

The relative contributions of species richness and species composition to ecosystem functioning

Nadine Sandau, Yvonne Fabian, Odile T. Bruggisser, Rudolf P. Rohr, Russell E. Naisbit, Patrik Kehrli, Alexandre Aebi and Louis-Félix Bersier

N. Sandau (nadine.sandau@gmail.com), Y. Fabian, O. T. Bruggisser, R. P. Rohr, R. E. Naisbit and L.-F. Bersier, Unit of Ecology and Evolution, Univ. of Fribourg, Chemin du Musée 10, VH-1700 Fribourg, Switzerland. NS also at: Ecosystem Dynamics, Swiss Federal Inst. for Forest, Snow and Landscape Research WSL, Birmensdorf, Switzerland. – P. Kehrli, Station de recherche Agroscope Changins – Wädenswil ACW, Nyon, Switzerland. – A. Aebi, Laboratory of Soil Biology, Univ. of Neuchâtel, Neuchâtel, Switzerland.

How species diversity influences ecosystem functioning has been the subject of many experiments and remains a key question for ecology and conservation biology. However, the fact that diversity cannot be manipulated without affecting species composition makes this quest methodologically challenging. Here, we evaluate the relative importance of diversity and of composition on biomass production, by using partial Mantel tests for one variable while controlling for the other. We analyse two datasets, from the Jena (2002–2008) and the Grandcour (2008–2009) Experiments. In both experiments, plots were sown with different numbers of species to unravel mechanisms underlying the relationship between biodiversity and ecosystem functioning (BEF). Contrary to Jena, plots were neither mowed nor weeded in Grandcour, allowing external species to establish. Based on the diversity–ecosystem functioning and competition theories, we tested two predictions: 1) the contribution of composition should increase with time; 2) the contribution of composition should be more important in non-weeded than in controlled systems. We found support for the second hypothesis, but not for the first. On the contrary, the contribution of species richness became markedly more important few years after the start of the Jena Experiment. This result can be interpreted as suggesting that species complementarity, rather than intraspecific competition, is the driving force in this system. Finally, we explored to what extent the estimated relative importance of both factors varied when measured on different spatial scales of the experiment (in this case, increasing the number of plots included in the analyses). We found a strong effect of scale, suggesting that comparisons between studies, and more generally the extrapolation of results from experiments to natural situations, should be made with caution.

The current global decline of biodiversity is causing concern over how ecosystem functioning and the provision of ecosystem services will be affected. In order to understand the consequences of biodiversity loss, manipulative biodiversity experiments (e.g. Cedar Creek, BIODEPTH or the Jena Experiment; Tilman et al. 1996, Hector et al. 1999, Roscher et al. 2004) have tried to explore the dependent relationships between biodiversity and ecosystem functioning (Hooper et al. 2012). While some experiments have analysed ecosystem multifunctionality (Gamfeldt et al. 2008, Zavaleta et al. 2010, Mouillot et al. 2011), most studies have concentrated on the species richness–biomass relationship. Although positive relationships between diversity and productivity are frequently reported (Hector et al. 1999, Hooper et al. 2005, 2012, Roscher et al. 2005, Cardinale et al. 2012, Isbell et al. 2015), the topic remains contentious as the direction of the relationship varies substantially between natural ecosystems and experimentally assembled plant communities (Chapin III et al. 1997, Grace et al. 2007, Assaf et al. 2011, Wardle

2016). While the idiosyncratic hypothesis anticipates that ecosystem functioning changes unpredictably with varying species richness because each species contributes unequally to productivity (Lawton 1994), the mass ratio hypothesis states that the characteristics of the dominant species in the community govern ecosystem properties. As a consequence, not only the number of species, but also species identity and therefore composition, is essential when members of a community disappear (Aarssen 2001).

Analyses of the role of composition in modulating the diversity–productivity relationships have yielded contrasting results, from studies reporting commensurate importance of diversity and composition (Hector et al. 2011), to those emphasizing the role of composition (Harvey et al. 2013, Sandau et al. 2014). Attempts to disentangle diversity effects from compositional effects are complicated because of the very different nature of the two variables: plant species richness is a quantitative variable that is easy to include in statistical models; whereas composition can be described

by the presence/absence or abundance of each species in the experimental plots. Composition is thus inherently multivariate, and each plot can be visualized as a point in an n -dimensional space, where n corresponds to the total number of species in the study. The contrasting nature of richness and composition raises analytical problems when incorporated together into a statistical model (Sandau et al. 2014).

The classical approach in biodiversity–ecosystem functioning analyses is to consider species composition as a random variable, with each different sown mixture considered as a different level drawn from the set of all possible compositions from the experimental species pool (Hector et al. 2011). Alternative models have been developed, e.g. approaches based on the Price equation (Fox 2006, Fox and Harpole 2008), or on pairwise interactions between plant species (the diversity interaction model and its variants; Kirwan et al. 2007, Connolly et al. 2013). Also, an experimental design has been devised specifically to separate the effects of species richness and of species identity (Bell et al. 2009). These models are tailored for experimental studies and their application to observational data is difficult or even impossible. Kahmen et al. (2005) and Sandau et al. (2014) have proposed statistical methods suitable for such situations. The first one uses an ordination method to reduce the multidimensionality of species composition to few axes; the coordinates of the plots on these axes can then be used as explanatory variables in regression models. The second approach assumes that the similarity in species composition between plots induces statistical dependence between the observations; this dependence is included in regression models as a correlation structure in a way similar to phylogenetic regression correcting for species relatedness.

All the above methods allow a test of the statistical effect of species richness or of species composition on productivity, however they cannot measure their relative contributions because the two variables are included differently in the models. To measure the relative contributions of species richness and species composition, Hector et al. (2011) used a method that considers all explanatory variables as random factors (Gelman 2005). They used a Monte Carlo Markov chain (MCMC) approach to estimate the parameters of both species composition and species richness, and computed the standard deviation of their posterior distributions as a measure of contribution of each variable. Based on the data of several biodiversity–ecosystem functioning experiments, they concluded that the contributions of species richness and of species composition were very similar. This article has however elicited further discussion (Connolly et al. 2011), and the use of a continuous variable (i.e. species richness) as a random factor is disputable (Bates 2010).

A straightforward method to explore the relative importance of richness and composition is to use Mantel tests. In this way, both factors are treated as pairwise distances among sites. Information on productivity, species richness and species composition are expressed in matrices containing the Euclidean distances in productivity between plots (the ‘response variable’), and the Euclidean distances in species richness, as well as distances in species composition, respectively (the ‘explanatory variables’). It is important to

understand that Mantel tests do not assume a particular relationship between variables (e.g. a positive or negative linear relationship between species richness and biomass); they simply assess the following question: do pairs of plots with similar richness/composition also tend to have similar biomass? This is achieved by evaluating the linear relationship between the pairwise distance values of the matrices. By using partial Mantel tests, it is possible to partial out the effect of each variable in turn (in this way accounting for a possible correlation between the variables), and thus to assess their relative importance (Legendre and Legendre 1998, pp. 557). The pros and cons of partial Mantel tests have been the subject of several articles (Harmon and Glor 2010); one notable concern is its low statistical power compared to other approaches. In our case, our aim is not to find the best model to explain variation in biomass between plots, but to assess the relative contributions of diversity and composition as explanatory variables. In this case, a decisive advantage of Mantel tests is that both explanatory variables are treated in a similar way, i.e. as dyadic data between observations.

