

Biodiversity manipulation experiments: studies replicated at multiple sites

A. Hector, M. Loreau, B. Schmid, and the BIODEPTH project¹

4.1 Introduction

Biodiversity is usually defined in a general sense as a collective term for all biological differences at scales ranging from genes to ecosystems (e.g. Harper and Hawksworth 1994). This breadth of meaning is both a strength and weakness of the term, which poses many challenges in its study. In this chapter, we focus on the BIODEPTH project, presenting an overview of published analyses of the combined site datasets from Hector *et al.* (1999) and Loreau and Hector (2001), and re-examining the debate caused by the early results (Hector *et al.* 1999, 2000b; Huston *et al.* 2000). We present some additional material and new methods and analysis that we hope will clarify some points of debate and draw comparisons with the small number of other multi-site biodiversity experiments that have been conducted so far. Many of the issues we discuss are not specific to biodiversity studies but apply widely in ecology and should therefore be of broad interest. These can be classified into two broad groups:

1. The importance of specifying precise, testable hypotheses and identifying causality when examining complex explanatory and response variables,

in this case biodiversity and ecosystem processes. This is particularly important where explanatory variables have multiple co-varying components, as for the many compositional and richness aspects of diversity.

2. Identifying the effects of extrinsic factors (e.g. climate) that vary and act at larger scales, and between systems, from intrinsic factors (including community processes) that act at a local scale within systems.

4.2 An overview of the BIODEPTH experiment and results

We present only a brief overview of the combined site analyses here as the BIODEPTH project is described extensively in many other publications: the overall project and analysis is described in Hector *et al.* (1999) and the details of the individual site designs can be found in the following papers for Switzerland (Diemer *et al.* 1997; Joshi *et al.* 2000; Koricheva *et al.* 2000; Spehn *et al.* 2000a,b; Stephan *et al.* 2000; Diemer and Schmid 2001), Greece (Troumbis *et al.* 2000), Sweden (Mulder *et al.* 1999; Koricheva *et al.* 2000), Silwood (Hector *et al.* 2000a, 2001a,b), Portugal (Caldeira *et al.* 2001), Germany (Scherer-Lorenzen 1999) and for the theoretical modelling component (Loreau 1998a, 2000a; Loreau and Hector 2001). While our analysis of combined datasets from all sites has concentrated on above-ground biomass production, these papers examine the effects of plant diversity on other processes at individual sites including: above-ground

¹ C. Beierkuhnlein, M. C. Caldeira, M. Diemer, P. G. Dimitrakopoulos, J. A. Finn, H. Freitas, P. S. Giller, J. Good, R. Harris, P. Höglberg, K. Huss-Danell, J. Joshi, A. Jumpponen, C. Körner, P. W. Leadley, A. Minns, C. P. H. Mulder, G. O'Donovan, S. J. Otway, J. S. Pereira, A. Prinz, D. J. Read, M. Scherer-Lorenzen, E-D. Schulze, A-S. D. Siamantziouras, E. M. Spehn, A. C. Terry, A. Y. Troumbis, F. I. Woodward, S. Yachi, and J. H. Lawton.

space-filling and canopy structure; root production; decomposition and nutrient cycling; soil biodiversity, activity and other below-ground processes; community water use efficiency, invasion resistance and the diversity and abundance of insects and other invertebrates.

A major aim of the BIODDEPTH project was to explicitly test for statistically significant differences between locations by combining data from all sites in single analyses. To date, our published multisite analyses have largely examined above-ground biomass production (but see Joshi *et al.* 2001; Spehn *et al.* 2002). To summarize, the BIODDEPTH project replicated the same basic experiment at eight different locations around Europe. At each site, we established experimental plant communities from seed so that we could control numbers and types of species and functional groups (species predicted to have similar effects on processes). The primary aim was to generate a gradient of species richness at each location with five levels ranging from monocultures to numbers of species estimated to be found in 4 m²

plots in unmanipulated grasslands at each site. Each level of species richness was replicated with several plant communities with different compositions, and each particular species combination was also replicated (see also Bell 1990; McGrady-Steed *et al.* 1997; Petchey *et al.* 1999; McGrady-Steed and Morin 2000) in order to separate these two components of diversity (Givnish 1994). Species were chosen from the pool of co-occurring species at each site at random with certain constraints (detailed in the above publications). Primary amongst these was constraining the functional composition of the communities so that we could also examine this aspect of diversity, at least in part. We then used standardized protocols to examine responses in a suite of ecosystem and community processes, including above-ground biomass production on which we concentrate here.

