OIKOS 92: 123–134. Copenhagen 2001

Separating the chance effect from other diversity effects in the functioning of plant communities

Jan Lepš, Valerie K. Brown, Tomas A. Diaz Len, Dagmar Gormsen, Katarina Hedlund, Jana Kailová, Gerard W. Korthals, Simon R. Mortimer, Claudino Rodriguez-Barrueco, Jacques Roy, Ignacio Santa Regina, Cornelis van Dijk and Wim H. van der Putten

Lepš, J., Brown, V. K., Diaz Len, T. A., Gormsen, D., Hedlund, K., Kailová, J., Korthals, G. W., Mortimer, S. R., Rodriguez-Barrueco, C., Roy, J., Santa Regina, I., van Dijk, C. and van der Putten, W. H. 2001. Separating the chance effect from other diversity effects in the functioning of plant communities. – Oikos 92: 123–134.

The effect of plant species diversity on productivity and competitive ability was studied in an experiment carried out simultaneously in five European countries: Czech Republic (CZ), the Netherlands (NL), Sweden (SE), Spain (SP), and United Kingdom (UK). The aim was to separate the 'chance' or 'sampling effect' (increasing the number of sown species increases the probability that a species able 'to do a job' will be included) from the complementarity effect (species-rich communities are better able to exploit resources and to take care of ecosystem functions than species-poor communities). In the experiment, low diversity (LD) and high diversity (HD) mixtures of grassland species were sown into fields taken out of arable cultivation. The HD mixture consisted of five grass species, five legumes and five other forbs. The LD mixtures consisted of two grasses, one legume and one other forb, with different plant species combinations in each replicate block. The design of the experiment was constructed in such a way that the total number of seeds of each species over all the replications was exactly the same in HD and LD treatments, and the total number of grass seeds, leguminous seeds and other forb seeds were the same in both LD and HD. The responses measured were the total above-ground biomass (as a measure of productivity) and the average number of naturally establishing species in a plot (as a measure of the competitive ability of the mixture), both measured in the third year of the experiment.

The results show that, on average, the HD plots performed better (i.e., attained higher biomass, had better weed suppression), but that the best LD mixture was as good as the best HD mixture. On the contrary, the worst LD mixture was always less successful than the worst HD replicate. The performance of particular species in the HD mixtures was a good predictor of the success of a certain species combination in a LD mixture (explaining 61% of variability between particular LD mixtures). In all sites, the LD mixture composed of species which were the most abundant in HD mixtures was as efficient in suppressing weeds as the HD mixture itself. It is argued that the performance of a species assemblage is influenced mostly by the identity of species and the diversity effect is mainly due to the 'chance' or 'sampling' effect: with increasing number of species the probability that an important species will be included in the mixture increases. Caution is urged in interpreting experiments with manipulated diversity and the possible limitations of such experiments are discussed.

J. Lepš and J. Kailová, Dept of Botany, Faculty of Biological Sciences, Univ. of South Bohemia (JL also at Inst. of Entomology), Czech Academy of Sciences, Branišovská 31, CZ-370 05 České Budějovice, Czech Republic (suspa@tix.bf.jcu.cz). – V. K. Brown and S. R. Mortimer, Dept of Agriculture, Univ. of Reading, Early Gate, P.O. Box 236, Reading, UK RG6 6AT. – T. A. Diaz Len, C. Rodriguez-Barrueco and I. Santa Regina, Instituto de Recursos Naturales y Agrobiologia CSIC, Salamanca, Spain. – D. Gormsen and K. Hedlund, Dept of Animal Ecology, Lund Univ., Ecology Building, SE-22362 Lund, Sweden. – G. W. Korthals, C. van Dijk, and W. H. van der Putten, Netherlands Inst. of Ecology, P.O. Box 40, NL-6666 ZG, Heteren, the Netherlands. – J. Roy, Centre d'Ecologie Fonctionnelle et Evolutive (CNRS), F-34293 Montpellier Cedex 5, France.