Here we analyse used the results of two biodiversity experiments: the Jena Experiment (Weigelt et al. 2010), and the Grandcour Experiment, which consisted of 12 sown wildflower strips in Switzerland (Fabian et al. 2013, Sandau et al. 2014). A major difference between both experiments is that the latter was not mown nor weeded, meaning that colonizer species could establish in the experimental plots.

In addition to the general question of which variable (diversity versus composition) is more important in explaining variation in productivity, we explore several hypotheses regarding how the relative importance of both factors evolves with time, differs between study types, and changes with spatial scale. Specifically, our three research questions are:

- 1) How does the relative importance of species richness and of plant composition change with time in the Jena Experiment? We hypothesize that the contribution of species richness will decrease and that of plant composition will increase. As less competitive species will be excluded by dominant species (Tilman and Downing 1994, Aarssen et al. 2003), species richness will tend to homogenise; at the same time, plots with the same dominant species will tend to have similar productivity, and hence increase the contribution of composition.

- 2) Is the relative importance of species richness and species composition different in controlled experiments with weeding than in experiments where colonizing species are allowed? We hypothesize that species richness is less important in experiments without weeding because variability in species richness between the plots will be reduced due to compensatory effects of colonizing species modulated to biotic resistance (i.e. colonizers can more readily become established within in species poor communities).

- 3) Is the relative importance of species richness and of plant composition different at different spatial scales? This question can be tackled in the Grandcour Experiment that was set up as a metacommunity experiment. We hypothesize that the observed values can vary strongly with spatial scale, depending on the variability in species richness and in species composition at different spatial scales.

Methods

Datasets

We used two datasets, the first from an experiment conducted in 12 wildflower strips in Switzerland, the Grandcour Experiment (Fabian et al. 2012, Bruggisser et al. 2012, Sandau et al. 2014); the second from the Jena Experiment (Weigelt et al. 2010). In the Grandcour Experiment, the number of species was manipulated by sowing different mixtures in 2007. A treatment to control the main herbivores and their predators was applied with fences (with three levels: no fence, small-meshed fence and large-meshed fence). Each of the 12 experimental fields contained the three levels of herbivore–predator control treatment, and in each level four 9×6 m subplots were sown with different plant mixtures (2, 6, 12 and 20 sown plant species, all from the tall herb functional group), resulting in a total of 144 subplots (further information in the Supplementary material Appendix 1). An important difference from classical biodiversity–ecosystem functioning experiments was the absence of weeding and of mowing; consequently, species richness is expressed as the total number of species. Plant species composition was described by the percentage cover of all plant species in the subplots, estimated by visual inspection. Total aboveground biomass was assessed from leaf area index measurements, calibrated with results derived by the classical clipping method in eight subplots (Pearson correlation between both measurements was 0.89). Species percentage cover, species richness and biomass were measured in late summer 2008 and 2009.

In the Jena Experiment, temperate grassland communities with changing species richness and functional group number were created from a pool of 60 species. Different species richness levels (2, 4, 8, 16 and 60) were sown in 66 plots of 20×20 m organized in four blocks (note that we do not consider plots sown with zero or one species). Biomass was harvested in 3–4 subplots of 0.2×0.5 m in each plot in late spring and late summer, usually in May and August and sorted to species. Plant species cover was estimated in a separate 9-m^2 subplot. Plots were mown and weeded twice a year. For our analyses we used the published data from 2002 to 2008, totalling 12 sampling sessions (for more information see Weigelt et al. 2010).

Statistical analyses

For the purpose of comparing the relative contribution of species richness versus composition, we use partial Mantel tests. Mantel tests allow for the comparison of two distance (or similarity) matrices measured on the same observational units (e.g. geographic distances versus differences in leukaemia incidences for a selection of localities; Mantel 1967). This is achieved simply by measuring the correlation r between all corresponding entries of both matrices, and using permutations of the columns and their corresponding row to assess the statistical significance of r . Partial Mantel tests, based on partial correlation coefficients (Legendre and Legendre 1998), were devised to account for one or several additional distance matrices, with the aim of partialling out their contribution in the correlation between the two matrices of interest.

The partial correlation coefficients can be used to assess the relative importance of each set of variables from which the distance matrices are computed, and the unique and common importance of each set can be evaluated. Note however that the total amount of variation, and thus the percentage of unexplained variance (Borcard et al. 1992, Fabian et al. 2012), cannot be measured with this approach.

We computed distance matrices containing the differences (Euclidean distance) in biomass between (sub)plots, the differences in realized species richness (i.e. based on the observed and not the sown number of species), and Bray–Curtis floristic coefficients. For the latter, we used raw percentage cover of the plant species (using relative cover yielded highly similar results). Matrices were computed for each sampling session. Distance matrices were computed with the `vegdist` function of the `vegan` package (Oksanen et al. 2015) in R (<www.r-project.org>). In our analyses, we also considered the random effects created by fields in the Grandcour Experiment or by blocks in the Jena Experiment, by adding a design matrix defining field or block membership of the plots (Legendre and Legendre 1998). A distance of 0 was given to pairs of (sub)plots if they belonged to the same field or block, and 1 if otherwise. For the Grandcour Experiment, we also accounted for the herbivore–predator control treatment by adding three design matrices defining the within-field membership to each level (no fence, small-meshed fence and large-meshed fence), in this way considering this treatment as a random factor.

We used the function `mantel` of the `ecodist` package (Goslee and Urban 2007) to perform the partial Mantel tests with 10 000 permutations based on Pearson correlations r . We used two models, both with distances in biomass as ‘dependent variable’, to compare 1) the effect of species richness while controlling for composition, and 2) the effect of composition while controlling for species richness. The block membership matrices for the Jena Experiment, or the four field and treatment membership matrices for the Grandcour Experiment, were always included. We reported the one-tailed p -values, testing the null hypothesis $r \leq 0$, since negative r indicate that pairs of (sub)plots with similar richness (or composition) have divergent biomasses. As the two explanatory matrices representing the distance in species richness and species composition may be correlated, we determined the overlap of their explanatory power according to the formula:

$$\text{overlap} = \frac{(r_1^2 + r_2^2) - (r_{p1}^2 + r_{p2}^2)}{2}$$

with r_{p1} and r_{p2} the partial Mantel coefficients as defined above, and r_1 and r_2 the partial Mantel coefficients for the two Mantel tests models with either species richness (species composition), while not controlling for the effect of species composition (species richness), (note that these are partial coefficients since the experimental structure was always included in the models).

