some species act as ecosystem engineers (Jones *et al.* 1997). As with the analysis of production, this constitutes the most liberal test possible for detecting a relationship between mutualism and species richness without confounding species richness with species composition.

REFERENCES

Aarsen, L. W. 1997 High productivity in grassland ecosystems: effected by species diversity or productive species? *Oikos* **80**, 183–184.

Callaway, R. M. (and 12 others) 2002 Positive interactions among alpine plants increase with stress. *Nature* **417**, 844–848.

Cardinale, B. J., Palmer, M. A. & Collins, S. L. 2002 Species diversity enhances ecosystem functioning through interspecific facilitation. *Nature* **415**, 426–429.

Federer, W. T. 1993 *Statistical design and analysis for intercropping experiments, vol. I: two crops.* New York: Springer.

Federer, W. T. 1998 *Statistical design and analysis for intercropping experiments, vol. II: three or more crops.* New York: Springer.

Fridley, J. D. 2001 The influence of species diversity on ecosystem productivity: how, where, and why? *Oikos* **93**, 514–526.

Hastings, N. A. J. & Peacock, J. B. 1975 *Statistical distributions.* New York: Wiley.

Hector, A. (and 33 others) 1999 Plant diversity and productivity experiments in European grasslands. *Science* **286**, 1123–1127.

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.

Huston, M. A. 1997 Hidden treatments in ecological experiments: re-evaluating the ecosystem function of biodiversity. *Oecologia* **110**, 449–460.

Huston, M. A. (and 11 others) 2000 No consistent effect of plant diversity on productivity (Technical comment). *Science* **289**, 1255a.

Jones, C. G., Lawton, J. H. & Shachak, M. 1997 Positive and negative effects of organisms as physical ecosystem engineers. *Ecology* **78**, 1946–1957.

Kinzig, A. P., Pacala, S. W. & Tilman, D. (eds) 2001 *The functional consequences of biodiversity.* Princeton University Press.

Loreau, M. 1998 Biodiversity and ecosystem functioning: a mechanistic model. *Proc. Natl Acad. Sci. USA* **95**, 5632–5636.

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

Mooney, H. A., Cushman, J. H., Medina, E., Sala, O. E. & Schulze, E.-D. (eds) 1993 *Functional roles of biodiversity*. Chichester, UK: Wiley.

Naeem, S., Thompson, L. J., Lawler, S. P., Lawton, J. H. & Woodfin, R. M. 1995 Empirical evidence that declining species diversity may alter the performance of terrestrial ecosystems. *Proc. R. Soc. Lond.* B **347**, 249–262.

Olkin, I., Gleser, L. J. & Derman, C. 1994 *Probability models and applications*, 2nd edn. New York: Macmillan.

Tilman, D. 1999 Ecological consequences of changes in biodiversity: a search for general principles. *Ecology* **80**, 1455–1474.

Tilman, D., Wedin, D. & Knops, J. 1996 Productivity and sustainability influenced by biodiversity in grassland ecosystems. *Nature* **379**, 718–720.

Tilman, D., Lehman, C. L. & Thompson, K. T. 1997 Plant diversity and ecosystem productivity: theoretical considerations. *Proc. Natl Acad. Sci. USA* **94**, 1857–1861.

Vandermeer, J. 1989 *The ecology of intercropping*. Cambridge University Press.