As we examined the productivity-diversity relationships emerging at individual sites, it was clear that there were differences between sites (Fig. 4.1(b), Table 4.1; Loreau *et al.* 2001); we return to these

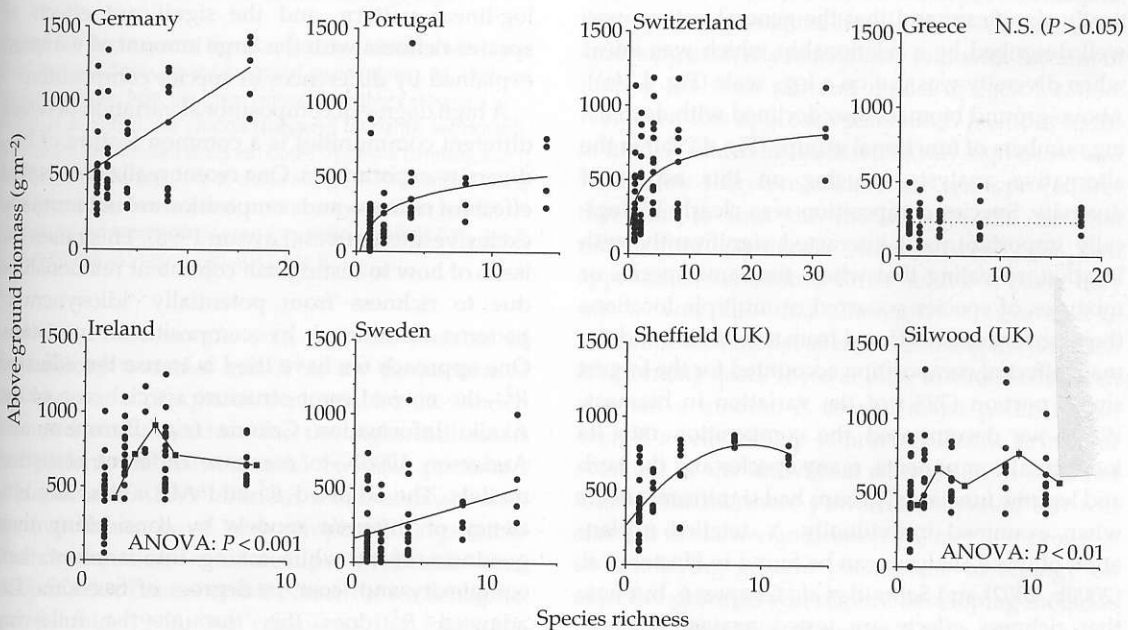


Figure 4.1 The different relationships between biodiversity and above-ground biomass production found in the analyses of individual site patterns when the highest adjusted R^2 criterion was used to select chosen models. Points are individual plot values and lines are slopes from linear (Germany, Sweden), log-linear (Portugal, Switzerland) or quadratic (Sheffield) models, join richness level means (solid squares with SEMs) for significant ANOVA (Ireland, Silwood), or show non-significant relationships (Greece, broken line).

Table 4.1 Summary of individual site analyses comparing major diversity-function hypotheses. Table entries are adjusted R^2 of different models scaled relative to the grand mean fitted in model 1 (e.g. adjusted R^2 of a given model, $N_i = (\text{rms}[\text{model 1}] - \text{rms}[\text{model } N_i]) / \text{rms}[\text{model 1}]$, where rms = residual mean square), which we used to compare the five models to take account of both their explanatory power (proportion of total variance) and complexity (degrees of freedom). The selected best model for each site when analysed individually is highlighted in bold. Model 1: Productivity = Grand mean; Model 2: Productivity = Constant + b (Species number); Model 3: Productivity = Constant + $b \log$ (Species number); Model 4: Productivity = Constant + b (Species number) + c (Species number)²; Model 5: Productivity = Grand mean + a_i (where a_i is the effect of species level i)

Location	Model 1 null	Model 2 linear	Model 3 log-linear	Model 4 quadratic	Model 5 ANOVA
Germany	0	0.3031	0.2652	0.2919	0.2845
Ireland	0	0.0422	0.1312	0.2072	0.3039
Silwood	0	0.0461	0.0846	0.1031	0.1661
Sheffield	0	0.4161	0.4672	0.4895	0.4784
Switzerland	0	0.1145	0.1726	0.1347	0.1541
Portugal	0	0.1605	0.1755	0.1593	0.1588
Sweden	0	0.1165	0.0840	0.1058	0.1053
Greece	0	-0.0032	-0.0046	-0.0236	-0.0662

differences below. However, when combined in an overall analysis we found that these differences in the species richness relationships were not statistically significant and that the general pattern was well described by a relationship which was linear when diversity was put on a log₂ scale (Fig. 4.2(a)). Above-ground biomass also declined with decreasing numbers of functional groups (Fig. 4.2(b)) in the alternative analysis focusing on this aspect of diversity. Species composition was clearly biologically important as it interacted significantly with location, revealing that where the same species or mixtures of species occurred at multiple locations their performance differed from site-to-site, and the main effect of composition accounted for the largest single portion (39%) of the variation in biomass. When we decomposed the composition into its individual components, many species and the herb and legume functional groups had significant effects when examined individually. A detailed explanation of these analyses can be found in Hector *et al.* (2000b, 2002) and Schmid *et al.*, Chapter 6, but note that richness effects are tested against compositional effects (Tilman 1997a,b) and, by the same logic, interactions between richness and other terms against the corresponding interaction with composition.

The results reported above pose an apparent problem: how can we reconcile the differences between individual site analyses with the overall log-linear pattern, and the significant effects of species richness with the large amount of variation explained by differences in species composition?