Additionally, for the Jena dataset, we performed two complementary analyses useful for the interpretation of the Mantel tests’ results. Firstly, we analysed if plant composition and species richness between plots tended to become more homogeneous or heterogeneous with time. For each sampling session, we computed the average pairwise Bray–Curtis distances, and the Euclidean distances in average

species richness, between all plots of the same sown diversity level. Generalized least square models with AR1 correlation structure were applied to detect trends (Zuur et al. 2009), using the function gls of the nlme package (Pinheiro et al. 2013).

Secondly, we analysed the relationship between the average biomass of the species and their tendency to increase or decrease during the six years of data. Since globally the total biomass decreased in the course of the experiment, we used relative biomass in each plot and at each sampling session to analyse the temporal trends of the species. This trend was estimated separately for the spring and summer sampling sessions, as there is a strong seasonal effect (in general more biomass in spring). We computed non-parametric Spearman correlations between relative biomass and sampling session date for each species in each plot to describe its temporal trend. We counted the number of times each species decreased or increased in the plots where it was present. The average biomass of each species was computed from the relative biomass over all plots and sampling sessions; these values were then ranked. These ranked averages were used as explanatory variable for the binomial response variable ‘number of increases versus decreases in proportional biomass’. We used a generalized mixed-effects model for this analysis, with season as random factor and quasi-binomial family (function glmmPQL of the MASS package; (Venables and Ripley 2002)).

Finally, for the Grandcour Experiment, we performed partial Mantel tests as above to assess the effect of spatial scale on the relative contribution of richness and composition on productivity. To simulate varying scales of the experiment, we first computed the partial Mantel tests for the 12 individual fields (fine scale analysis), then on combinations of 2, 4, 6, 8 and 10 fields pooled, effectively broadening the scale of observation. All possible combinations of fields were considered, resulting in 66, 495, 924, 495 and 66 tests, respectively.

Data deposition

Data available from the Dryad Digital Repository: <<http://dx.doi.org/10.5061/dryad.44bm6>> (Sandau et al. 2016).

Results

Temporal trends in contributions of species richness and composition

In the Jena Experiment, across all years, we found that richness contributed to 62% and composition to 38% of the similarity in biomass production (Table 1). Nevertheless, the contribution of composition relative to richness changed with time since the establishment of the experiment. We found a strong temporal trend in the contributions of species richness and species composition to biomass, with species richness becoming more important with time (Fig. 1). This trend, expressed as the percentage contribution of species richness (Table 1), was significant (generalized least squares model on logit transformed data with AR1 correlation structure; slope for richness contribution = 0.089, p-value = 0.006).

The decrease in the contribution of plant composition was not merely due to a homogenization of the different communities with time. On the contrary, plots sown with 8 and 16 species tended to become more different in composition (Table 2). In addition, the increase in contribution of species richness was not due to an increase in the heterogeneity of species number; only plots with 16 sown species showed a significant positive trend (Table 2). Finally, we examined if there was a relationship between species proportional abundance and their propensity to decrease or increase during the course of the experiment. We found weak and non-significant negative trend (Fig. 2). Note that these analyses were not possible for the Grandcour Experiment since data were collected only during two sampling sessions.

Difference between weeded and non-weeded experiments

In the Grandcour Experiment, the partial Mantel test for the effect of species richness on total biomass was not significant, whereas the effect of composition was highly significant (Table 1). These results were consistent for both years. Species richness explained little variation in total biomass

Table 1. Results of partial Mantel tests for the Grandcour and the Jena datasets. The *r*-value is the partial Pearson coefficient. For species richness, it is interpreted as the correlation between pairwise distances in species richness and in biomass, once distances in species composition have been accounted for; and vice versa for species composition. Statistically significant results are in boldface type.

Experiment/month or year	Species richness		Species composition		Percentage contribution		
	<i>r</i> -value	p-value	<i>r</i> -value	p-value	Species richness	Species composition	Overlap
Jena/13	0.058	0.135	0.140	0.002	28.9%	69.3%	1.8%
Jena/16	0.148	0.003	0.112	0.001	56.3%	42.5%	1.2%
Jena/25	−0.031	0.730	0.198	0.0001	0.0%	100.0%	0%
Jena/28	0.039	0.206	0.145	0.0002	20.7%	77.6%	1.6%
Jena/37	0.089	0.051	0.113	0.003	44.1%	55.9%	0%
Jena/40	0.153	0.007	0.051	0.075	75.0%	25.0%	0%
Jena/50	0.070	0.112	0.060	0.081	54.0%	46.0%	0%
Jena/52	0.339	0.0002	−0.004	0.548	100.0%	0.0%	0%
Jena/61	0.210	0.005	0.059	0.099	78.1%	21.9%	0%
Jena/64	0.176	0.007	0.024	0.272	88.1%	11.9%	0%
Jena/73	0.368	0.0001	−0.068	0.947	100.0%	0.0%	0%
Jena/76	0.186	0.006	0.008	0.421	94.6%	3.9%	1.5
Grandcour/2008	0.032	0.215	0.150	0.001	16.6%	78.5%	5.0%
Grandcour/2009	0.014	0.353	0.102	0.002	11.6%	82.9%	5.5%

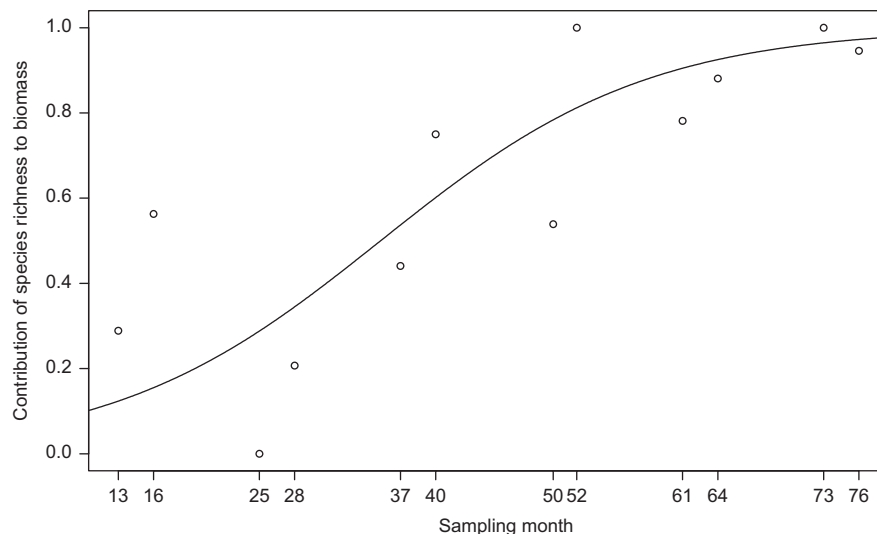


Figure 1. Evolution of the percentage contribution of species richness to total biomass in the Jena Experiment. Month 0 corresponds to the start of the experiment when plants were sown. The fitted line is the result of a generalized least-square regression with AR1 correlation structure on the logit-transformed percentages.

(14% on average) compared to floristic composition (81% on average).