Accepted 7 September 2000

ISSN 0030-1299

The belief that 'diversity begets stability' can be traced back to Darwin (1859) and was one of the leading ideas of ecology in the fifties, sixties and early seventies of the past century (MacArthur 1955, Elton 1958, Odum 1971). However, even during this time the belief was questioned; see, for example, the discussion between McNaughton (1967) and Singh and Misra (1969). May's book (May 1973) showed, albeit using simplified and unrealistic models, that stability is not a necessary consequence of complexity. Theoretical models can also be constructed that predict a positive stability-diversity relationship (Doak et al. 1998); however, this prediction may disappear with a slight change in the underlying assumptions (Tilman et al. 1998). The diversity of a community may affect various ecosystem functions, including, for example, productivity (Naeem et al. 1996, Tilman et al. 1997a, b) or resistance to invasions (Rejmánek 1996). The causality of such relationships is, however, often questioned. From the early eighties, the need for precise formulation of hypotheses (Pimm 1984) and for experimental assesment of the relationships (e.g. Van Voris et al. 1980) increasingly dominate studies of diversity effects.

Several empirical studies (e.g. Lepš et al. 1982, Lepš 1990, MacGillivray et al. 1995, Grime 1997, Hooper and Vitousek 1997, 1998) have demonstrated that ecosystem function is mainly a consequence of the prevailing strategies of constituent species, in interaction with the abiotic environment. Generally, the presence of keystone species will influence the ecosystem functioning more than number of species per se (Mooney 1997, Roy 2001). Even some of the experimental effects that are ascribed to diversity are probably a consequence of the effect of abiotic conditions and species ecophysiology (Huston 1997). However, the fact that species characteristics are more important than diversity per se does not mean that diversity has no effect on ecosystem function. Indeed, the majority of ecologists are convinced about the importance of diversity for ecosystem processes (Schläpfer et al. 1999). The evidence supporting the diversity effect is now the focus of novel ecological research and is intensively debated (Johnson et al. 1996, Huston 1997, Grime 1998, Hodgson et al. 1998, Lawton et al. 1998, Tilman 1999, Hector et al. 1999, Naeem et al. 1999, Naeem 2000, Wardle et al. 2000).

Many ecosystem functions are connected with productivity and, because of the strong effect of productivity on diversity (Al-Mufti et al. 1977), it is clear that the correlative approach (i.e. comparing natural communities differing in their species richness) has strong limitations, and that the question has to be tackled experimentally. If two communities in nature differ in their species richness, they also differ in other characteristics. For example, with increasing nutrient load in meadows, the species diversity usually decreases. However, there is a concurrent replacement of stress tolerant species (sensu Grime 1979) by species with a competitive strategy. As a consequence, diversity and composition are confounded and it is difficult or impossible to separate their effects. Consequently, the effects of diversity are usually tested using artificially created assemblages of species. The number of species sown or planted is manipulated directly, and sometimes controlled by hand-weeding (e.g. Tilman 1997, Hector et al. 1999). Nevertheless, even in manipulative experiments, any change in species richness can be achieved only by changing species composition, and consequently also (at least some) functional traits.

Experimental manipulations can confound the effect of species identity with the effects of diversity per se (Huston 1997). Some species are often better at 'doing the job' than others. Therefore, experiments need to be designed carefully to avoid the possibility that the effect of species is mistakenly interpreted as an effect of diversity. The best solution is an experimental design in which the low diversity treatments are composed of various subsets of species in the high diversity treatment. The design should be such that each species is equally important at all the diversity levels examined. In addition, some species are more similar to each other than to other species, leading to the concept of functional groups. For example, two narrow-leaved grasses are functionally more similar to each other than grasses and forbs. Consequently, the diversity effect (if any) has a hierarchical nature (e.g. the diversity of functional groups, species diversity within functional groups, etc.).