A high degree of compositional variation between different communities is a common feature of biodiversity experiments. One recent realization is that effects of richness and composition are not mutually exclusive alternatives (Lawton 1998). This raises the issue of how to distinguish consistent relationships due to richness from potentially 'idiosyncratic' patterns dominated by compositional variation. One approach we have tried is to use the adjusted R^2 —the normal error structure special case of the Akaike Information Criteria (e.g. Burnham and Anderson 1998)—to compare different statistical models. The adjusted R^2 and AIC assess the efficiency of different models by considering their goodness of fit while taking into account their complexity and 'cost' in degrees of freedom. The adjusted R^2 does this through the following formulation:

$$1 - \frac{n-1}{n-p-1} (1 - R^2)$$

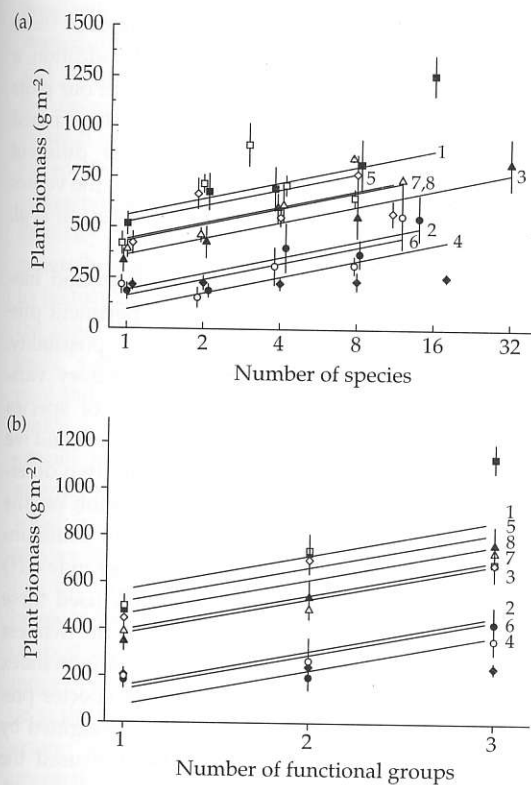


Figure 4.2 Above-ground biomass production increases with higher plant species richness (a) and functional group richness (b). Lines are regression slopes from the linear regression analyses presented in Hector *et al.* (1999) and symbols (staggered for clarity) are richness level means and standard errors for: closed squares = Germany, line 1; closed circles = Portugal, line 2; closed triangles = Switzerland, line 3; solid diamonds = Greece, line 4; open squares = Ireland, line 5; open circles = Sweden, line 6; open diamonds = Sheffield (UK), line 7; open diamonds = Silwood Park (UK), line 8.

where n is the sample size, p is the number of explanatory variables and R^2 is the proportion of the total sums of squares explained by the model. We used this procedure for the eight individual site patterns of above-ground biomass presented in Hector *et al.* (1999). Various schemes of alternative hypotheses have been proposed as possible diversity-function relationships (Vitousek and Hooper 1993; Lawton 1994; Naeem 1998; Schlöpfer and Schmid 1999). We compared the following five statistical models: a grand mean; a linear effect of species richness; a linear effect of species richness when diversity is put on a \log_2 scale (curvilinear on untransformed axes); a quadratic

effect of species richness (maximum or minimum productivity at intermediate levels of species richness); and when significant differences between factor levels in an analysis of variance are more efficient in explaining variation than simple linear or curvilinear relationships.

Table 4.1 shows that each of these five relationships was the most efficient model, using this criterion, for at least one site of the BIODDEPTH experiment. How can we reconcile this picture of differences in species richness responses, presented in Fig. 2 of Hector *et al.* (1999), with the picture of consistency presented in Fig. 1 of the same paper (compare Figs 4.1 and 4.2 of this chapter)? The differences stem in part from the different aims of hypothesis testing—which aims for a model with the minimal number of significant terms—versus model selection procedures that focus on finding the model that explains the most variation, or which is most efficient in explaining the most variation for the least cost in degrees of freedom (the adjusted R^2 and AIC approaches; note that there are other methods for doing this which differ in how they assess the most efficient model). For our data, the approach presented here in Table 4.1 focuses on analysing each site *individually* and with the aim of finding the most efficient model at each location. However, this approach sometimes includes terms in models which are not statistically significant and which are therefore eliminated with procedures that aim to find the model with the minimal number of significant terms, hence the two approaches sometimes differ in the models they select. Note also that although we were able to select one best model at each site using the adjusted R^2 , in many cases there is little to choose between alternative models. A second major point, is that even if a relationship is non-significant at one (or even all) of the individual sites tested, it does not follow that the more powerful combined test will also fail to show significance. We feel that progress in identifying general phenomena in ecology on objective grounds will require developing methods for testing of statistical consistency or differences in effects at different locations. Clearly, the challenge in doing this will be in identifying analysis strategies that can identify general patterns while also taking into account detailed differences between

locations. While there is clearly some way to go in this process, it is nonetheless underway in other areas of ecology (e.g. Reader *et al.* 1994; Zak *et al.* 1994; Gough *et al.* 2000; Gross *et al.* 2000). Results for multisite biodiversity experiments show a mixture of significant differences between locations (e.g. Van der Putten *et al.* 2000; Leps *et al.* 2001) and consistent general patterns (e.g. Bullock *et al.* 2001).