In the Jena Experiment, the contribution due to species richness alone was higher (62% on average); this value was significantly larger than the average percentage in the Grandcour Experiment (difference in percentage = 48%, p -value = 0.047, permutation test with 2000 simulations). However, when comparing the percentage contributions in Grandcour with those of the first two years in Jena (to match the Grandcour Experiment with measurements during two years), the contribution was still larger in Jena, but the difference between the two experiments was less marked (average of 14% compared to 26%, respectively).

A lower contribution of species richness in the non-weeded experiment could be due to decreasing variability in species richness between the plots with time. Figure 3 presents the coefficient of variation for species richness in the Jena and in the Grandcour Experiment. As hypothesized, the decrease in the coefficient of variation from the start of the experiment for species richness was much larger in the Grandcour than in the Jena Experiment (Fig. 3).

Table 2. Temporal trends in average distances (as a measure of heterogeneity) for species composition and for species richness between all pairs of plots. Parameters are results of AR1 regression models for average distances against sampling session in the Jena Experiment. Averages are computed for plots with similar sown species richness. A positive (negative) slope indicates that species composition or species richness became more heterogeneous (homogeneous) with sampling session. Statistically significant slopes are in boldface type.

Sown species richness	Species composition		Species richness	
	Slope	p-value	Slope	p-value
2	0.0007	0.427	-0.005	0.739
4	-0.002	0.051	-0.005	0.741
8	0.005	<0.001	0.009	0.403
16	0.008	0.004	0.018	<0.001
60	0.005	0.400	-0.0003	0.963

Change in the contribution of composition at varying spatial scales

We explored if spatial scale affected the contributions of species richness and composition by analysing the Grandcour data of 2008 for each field separately, and in combinations of 2, 4, 6, 8 and 10 fields, in order to simulate increasing scale (Fig. 4). Within the 12 individual fields, species

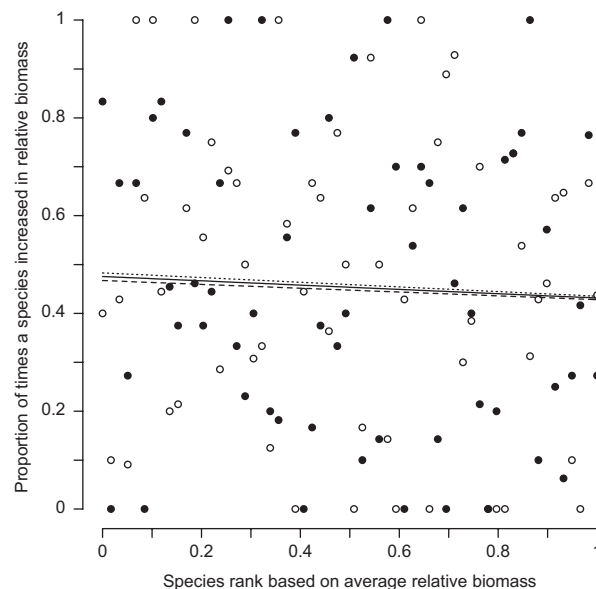


Figure 2. Relationship between the proportion of times each plant species was found to increase in biomass during the course of the Jena Experiment, and its average relative biomass in the plots where it occurred. Biomass is expressed as the rank of species ranged between 0 and 1. We considered separately spring and summer sampling periods. The solid line is the result of a generalized linear mixed-effects model with quasi-binomial family and sampling season as random factor (slope = -0.18, p = 0.64); white dots and dotted line are for the spring sampling sessions, black dots and dashed line for the summer sampling sessions.

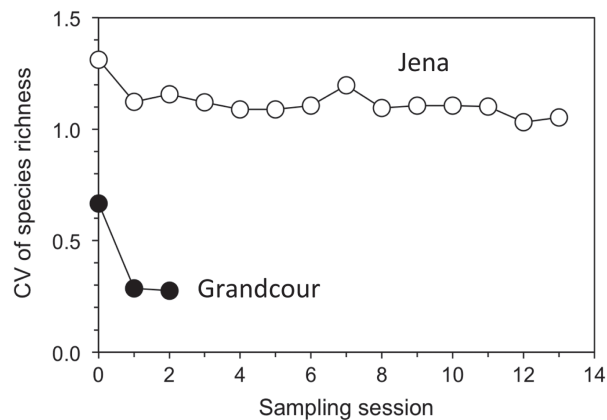


Figure 3. Coefficient of variation (CV) of species richness in the 66 plots of the Jena Experiment (open circles) and in the 144 plots of the Grandcour Experiment (solid circles) for each sampling session. Sampling session 0 refers to the initial sown species richness.

richness was significant four times at the 5% level, while plant composition was significant five times. When broadening the scale of observation, the median values of the Mantel r did not change markedly for species richness, and decreased slightly for plant composition. However, the p-values showed a clear trend: they became large and non-significant for species richness, while the opposite was apparent for plant composition. When measuring the relative contribution of species richness (Fig. 4e), we found that median values for the different numbers of fields varied around the 17% obtained for the complete dataset (Table 1); note the large variability of the estimated contributions when considering all possible combinations of subplots.

Discussion

Temporal trends in contributions of species richness and composition

Contrary to our expectations, the contribution of community composition to biomass decreased over the six years of the Jena Experiment, while the contribution of species richness increased. Thus, at the start of the experiment, plots with similar sown mixtures tended to have similar biomass, while after six years this effect vanished. Concerning species richness, plots with similar sown diversity had different productivity at the beginning of the experiment, but harboured similar biomasses already after three years. Moreover, we showed that the observed trends in the relative contributions of composition and of species richness are not mere consequences of a decrease in between-plots heterogeneity in composition and/or an increase in between-plots heterogeneity in species richness (Table 2).

These observed trends contradict our expectations, which were based on a hypothesized increase in competitive effects through time (Tilman and Downing 1994, Aarssen et al. 2003). On the one hand, plots would lose the less competitive species with time, and one could thus expect a homogenization in between-plot species richness. On the other hand, the more competitive species would increase

in biomass, and thus plots with the same dominant species would tend to have similar productivity. As said above, these expectations are rejected by our supplementary analyses (Table 2). Secondly, the dominant species in the experiment did not tend to increase in biomass with time (Fig. 2).

Several non-exclusive explanations might underlie our observations. Firstly, dominant plant species could be preferentially attacked by herbivores or other enemies, which could control their expansion (Janzen–Connell hypothesis; Sedio and Ostling 2013). Secondly, intraspecific competition may increase for a dominant species to a point at which individual fitness decreases, allowing other species to establish (Deng et al. 2012). Thirdly, complementarity effects could become more prevalent with time, which was suggested by Marquard et al. (2009) for the Jena Experiment. Notably, Roscher et al. (2013b) found an increase in functional trait diversity through time in the Jena experiment, with a change from traits associated with fast growth towards traits related to conservation of growth-related resources and successful reproduction. They concluded that assembly processes in their experiment appeared to maximize functional diversity through niche differentiation, a mechanism that did not exclude the loss of species with low competitive abilities. Fourth, one cannot, however, dismiss the effect of disturbance by repeated weeding and mowing (Wardle 2001, Roscher et al. 2009, 2013a). By creating open space, it could allow less competitive species (i.e. good colonizers in the scheme of the competition–colonization tradeoff; Cadotte 2007) to persist and possibly increase in such systems. It is however worth noting that a comparison of the weeded and non-weeded subplots in the Jena Experiment did not detect a change in functional composition between these treatments (Roscher et al. 2013b).