There are two possible basic mechanisms for the effect of diversity: (1) the 'chance' or 'sampling' effect (Aarssen 1997, Huston 1997, Tilman et al. 1997a, b) and (2) the 'complementarity' effect (Loreau 1998a):

(1) The chance effect is simply a probabilistic effect: with increasing numbers of species present, there is an increasing probability that the species mixture will include the species which are able to 'do the job' (e.g. species which are better able to suppress other species, are more efficient in resource capture, etc.). If the increase in average productivity is due the chance effect, then the best of the monocultures or low diversity communities should be as efficient as the high diversity communities. A necessary condition for the sampling effect is interspecific differences in ability to capture resources.

(2) The complementarity effect results from interspecific differences in resource partitioning. This can be the case when species compete for more than a single limiting resource, and each of the species is better in exploitation of one or other of the resources, or each of them has a different response curve to some environmental factor which varies over space or time. In this case, high diversity mixtures should be more efficient at resource capture than the best of the monocultures or low diversity mixtures.

These two mechanisms are not mutually exclusive and probably operate simultaneously. This complicates their testing as separate hypotheses. In fact, the idea of complementarity is much older than Loreau (1998a). It has been long popular in agricultural research, as one of the bases of intercropping (Vandermeer 1989), and was a frequent topic of discussions, both from agricultural and ecological points of view (e.g. de Wit 1960, van den Bergh and Elberse 1970, Trenbath 1974). The goal of our study was to test experimentally if there is any effect of plant species diversity on the functioning of artificially-created grassland communities. In particular, we tested whether high diversity mixtures of grassland species are more efficient than low diversity mixtures in the suppression of arable weeds present in the seed bank, and if high diversity mixtures attain higher biomass. We tested if there are differences between averages of the diversity treatments, and whether the 'best' low diversity communities are comparable with the high diversity communities. Finally, to demonstrate the importance of species identity, we tested the ability to predict the performance of particular low diversity mixtures based on the performance of their constituent species in high diversity mixtures. This enabled us to evaluate the importance of species identity in diversity experiments.

In ecological experimentation, there are several trade-offs (Diamond 1986). Carefully designed greenhouse or growth-chamber experiments are easily replicated, with control of external environmental variables, but their relevance for the functioning of real ecosystems can be low. On the other hand, field experiments are prone to uncontrollable fluctuations in the environment, which decreases their replicability (Lepš et al. 1999). A reasonable way to deal with this problem is to have replications over a wide range of environmental conditions. The present study was part of a large project on enhancement of ecosystem development at abandoned arable fields throughout Europe (Van der Putten et al. 2000). Consequently, our experiments were carried out simultaneously on formerly cultivated land in five different European countries. The locations may, therefore, be regarded as independent replicates. We constructed communities with the same composition of functional groups, but differing in species diversity within functional groups. The functional groups used for the purpose of this study were grasses, legumes, and non-leguminous forbs.

Methods

Identical experiments were carried out on five formerly cultivated sites, one in each of five European countries (Table 1, see also Van der Putten et al. 2000). The last crop was harvested from each site in 1995, the sites were then ploughed or cultivated and in spring 1996 the treatments were established. At each site, the experiment consisted of five randomised complete blocks, each block consisting of four 10-m × 10-m plots, each with one of the following treatments: continued crop rotation (CCR), natural colonisation (NC), low diver-

Table 1. Basic characteristics of experimental sites. Climatic data are long-term averages.