One of the particular complexities that arises when multiple locations are compared (Hector *et al.* 1999; Van der Putten *et al.* 2000; Bullock *et al.* 2001; Emmerson *et al.* 2001) is the additional between-sites components for species richness effects. In addition, while different sites are probably unlikely to be identical in their species lists, they may not overlap in species composition at all (some of the sites in Emmerson *et al.* 2001) or they may partially overlap with some species and mixtures re-occurring at two or more locations. For BIODPTH, reoccurrence of the same species mixtures at different locations was simply a product of our within-site species selection. An alternative, if more complex approach, would be to deliberately control how mixtures are common or unique to sites in the design. In either case, the advantage is that it is possible to test for different effects of diversity at different locations, by testing the location-by-richness terms against the location-by-composition interaction for example (the analysis of BIODPTH and biodiversity experiments in general are discussed in Schmid *et al.*, Chapter 6, and Hector *et al.* (2002)). In BIODPTH, the advantage of this was that it allowed us to determine that although species richness effects varied in individual location analyses, these differences were non-significant compared to the compositional variation of the same species assemblages grown at different sites (Hector *et al.* 1999, 2000b, 2002).

4.2.1 Experimental planned levels of diversity versus observed values

One of the problems that arises from the multifaceted nature of biodiversity is that it is difficult, indeed impossible, to control all aspects at once in a single experimental design. For example, in the BIODPTH experiments, we controlled the initial numbers of species sown into our plots but we did

not try to control the relative abundances of these species after the initial seed sowing. In addition, a small number of species did not persist in our plots reducing the numbers present relative to the planned numbers. Might our results change if different measures of diversity are used instead of the values from the initial experimental design (Huston *et al.* 2000)?

We re-analysed the data on above-ground biomass from the second year of the experiment presented in Hector *et al.* (1999) to test this possibility. We calculated three alternative explanatory variables. We used the observed number of species present in the second year of the experiment and we also used the biomass data to calculate two diversity indices that incorporated information on the relative abundances of the species in our communities. We calculated the Shannon–Weaver index (H) and the Simpsons' index as $1 - D$. We used these two diversity indices that both combine the richness and evenness of species because the Shannon index can be weighted towards numbers of species present, whereas the Simpsons' index is weighted by dominant species (Magurran 1988). We used the indices on their original scales for comparison with $\log_2(\text{species richness})$, which we found described our data well, but the antilog of H' is a further possibility. Diversity indices can have properties, such as non-normal distributions, that can be problematical for analysis; we use them in these exploratory analyses as in practice they produced reasonable residual plots and ANOVA is a relatively robust technique.

We found, that just as for sown species richness, all three alternative explanatory variables had highly significant positive relationships with biomass (observed species richness: $F_{1,196} = 48.92$, $P < 0.001$; Shannon index: $F_{1,196} = 26.4$, $P < 0.001$; Simpsons' index: $F_{1,196} = 27.49$, $P < 0.001$) and that these relationships did not differ significantly between locations (all location-by-diversity interaction terms $F < 2$ and $P > 0.05$). Surprisingly, even after controlling for the Shannon index values there was a highly significant residual effect of the sown number of species ($F_{1,195} = 28.17$, $P < 0.001$). There are a number of possible explanations for this unexpected result. Setting aside the trivial point that it is likely that any survey of diversity at a particular place and

time will underestimate the true diversity, more importantly, a process measured at a certain place and point in time may be influenced not only by the diversity of organisms currently present at that location but also by neighbouring individuals and by species that were present but have now disappeared. This second point is not specific to biodiversity experiments but may be true more generally: the results of an experiment may reflect its history and not just the explanatory variables at the time of measurement (Harper 1977). Interestingly, a third possibility concerns the size of the pool and hence the diversity of functional traits from which a number of species is derived; a 'multi-species sampling effect' (Loreau *et al.* 2001). For example, consider two sets of four species, one of which is derived from an initial (sown) set of eight species and the other from a group of six. It may be that the subset derived from the set of eight could be more productive than that derived from the set of six if the set of eight species had a greater initial range of functional traits, some of which was passed on to the subset of four species, thereby providing a greater potential for complementary and positive interactions. Clearly, there is much to be gained from additional analyses examining different aspects of diversity but it appears that the planned species richness is not only the 'correct' explanatory variable in terms of the experimental design and hypotheses but that it is also a good reflection of longer-term effects and can have greater explanatory power than measurements of diversity at one point in space and time during the experiment.

4.2.2 Partitioning biodiversity effects

The other most controversial aspects of the early BIODEPTH results concerned identification of the mechanisms generating the negative relationship between above-ground biomass and declining diversity. Initial explanations for the effects of biodiversity on ecosystem processes focused on niche complementarity (Naeem *et al.* 1994a; Tilman *et al.* 1996) through the partitioning of resources. However, there is a simpler way, which was initially missed, in which diversity can affect ecosystem functioning even in the absence of resource use complementarity. Under the sampling effect hypothesis

(Aarssen 1997; Huston 1997; Tilman *et al.* 1997a), when communities are assembled at random from a pool of species, more diverse mixtures have a higher probability of containing a species with extreme traits which could become dominant and drive ecosystem functioning. A more general 'selection effect' (Loreau 1998a, 2000a; Loreau and Hector 2001) is obtained if two assumptions of the sampling effect are relaxed. First, a single species need not dominate completely and, second, dominance need not be perfectly positively correlated with increasing monoculture productivity or biomass. Thus, biodiversity effects can be grouped into two classes: 'complementarity effects' (including resource-partitioning complementarity, positive interactions and negative interference) and 'selection effects' that occur through dominance or subordination of species with particular traits. We have recently presented a new method that performs an additive partitioning of the individual contributions of the two effects in biodiversity experiments.