Note that we assume that the contribution of composition is linked to the selection effect, and that of species richness to complementarity effects. On the one hand, if selection effects are the only driver of the biodiversity–ecosystem functioning relationship, two plots with similar dominant species should both have similar biomass; at the same time, two plots with the same species richness may or may not have the same biomass. It is the presence of particular species that is important, and this will be captured by a high similarity in species composition between plots. Note however that this requires the use of quantitative (i.e. based on plant cover) measures of similarity or distance: with qualitative measures, the occasional rare species are likely to blur this effect. On the other hand, if complementarity is the key process, plots with greater numbers of species are more likely to better exploit the resources, whatever species identity, and thus attain greater biomass levels. Consequently, similarity in species richness is expected to be related to similarity in biomass. Of course, both processes are not exclusive (Loreau 2001), and this justifies the use of partial Mantel tests. In this respect, it is interesting to note that the contributions common to species richness and species composition are extremely small in the Jena Experiment (Table 1). This is expected in correctly balanced experimental designs. The greater but still moderate common contributions in the Grandcour Experiment can be attributed to the absence of weeding, and the resulting presence of non-planned colonizing species.

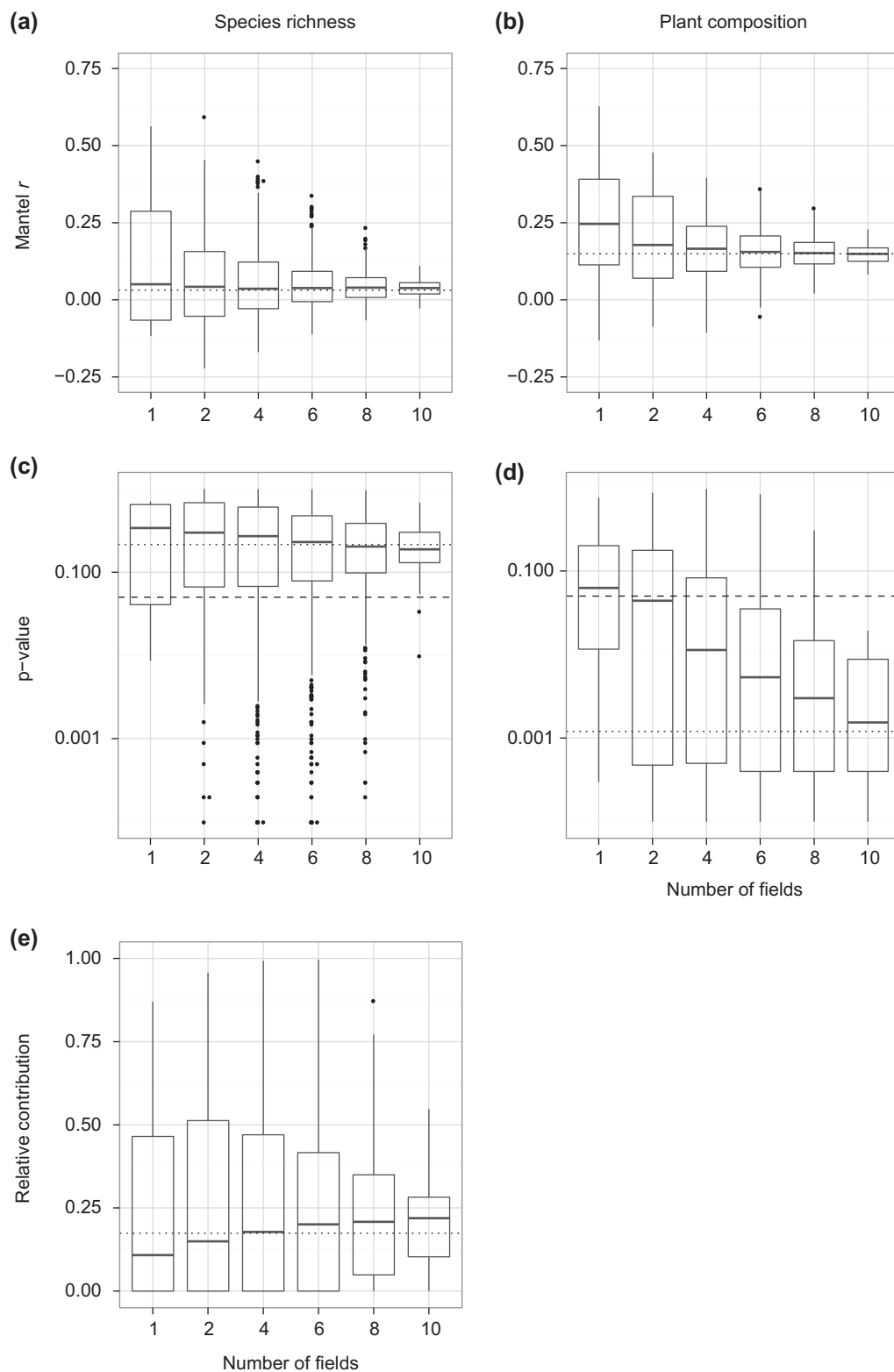


Figure 4. Results of the analyses of the effect of spatial scale. Boxplots showing the results of the partial Mantel tests for distances in plant species richness and total biomass, controlling for plant composition (left panels a and c), and for distances in plant composition and total biomass, controlling for plant species richness (right panels b and d). Panels a and b give the partial Mantel correlations r , panels c and d the corresponding p-values on a log scale. Panel e shows the relative contribution of species richness. The dotted line in each panel gives the value for the global dataset (12 fields); the dashed lines in c and d represent the 0.05 significance level. Spatial scale is represented by the x-axis, which gives the number of fields pooled in the analyses.

Differences between the Jena and the Grandcour Experiment

The comparison of the two experiments for relative contributions of composition and species richness yields results in accordance with our hypothesis. In the Grandcour Experiment, where the composition was not controlled by weeding, and thus natural colonization occurred, we found that composition is of overwhelming importance compared to species richness. In contrast, the Jena Experiment showed a more balanced contribution, with richness having a predominant importance when considered over all years. This discrepancy between experiments with and without weeding may simply arise from the much greater variability in community composition generated by the additional plant species, but also from the fast homogenization in species richness in the non-controlled experiments (Fig. 3).

A recent meta-analysis examined the relative importance of richness and composition factors in 12 controlled experiments (Hector et al. 2011). They found an equal contribution of the two similar to our results for the Jena Experiment when considered over all years. In the approach of Hector et al. (2011), variance contributions were estimated by treating species richness and species composition as random factors, and by computing the standard deviation of the posterior distributions for both variables obtained by a MCMC procedure. This method was commented on by Connolly et al. (2011), who reanalysed the data with a different method and found a greater contribution of species richness (2.7 times higher than that of species composition). This discrepancy highlights the difficulty in comparing species richness and species composition in regression models. For this reason, we believe that our simple approach with Mantel tests offers a sensible technique to estimate these contributions by treating explanatory factors in a similar way (i.e. as pairwise distances), with no assumption about the shape of the relationship between species richness and ecosystem functioning.