Our additive partition unifies and relates in a single equation previous measures based on the relative yields and proportional deviation from expected value approaches (Garnier *et al.* 1997; Hector 1998; Hooper 1998; Loreau 1998b; Emmerson and Raffaelli 2000; Dukes 2001), with a way of estimating selection in mixed populations analogous to the Price equation from evolutionary genetics (Price 1970, 1995; Frank 1997). The method provides absolute estimates of different biodiversity effects allowing quantitative comparison of their respective contributions. The Price equation (Price 1970, 1995; Frank 1997) is typically used to separate changes in character traits that are due to the direct effects of natural selection in altering the frequencies of different alleles from those due to interactions between alleles in the altered population; the fidelity of transmission. Analogously in our method, 'ecological selection' occurs when changes in the relative yields of species in a mixture are non-randomly related to their traits (e.g. yields) in monoculture causing dominance and subordination of species. The selection effect is therefore determined by the covariation between monoculture traits and relative abundance in mixtures in the same way as in the Price equation. Positive selection occurs if species with higher-than-average monoculture yields

dominate mixtures, and negative selection if the opposite is true. The complementarity effect measures change in the average relative yields of species in mixtures compared to their expected values under the null hypothesis that yields in mixture will equal the monoculture yield times the proportion of the species in mixture. That is, the expected relative yield of species i is

$$RY_{e,i} = P_i M_i$$

The observed relative yield is

$$RY_{o,i} = Y_{o,i} / M_i$$

and the deviation in relative yields equals

$$\Delta RY_i = RY_{o,i} - RY_{e,i}$$

where P_i is the proportion of species i in mixture; M_i is the monoculture yield of species i ; and $Y_{o,i}$ is the observed yield of species i in mixture. The method assesses whether increases in some species in mixtures are balanced by declines in others or whether there is evidence for complementary (resource partitioning), positive (facilitation) or negative (physical or chemical interference) interactions that shift the total yield away from the null prediction which assumes none of these additional interactions. A positive complementarity effect (resource partitioning or facilitation) occurs if the average deviations from expected values of the relative yields of the species in a mixture is higher than expected, and a negative complementarity effect (direct interference) if it is lower. The sum of the selection and complementarity effects gives the net biodiversity effect, which is the difference between the observed yield of a mixture and its expected yield under the null hypothesis that there is no selection effect or complementarity effect. This expected null value is the average of the monoculture yields of the component species weighted by their initial relative abundance in mixture, which for BIODEPTH and similar substitutive designs is $1/N$, where N is the number of species in the mixture and is thus a simple averaging of the single-species yields. The selection, complementarity and net biodiversity effects all have the dimension of the ecosystem property in question (such as yield)

and an expected value of zero under the null hypothesis of no biodiversity effect. Full details of the additive partitioning method can be found in Loreau and Hector (2001) but the basic equation expressed in the terms described above is

$$\text{Net biodiversity effect} = N \overline{\Delta RY} \overline{M} + N \text{cov}(\Delta RY, M)$$

All three biodiversity effects can be positive or negative, and complementarity and selection effects can therefore fully or partially cancel each other. There are therefore nine possible qualitative outcomes that arise from the combinations of positive, zero or negative complementarity and selection effects.

We provide a worked example of the method in Table 4.2 and Fig. 4.3 which illustrates these nine possible qualitative outcomes (cases (a)–(i) in the following all refer to both Table 4.2 and Fig. 4.3). The null situation refers to no effect of complementarity or selection and therefore no net effect (e). When they arise, biodiversity patterns can be either purely due to selection or complementarity effects with zero values of the other biodiversity effect ((b) and (d)); note that no selection effect arises only when the relative effects of complementarity are equally distributed between species), or a combination of reinforcing positive (a) or negative (i) complementarity and selection effects. Negative selection effects occur when dominance is by a species with a lower-than-average monoculture biomass ((c), (f), (i)), which can then hide positive complementarity (c). Positive and negative effects may even cancel each other out exactly producing no net effect (c). Negative complementarity effects ((g)–(i)) indicate direct interference between species.