Change in the contribution of composition at varying spatial scales

The relative contribution of species richness to overall biomass increases when larger spatial scales are considered (i.e. when an increasing number of fields is used in the analysis) (Fig. 4e). This trend is due to a decrease in the contribution of species composition, while the importance of species richness remains almost constant (Fig. 4a–b). However, the statistical significance of the partial Mantel correlations (Fig. 4c–d) reveals that species composition becomes predominant at larger spatial scales, while, on the contrary, the significance of species richness tends to weaken. These results can be understood in the context of the experimental setting of the Grandcour Experiment: each field contains the full range of sown species richness (from 2 to 20), but the same species were sown in the three repetitions (blocks) of the four diversity levels (Supplementary material Appendix 1, Fabian et al. 2012). Thus, expanding the number of fields in the analysis does not greatly affect the range of species richness, but increases the variability in plant species compositions, firstly by increasing the β -diversity of sown species

between fields, and secondly by including a greater diversity of colonizing species (the total number of species recorded in 2008 was 113, compared to the 20 species sown). These results illustrate a general issue when comparing the contributions of different sets of variables, namely that the results are context (or scale) dependent (Wardle and Zackrisson 2005, Lepš et al. 2007, Connolly et al. 2011). Thus, the assumption that results from experiments would reflect the relative contributions of diversity and of composition in natural systems is unwarranted; also, the results from different experiments should be compared with caution. However, comparisons of contributions in the same experiment through time are exempt from this problem.

Conclusions

In all, our results firstly show that care should be taken in evaluating the relative contributions of explanatory factors across different experiments, as the context (in our case, the spatial scale) of the experiment matters (e.g. the range in species richness, the particular choice of species, the diversity of traits of the species). Secondly, we observe a positive temporal trend in the importance of species richness and, according to our assumption, of complementarity effects. This result is in line with other studies (Marquard et al. 2009) and we believe that it is ecologically important, as it indicates that biotic interactions between species may not favour particular species at the expense of others, but rather lead to more equitable communities. To uncover the processes underlying this type of evolution in plant community structure, projects like the Jena Experiment are particularly valuable. Also, conducting long-term experiments in conditions where immigration can be strictly prevented would be useful to avoid disturbances generated by the maintenance of species richness. Finally, extending research to other taxonomic groups and environments is vital to assess the generality of the mechanisms behind the biodiversity–ecosystem functioning relationship.

Acknowledgements – We thank Loïc Pellissier, Sarah M. Gray and Jan Lepš for helpful suggestions on an earlier version of the manuscript. We also are grateful for the participation of the farmers from Grandcour and all the help provided by friends and family members.

Funding – This study was supported by the Swiss National Science Foundation (Grant 31003A_138489 to LFB).

References

- Aarssen, L. W. 2001. On correlations and causations between productivity and species richness in vegetation: predictions from habitat attributes. – *Basic Appl. Ecol.* 2: 105–114.
- Aarssen, L. W. et al. 2003. Is the productivity of vegetation plots higher or lower when there are more species? Variable predictions from interaction of the “sampling effect” and “competitive dominance effect” on the habitat templet. – *Oikos* 102: 427–432.
- Assaf, T. et al. 2011. The relationship between plant diversity and productivity in natural and in managed grassland. – *Appl. Ecol. Environ. Res.* 9: 157–166.