In general qualitative terms, joint positive biodiversity effects (a) appear to be the situation for the BIODEPTH experiments in Ireland the Sheffield (Loreau and Hector 2001) and for Cedar Creek (see Tilman *et al.* Chapter 3, 2001; Tilman 2001). Counteracting positive complementarity and negative selection (c) appear to produce a zero-to-negative net effects in a Californian Serpentine Grassland (Hooper and Vitousek 1997; Hooper 1998) and a positive net effect at the Portuguese BIODEPTH site. The overall pattern across all eight BIODEPTH

Table 4.2 Worked examples of different scenarios for the additive partitioning of biodiversity effects showing the different qualitative outcomes illustrated in Fig. 4.3

	M_i	$Y_{e,i}$	$Y_{o,i}$	$RY_{e,i}$	$RY_{o,i}$	ΔRY	NE	CE	SE
<i>(a) Positive SE, positive CE (transgressive overyielding)</i>									
Species A	800	400	700	0.5	0.875	0.375	—	—	—
Species B	200	100	150	0.5	0.75	0.25	—	—	—
Total mixture	—	500	850	1	1.625	—	350	312.5	37.5
Mean	500	—	—	—	—	0.3125			
<i>(b) No SE, positive CE (non-transgressive overyielding)</i>									
Species A	800	400	600	0.5	0.75	0.25	—	—	—
Species B	200	100	150	0.5	0.75	0.25	—	—	—
Total mixture	—	500	750	1	1.5	—	250	250	0
Mean	500	—	—	—	—	0.25			
<i>(c) Negative SE, positive CE</i>									
Species A	800	400	350	0.5	0.4375	-0.0625	—	—	—
Species B	200	100	150	0.5	0.75	0.25	—	—	—
Total mixture	—	500	500	1	1.1875	—	0	93.75	-93.75
Mean	500	—	—	—	—	0.09375			
<i>(d) Positive SE, no CE (no overyielding)</i>									
Species A	800	400	600	0.5	0.75	0.25	—	—	—
Species B	200	100	50	0.5	0.25	-0.25	—	—	—
Total mixture	—	500	650	1	1	—	150	0	150
Mean	500	—	—	—	—	0			
<i>(e) No SE, no CE (Null hypothesis)</i>									
Species A	800	400	400	0.5	0.5	0	—	—	—
Species B	200	100	100	0.5	0.5	0	—	—	—
Total mixture	—	500	500	1	1	—	0	0	0
Mean	500	—	—	—	—	0			
<i>(f) Negative SE, no CE</i>									
Species A	800	400	200	0.5	0.25	-0.25	—	—	—
Species B	200	100	150	0.5	0.75	0.25	—	—	—
Total mixture	—	500	350	1	1	—	-150	0	-150
Mean	500	—	—	—	—	0			
<i>(g) Positive SE, negative CE (interference)</i>									
Species A	800	400	500	0.5	0.625	0.125	—	—	—
Species B	200	100	25	0.5	0.125	-0.375	—	—	—
Total mixture	—	500	525	1	0.75	—	25	-125	150
Mean	500	—	—	—	—	-0.125			
<i>(h) No SE, negative CE</i>									
Species A	800	400	400	0.5	0.5	0	—	—	—
Species B	200	100	100	0.5	0.5	0	—	—	—
Total mixture	—	500	500	1	1	—	0	0	0
Mean	500	—	—	—	—	0			
<i>(i) Negative SE, negative CE</i>									
Species A	800	400	200	0.5	0.25	-0.25	—	—	—
Species B	200	100	75	0.5	0.375	-0.125	—	—	—
Total mixture	—	500	275	1	0.625	—	-225	-187.5	-37.5
Mean	500	—	—	—	—	-0.1875			

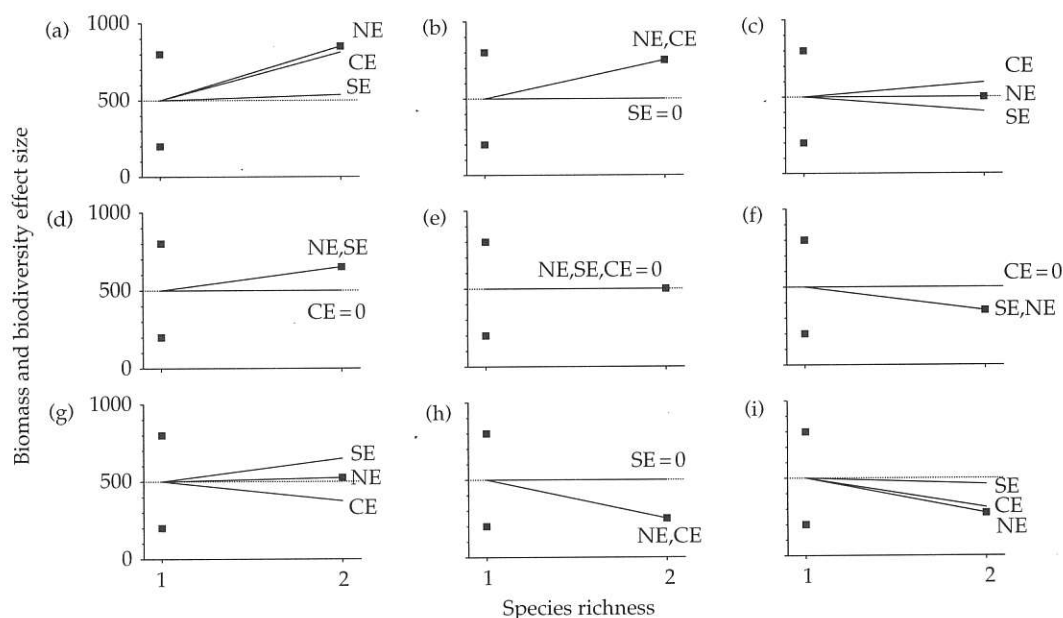


Figure 4.3 Examples of the nine different possible qualitative outcomes of the additive partitioning method presented in Table 4.2: (a) positive selection, complementarity and net effects; (b) zero selection effect, positive complementarity and net effect; (c) equal negative selection effect and positive complementarity effect producing zero net effect (example 1 from Loreau 1998); (d) positive selection effect and zero complementarity effect producing positive net effect; (e) zero selection, complementarity and net effects; (f) negative selection effect and zero complementarity effect producing negative net effect; (g) positive selection effect and negative complementarity effect; (h) zero selection effect and negative complementarity effect producing a negative net effect; (i) negative selection, complementarity and net effects. Symbols are monoculture and two-species mixture total yields, solid lines are biodiversity effects and broken lines show the null scenario of an averaging of monoculture yields (note that biodiversity effects have been scaled relative to this null value of 500 rather than zero so that they can be shown on the same figure as the observed yields).

sites pointed to a positive net effect generated by positive complementarity with a zero selection effect on average (b) as described in Loreau and Hector (2001). Since our approach requires a comparison between the performances of species in mixture and in monoculture, we restricted its application to the subset of experimental mixture plots that contained species for which monoculture yields were available. We discuss only the overall patterns across all sites here but individual site variations are described in Loreau and Hector (2001). The overall log-linear increase in above-ground biomass with species richness for the whole experiment was observed for this subset of the data. The net biodiversity effect was positive (the grand mean was significantly different from zero; Loreau and Hector 2001 Fig. 1(a)) and increased significantly with species richness beyond two species (Fig. 4.4(a)). However, the selection effect was variable across

individual localities and overall these variations cancelled out so that the grand mean was not significantly different from zero, and on an average the selection effect was unaffected by changes in species richness. The only factors that influenced the selection effect significantly were locality and species composition (Fig. 4.4(c)). Complementarity effects at individual locations were also variable but the combined analysis revealed significant locality and composition main effects and a significant positive relationship with species richness (Fig. 4.4(b)). The presence of legumes in mixtures had important impacts on their performance; in general, they tended to increase observed yields and the net and complementarity effects, and to generate more extreme selection effects, both positive and negative (Fig. 4.4(c)). However, species richness retained a significant log-linear effect on complementarity even when the presence of legumes was included

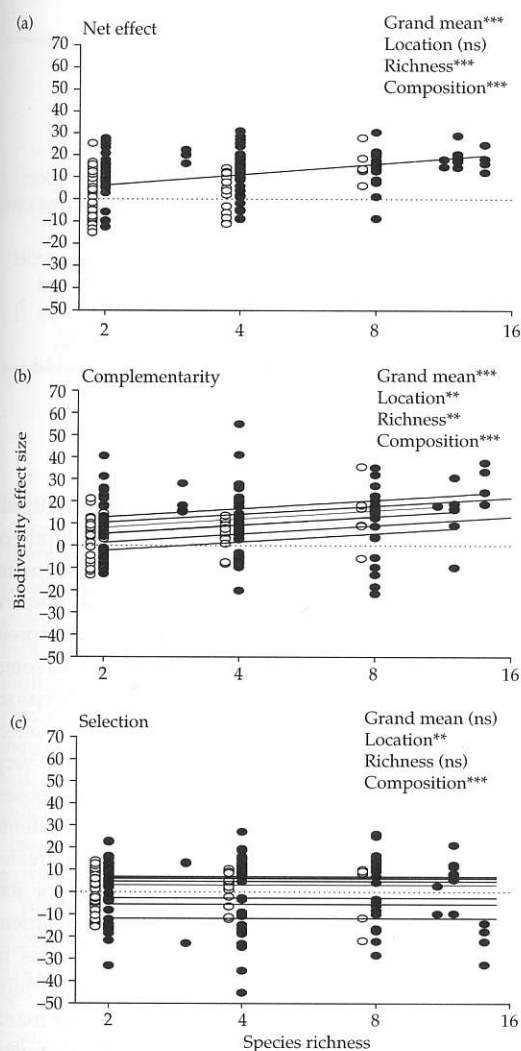


Figure 4.4 Net biodiversity effect ΔY , (a), Complementarity effect (b), $N\Delta RYM$, and Selection effect $Ncov(\Delta RY, M)$, (c), as functions of species richness for year two above-ground biomass production. Open circles are plots that do not contain any legume species; filled circles are plots that contain one or more species of legume. Lines are slopes from the multiple regression model using species richness on a log₂ scale. Complementarity effect lines from highest elevation to lowest are: Portugal, Switzerland, Silwood, Germany, Sheffield, Ireland, Greece, and Sweden. Selection effect lines from highest elevation to lowest are: Ireland, Germany, Sheffield, Greece, Sweden, Silwood, Switzerland, and Portugal. Values of the biodiversity effect (in g m⁻²) were square root transformed while preserving the original positive and negative signs to meet the assumptions of analyses. Results are summarized for the grand mean for the three biodiversity effects and for the influence of location, species richness and composition:

* = $P < 0.05$, ** = $P < 0.01$, *** = $P < 0.001$. Adapted from a figure in *Nature* with permission.

as an additional factor in our across-site analyses (this test is very conservative since part of the species richness effect is absorbed into the legume effect when the latter is fitted first in the analysis).