- Bates, D. 2010. Linear mixed model implementation in lme4. – *Statistics* 2010: 1–32.
- Bell, T. et al. 2009. A linear model method for biodiversity–ecosystem functioning experiments. – *Am. Nat.* 174: 836–849.
- Borcard, D. et al. 1992. Partialling out the spatial component of ecological variation. – *Ecology* 73: 1045–1055.
- Bruggisser, O. T. et al. 2012. Direct and indirect bottom–up and top–down forces shape the abundance of the orb-web spider *Argiope bruennichi*. – *Basic Appl. Ecol.* 13: 706–714.
- Cadotte, M. W. 2007. Competition–colonization tradeoffs and disturbance effects at multiple scales. – *Ecology* 88: 823–829.
- Cardinale, B. J. et al. 2012. Biodiversity loss and its impact on humanity. – *Nature* 486: 59–67.
- Chapin III, F. S. et al. 1997. Biotic control over the functioning of ecosystems. – *Science* 277: 500.
- Connolly, J. et al. 2011. Comment on: BUGS in the analysis of biodiversity experiments: species richness and composition are of similar importance for grassland productivity. – *PLoS ONE* 6: e17434.
- Connolly, J. et al. 2013. An improved model to predict the effects of changing biodiversity levels on ecosystem function. – *J. Ecol.* 101: 344–355.
- Deng, J. et al. 2012. Models and tests of optimal density and maximal yield for crop plants. – *Proc. Natl Acad. Sci. USA* 109: 15823–15828.
- Fabian, Y. et al. 2012. Diversity protects plant communities against generalist molluscan herbivores. – *Ecol. Evol.* 2: 2460–2473.
- Fabian, Y. et al. 2013. The importance of landscape and spatial structure for hymenopteran-based food webs in an agro-ecosystem. – *J. Anim. Ecol.* 82: 1203–1214.
- Fox, J. W. 2006. Using the Price equation to partition the effects of biodiversity loss on ecosystem function. – *Ecology* 87: 2687–2696.
- Fox, J. W. and Harpole, W. S. 2008. Revealing how species loss affects ecosystem function: the trait-based price equation partition. – *Ecology* 89: 269–279.
- Gamfeldt, L. et al. 2008. Multiple functions increase the importance of biodiversity for overall ecosystem functioning. – *Ecology* 89: 1223–1231.
- Gelman, A. 2005. Analysis of variance? – Why it is more important than ever. – *Ann. Stat.* 33: 1–53.
- Goslee, S. and Urban, D. 2007. The ecodist package: dissimilarity-based functions for ecological analysis. – *J. Stat. Softw.* 22: 1–19.
- Grace, J. B. et al. 2007. Does species diversity limit productivity in natural grassland communities? – *Ecol. Lett.* 10: 680–689.
- Harmon, L. J. and Glor, R. E. 2010. Poor statistical performance of the Mantel test in phylogenetic comparative analyses. – *Evolution* 64: 2173–2178.
- Harvey, E. et al. 2013. Identity effects dominate the impacts of multiple species extinctions on the functioning of complex food webs. – *Ecology* 94: 169–179.
- Hector, A. et al. 1999. Plant diversity and productivity experiments in European grasslands. – *Science* 286: 1123–1127.
- Hector, A. et al. 2011. BUGS in the analysis of biodiversity experiments: species richness and composition are of similar importance for grassland productivity. – *PLoS ONE* 6: e17434.
- Hooper, D. U. et al. 2005. Effects of biodiversity on ecosystem functioning: a consensus of current knowledge. – *Ecol. Monogr.* 75: 3–35.
- Hooper, D. U. et al. 2012. A global synthesis reveals biodiversity loss as a major driver of ecosystem change. – *Nature* 486: 105–108.
- Isbell, F. et al. 2015. Biodiversity increases the resistance of ecosystem productivity to climate extremes. – *Nature* 526: 574–577.
- Kahmen, A. et al. 2005. Effects of plant diversity, community composition and environmental parameters on productivity in montane European grasslands. – *Oecologia* 142: 601–615.
- Kirwan, L. et al. 2007. Evenness drives consistent diversity effects in intensive grassland systems across 28 European sites. – *J. Ecol.* 95: 530–539.
- Lawton, J. H. 1994. What do species do in ecosystems? – *Oikos* 71: 367–374.
- Legendre, P. and Legendre, L. 1998. Numerical ecology. – Elsevier.
- Lepš, J. et al. 2007. Long-term effectiveness of sowing high and low diversity seed mixtures to enhance plant community development on ex-arable fields. – *Appl. Veg. Sci.* 10: 97–110.
- Loreau, M. 2001. Biodiversity and ecosystem functioning: current knowledge and future challenges. – *Science* 294: 804–808.
- Mantel, N. 1967. The detection of disease clustering and a generalized regression approach. – *Cancer Res.* 27: 209–220.
- Marquard, E. et al. 2009. Plant species richness and functional composition drive overyielding in a six-year grassland experiment. – *Ecology* 90: 3290–3302.
- Mouillot, D. et al. 2011. Functional structure of biological communities predicts ecosystem multifunctionality. – *PLoS ONE* 6: e17476.
- Oksanen, J. et al. 2015. vegan: community ecology package. – R package ver. 2.2–1.
- Pinheiro, J. et al. 2013. nlme: linear and nonlinear mixed effects models. – R packag. ver. 1–86.
- Roscher, C. et al. 2004. The role of biodiversity for element cycling and trophic interactions: an experimental approach in a grassland community. – *Basic Appl. Ecol.* 121: 107–121.
- Roscher, C. et al. 2005. Overyielding in experimental grassland communities - irrespective of species pool or spatial scale. – *Ecol. Lett.* 8: 419–429.
- Roscher, C. et al. 2009. Community assembly and biomass production in regularly and never weeded experimental grasslands. – *Acta Oecol.* 35: 206–217.
- Roscher, C. et al. 2013a. What happens to the sown species if a biodiversity experiment is not weeded? – *Basic Appl. Ecol.* 14: 187–198.
- Roscher, C. et al. 2013b. A functional trait-based approach to understand community assembly and diversity–productivity relationships over 7 years in experimental grasslands. – *Perspect. Plant Ecol. Evol. Syst.* 15: 139–149.
- Sandau, N. et al. 2014. Including community composition in biodiversity–productivity models. – *Meth. Ecol. Evol.* 5: 815–823.
- Sandau, N. et al. 2016. Data from: The relative contribution of species richness and species composition to ecosystem functioning. – Dryad Digital Repository, <<http://dx.doi.org/10.5061/dryad.44bm6>>.
- Tilman, D. G. and Downing, J. A. 1994. Biodiversity and stability in grasslands. – *Nature* 367: 363–365.
- Tilman, D. G. et al. 1996. Productivity and sustainability influenced by biodiversity in grassland ecosystems. – *Nature* 379: 718–720.
- Venables, W. N. and Ripley, B. D. 2002. Modern applied statistics with S. – Springer.
- Wardle, D. A. 2001. Experimental demonstration that plant diversity reduces invasibility – evidence of a biological mechanism or a consequence of sampling effect. – *Oikos* 95: 161–170.
- Wardle, D. A. 2016. Do experiments exploring plant diversity–ecosystem functioning relationships inform how biodiversity loss impacts natural ecosystems? – *J. Veg. Sci.* 27: 646–653.

- Wardle, D. A. and Zackrisson, O. 2005. Effects of species and functional group loss on island ecosystem properties. – *Nature* 435: 806–810.
- Weigelt, A. et al. 2010. The Jena Experiment: six years of data from a grassland biodiversity experiment. – *Ecology* 91: 930–931.
- Zavaleta, E. S. et al. 2010. Sustaining multiple ecosystem functions in grassland communities requires higher biodiversity. – *Proc. Natl Acad. Sci. USA* 107: 1443–1446.
- Zuur, A. F. et al. 2009. Mixed effects models and extensions in ecology with R. – Springer.

Supplementary material (available online as Appendix oik-03901 at <www.oikosjournal.org/appendix/oik-03901>). Appendix 1.

Sandau, N., Fabian, Y., Bruggisser, O. T., Rohr, R. P., Naisbit, R. E., Kehrli, P., Aebi, A. and Bersier, L.-F. 2016. The relative contribution of species richness and species composition to ecosystem functioning. – Oikos doi: 10.1111/oik.03901

Appendix 1

Description of field methods for the Grandcour Experiment

We set up a multitrophic experiment in spring 2007 (May–June) by sowing experimental wildflower strips. The experimental fields were located around the village of Grandcour, Switzerland (46° 52' N, 6° 56' E, elevation 479 m a.s.l.). According to the Bundesamt für Meteorologie und Klimatologie MeteoSchweiz the annual precipitation in the region is 941 mm while the local mean annual temperature is 10.1°C. The landscape is a mosaic, typical for the Swiss lowland, where agricultural elements are interspersed with small traditional fruit orchards, gardens, forest patches and grassland. Soils can be classified as cambisol/arenosol and calcaric cambisols (Ducommun Dit-Boudry 2010).

Experimental design

The 12 wildflower strips (72×9 m or 108×6 m, 648 m^2) were each equally divided into three trophic treatments: 1) control (*C*, without fence), 2) vertebrate predator exclusion (*PE*, 25 mm mesh fence), and 3) vertebrate predator and major herbivore exclusion (*PHE*, 8 mm mesh fence). The full details of the experimental design are already published, and can be found in Bruggisser et al. (2012) and Fabian et al. (2012). Within each field treatment order was randomly assigned. We further divided each treatment into 4 subplots of 54 m^2 (6×9 m), which we then assigned to one of four plant species richness levels. This yielded a total of 144 subplots.

The species pool for the seed mixtures contained 20 plant species that are part of the wildflower mixture used in Swiss nature restoration schemes (Haaland et al. 2011, Haaland and Bersier 2011). The mixtures consisted only of species of the tall herb functional group (Roscher et al. 2004) (Table S1), as we excluded two legumes, a small herb species, and an exotic species, which are present in the conventional mixture. We assembled the four plant richness levels of 2, 6, 12 and 20 with regard to equal frequency by constrained random draw. Within each experimental wildflower strip each diversity level was repeated three times, in each trophic treatment once, so

that sown composition was kept constant within fields, but varied among fields. While the 20 species mixture was reproduced in all wildflower strips, the other diversity levels were represented by 12 different experimentally sown mixtures. In total this resulted in 37 different sown mixtures.