To summarize, the results of the overall across-site analysis of the complementarity effect showed a much closer match with the log-linear relationship found for the above-ground biomass patterns than did the selection effect. The increased complementarity in species-rich mixtures involved not only complementarity between legumes and other plant types, but also complementarity between species within each of these groups. Therefore, our analysis suggests that the positive relationship between above-ground biomass production and increasing diversity was driven by the complementarity effect and could not be explained by the selection effect. The additive partitioning method has not yet been applied to other multisite studies but it appears that in all cases both sampling and complementarity effects act in combination (Van der Putten *et al.* 2000; Leps *et al.* 2001; Bullock *et al.* 2001; Emmerson *et al.* 2001).

4.2.3 Dominance and above-ground biomass production

A central component of the sampling effect hypothesis concerns the dominance of plant communities by the most productive species in monoculture. Indeed, several researchers have predicted that biomass production in plant communities should be greater for communities that are strongly dominated than for those with a more even distribution of species (e.g. Huston 1997; Grime 1998; Huston *et al.* 2000; Wardle *et al.* 2000b). To test this hypothesis for the BIODDEPTH data, we calculated the proportion of total community biomass of the dominant species in the third year of the experiment and separated communities that were less strongly dominated by a single species (<70% of total above-ground biomass) from those that were more strongly dominated, splitting the dataset approximately in half for a balanced analysis. Counter to predictions, we found a stronger positive effect of increasing species richness on biomass for communities that were less strongly dominated and than for those that were more dominated by a single species (Fig. 4.5). This

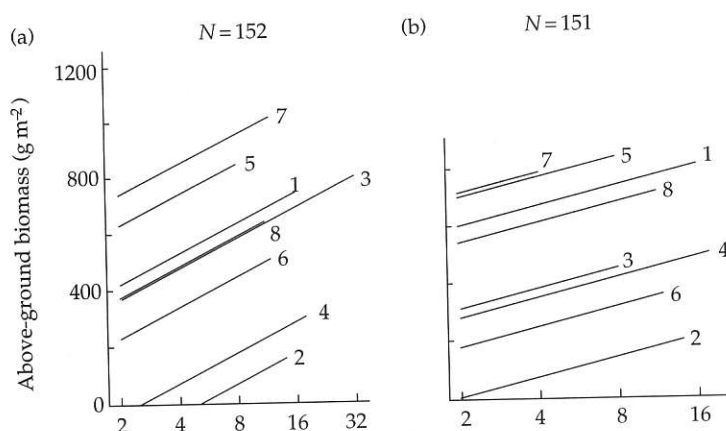


Figure 4.5 The positive effect of species richness on above-ground biomass production is stronger for communities that are less strongly dominated by a single species ((a) most dominant species <70% total biomass; \log_2 slope = $107.6 \pm 16.7 \text{ g m}^{-2}$) than for those that are more strongly dominated (b) >70%; \log_2 slope = $67.8 \pm 19.2 \text{ g m}^{-2}$). Sample sizes are given as 'N'.

may have been due to increased dominance in low-productivity sites such as Sweden and Portugal due to the effects of drought at the former and frost and drought in the latter in eliminating individuals of vulnerable species, thereby causing greater relative dominance by the remaining resistant species. These analyses concentrate only on the single most dominant species and we are now examining dominance and evenness over all species. However, this analysis also suggests that dominance by individual species cannot explain productivity patterns in our experiment supporting a role for complementarity interactions.

4.3 Summary

This chapter has focused on the unique aspects of multisite biodiversity experiments. Although only a handful of these experiments exist, all show that changes in biodiversity impact a wide variety of ecosystem processes to some degree. All of these studies also seem to produce results that are generated by a combination of selection and complementarity effects but with variation in which of these biodiversity effects dominates. This work clearly illustrates the broader importance in ecological studies of separating sampling processes from the effects of biotic interactions (e.g. Oksanen 1996;

Stevens and Carson 1999). Some of the multisite studies show overall consistent patterns while others reveal significant differences between locations in diversity effects. We have explored several complex issues related to the design, analysis and interpretation of biodiversity experiments that leads to some recommendations. First, because biodiversity has so many overlapping components the interpretation of studies will be greatly helped by setting very specific hypotheses, together with the proposed mechanisms; a point that applies to many questions in ecology. Second, ecology clearly needs a greater understanding and discussion of methods for statistically testing differences in biological relationships at different sites if we are to progress in distinguishing general from variable relationships across multiple locations. Finally, biodiversity manipulation experiments clearly need to be integrated with more classical correlational studies of environmental gradients; this is a clear priority for new analyses of existing datasets from multisite biodiversity experiments. A new generation of multisite studies that are specifically designed for this purpose could play a key role in achieving this important goal.

Our thanks to A. Kinzig for suggestions on presenting the partitioning method.