To assemble the experimental mixtures, we applied a substitutive design (Jolliffe 2000) where each plant species was added in the same proportion to result in maximum evenness. Taking into account the individual germination rates of the species predicted by the commercial seed provider (UFA Samen Lyssach, Switzerland), we corrected seed density to 1000 germinable seeds m^{-2} (Roscher et al. 2004).

Fields were prepared similar to conventional wildflower strips. The farmers harrowed the fields twice before sowing and sprayed them with the non-selective herbicide Roundup® (Glyphosate) to eliminate weeds and to reduce establishment from the seed bank. We scattered the seeds by hand; therefore soya groats were added as bulking agent to facilitate an even distribution within the subplots. As several studies have suggested that weeding might increase invasion (Wardle 2001, Roscher et al. 2009), we kept weeding to a minimum. Further, constant weeding would have disturbed other projects in our study system concerned with the fauna (Fabian et al. 2012, Bruggisser et al. 2012). Only in 2007 we removed *Chenopodium album* and *Amaranthus retroflexus* to avoid competition for light with germinating seeds. To prevent spread to adjacent fields and associated loss of environmental subsidies for the farmer, we removed the harmful agricultural weeds *Rumex obtusifolium* and *Cirsium vulgare* throughout the whole experiment. Besides one field, where due to high pressure of *Echinochloa crus-galli* mowing was recommended in the first year, subplots were not mown. The absence of weeding and mowing distinguishes the Grandcour Experiment from other studies such as BIODDEPTH, the Jena Experiment or Cedar Creek, where grassland plots were cut or burned according to the management regime of the surrounding region (Hector et al. 1999, Roscher et al. 2004, Tilman et al. 2006).

We thus manipulated plant diversity by sowing different diversity levels but permitted establishment from external species; hence the communities consisted of the selected sown species complemented by locally existing species filling empty space.

Plant community assessment

We evaluated plant communities in a randomly selected 2×2 m quadrat in each subplot, taking care that it was placed at least 1 m from the border. Except for some vegetative Poaceae, all vascular plants were identified to species level. We estimated individual species cover (%) visually each year in early autumn. Community coverage may exceed 100% due to plants overlap. Total species richness S in the subplots varied from 6 to 42 species, with a mean of 22.8.

LAI measurements and biomass estimation

Aboveground biomass is often used as a substitute for (aboveground) productivity. Measuring biomass is generally done by clipping plant material, which is the most accurate method. However, it was difficult to apply in our experiment due to large amounts of biomass. Moreover, we wanted to avoid disturbance to the projects dealing with higher trophic levels. Therefore we measured the leaf area index (LAI) with a plant canopy analyser, which is an adequate alternative non-destructive method for estimating biomass. The canopy analyser measures light attenuation by the vegetation and a standard coefficient is then used to derive the LAI. In each subplot, we recorded above-canopy reference measurements and 24 below-canopy measurements at ground level in early fall 2008 and 2009. We covered the fish-eye sensor with a 270° black cap, which helped to account for the small subplot size and excluded the effect of the person taking the measurements, while it also prevented direct sunlight shining into the lens. We did not adjust measurements for leaf angles. We then calibrated the LAI measures with five randomly selected quadrats of 30 × 30 cm in eight subplots. Within the quadrats, all plants were clipped at ground level, stored in a paper bag and then dried to constant weight at 60 C. We established a linear regression between aboveground biomass and LAI in the eight subplots (Pearson product-moment correlation $r = 0.89$), which was used to convert average LAI measures to total aboveground biomass (dry weight $\text{g} \cdot \text{m}^{-2}$). Aboveground biomass TB ranged from 550.6 $\text{g} \cdot \text{m}^{-2}$ to 1799.2 $\text{g} \cdot \text{m}^{-2}$ with a mean of 1098.3 $\text{g} \cdot \text{m}^{-2}$. Note that our estimate of biomass is closely related to the "annual net primary production" (ANPP) measure that is typically analysed. As we only determined total community biomass, and because we did not establish monocultures for species arriving from the seed bank, analyses of biodiversity effects using additive partitioning to separate complementarity and selection effects (Loreau and Hector 2001, Fox 2006) were not feasible.

References

- Bruggisser, O. T. et al. 2012. Direct and indirect bottom-up and top-down forces shape the abundance of the orb-web spider *Argiope bruennichi*. – Basic Appl. Ecol. 13: 706–714.
- Ducommun Dit-Boudry, S.A.-L. 2010. Influence de la diversité floristique sur les paramètres biologiques et chimiques du sol dans les jachères florales, Neuchâtel.
- Fabian, Y. et al. 2012. Diversity protects plant communities against generalist molluscan herbivores. – Ecol. Evol. 2: 2460–2473.
- Fox, J. W. 2006. Using the Price equation to partition the effects of biodiversity loss on ecosystem function. – Ecology 87: 2687–2696.
- Haaland, C. and Bersier, L.-F. 2011. What can sown wildflower strips contribute to butterfly conservation? An example from a Swiss lowland agricultural landscape. – J. Insect Conserv.

15: 301–309.

Haaland, C. et al. 2011. Sown wildflower strips for insect conservation: a review. – *Insect Conserv. Divers.* 4: 60–80.

Hector, A. et al. 1999. Plant diversity and productivity experiments in European grasslands. – *Science* 286: 1123–1127.

Jolliffe, P. 2000. The replacement series. – *J. Ecol.* 88: 371–385.

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

Roscher, C. et al. 2004. The role of biodiversity for element cycling and trophic interactions: an experimental approach in a grassland community. – *Basic Appl. Ecol.* 121: 107–121.

Roscher, C. et al. 2009. Community assembly and biomass production in regularly and never weeded experimental grasslands. – *Acta Oecol.* 35: 206–217.

Tilman, D. G. et al. 2006. Biodiversity and ecosystem stability in a decade-long grassland experiment. – *Nature* 441: 629–632.

Wardle, D. A. 2001. Experimental demonstration that plant diversity reduces invasibility – evidence of a biological mechanism or a consequence of sampling effect. – *Oikos* 95: 161–170.

Table 1. List of sown plant species.

Species name
<i>Achillea millefolium</i> L.
<i>Agrostemma githago</i> L.
<i>Anthemis tinctoria</i> L.
<i>Centaurea cyanus</i> L.
<i>Centaurea jacea</i> L.
<i>Cichorium intybus</i> L.
<i>Daucus carota</i> L.
<i>Dipsacus fullonum</i> L.
<i>Echium vulgare</i> L.
<i>Hypericum perforatum</i> L.
<i>Leucanthemum vulgare</i> LAM.
<i>Malva moschata</i> L.
<i>Malva sylvestris</i> L.
<i>Origanum vulgare</i> L.
<i>Papaver rhoeas</i> L.
<i>Pastinaca sativa</i> L.
<i>Silene latifolia</i> POIR.
<i>Tanacetum vulgare</i> L.
<i>Verbascum lychnitis</i> L.
<i>Verbascum thapsus</i> L.