	Czech Republic	Netherlands	Sweden	Spain	United Kingdom
Abbreviation	CZ	NL	SE	SP	UK
Site	Benešov	Mossel	Trolleholm	Munovela	Bradenham
Coordinates	49.92N 15.00E	52.04N 05.45E	55.45N 13.15E	40.54N 5.45W	51.40N 0.48W
Altitude [m a.s.l.]	659	30	85	840	140
Mean temperature [°C]	6.4	9.4	7.5	10.8	9.6
Warmest month [°C]	July (16.4)	July (22.1)	July (17.1)	August (29.9)	July (16.5)
Coldest month [°C]	Jan. (−2.7)	Jan. (4.3)	Jan. (−0.9)	Jan. (0.8)	Jan. (3.6)
Average rainfall [mm yr^{-1}]	680	840	700	500	750
Wettest month [mm yr^{-1}]	July. (78)	Aug. (130)	July. (66)	Nov. (99)	Oct. (65)
Driest month [mm yr^{-1}]	Feb. (36)	Feb. (75)	Feb. (33)	June (17)	Feb. (41)
Soil texture	Loam	Sandy loam	Clay	Loamy clay	Loam
Soil classification	Brown soil	Brown soil	Brown soil	Chromic luvisol	Brown rendzina (chalk)
Surrounding vegetation	Arable land	Heath, mixed forest, abandoned arable land	Deciduous forest, cultured field	Dehesa-like woodland (*Quercus rotundifolia*)	Chalk grassland
soil pH H_2O	5.88	6.42	6.61	7.84	7.97
N-total [mg kg^{-1}]	1538	1330	1850	731	2000

Table 2. Seed densities [seeds m^{-2}] sown in particular blocks of LD treatments (LD1 to LD5) and density in all the HD plots. CZ site used as an example.

Group	Species	LD1	LD2	LD3	LD4	LD5	HD
Grasses	*Cynosurus cristatus*	0	1250	0	1250	0	500
	Festuca rubra	0	0	1250	0	1250	500
	Holcus lanatus	1250	0	0	1250	0	500
	Phleum pratense	1250	0	0	0	1250	500
	Trisetum flavescens	0	1250	1250	0	0	500
Legumes	*Lathyrus pratensis*	0	0	500	0	0	100
	Lotus corniculatus	0	0	0	0	500	100
	Medicago lupulina	0	500	0	0	0	100
	Trifolium dubium	0	0	0	500	0	100
	Trifolium pratense	500	0	0	0	0	100
Other forbs	*Centaurea jacea*	500	0	0	0	0	100
	Galium verum	0	0	0	500	0	100
	Lychnis flos-cuculi	0	0	500	0	0	100
	Plantago lanceolata	0	0	0	0	500	100
	Prunella vulgaris	0	500	0	0	0	100

sity seed mixture (LD), or high diversity seed mixture (HD). Only results from the NC, LD and HD treatments are presented in this paper. In the NC treatment, the plots were abandoned and allowed to colonise naturally. In the LD plots, a mixture of grassland species was sown, consisting of two grasses (density of each species 1250 seeds m^{-2}), one legume and one other forb species (both 500 seeds m^{-2}). In the HD plots the sown mixture consisted of five grass species (density of each species 500 seeds m^{-2}), five legumes and five other forbs (density of each species 100 seeds m^{-2}). Consequently, in all the sown plots, the density of added grasses was 2500 seeds m^{-2}, and of legumes and other forbs, 500 seeds m^{-2}. To avoid confounding the species identity effect with the diversity effect (Huston 1997), each low diversity treatment consisted of a different mixture of species. Each LD treatment consisted of four species that were a subset of the 15 species sown in the HD treatment, so that the total number of seeds of each species sown in all LD treatment plots was exactly the same as that sown in HD plots. In the Netherlands, the large seeded species *Vicia cracca* was sown at 20% of the standard rates. The design of the experiment is illustrated by the arrangement in the Czech site (Table 2).

In each 10-m × 10-m plot, twelve permanent 1-m^2 quadrats were sampled each year, with cover of all the species estimated on an ad hoc six-point scale (1: below 1%, 2: 1–4%, 2: 4–10%, 4: 10–25%, 5: 25–50%, 6: 50–100%). The scale is not linear, but roughly logarithmic. As a result of this, the arithmetic mean of cover scores corresponds roughly to the geometric mean of cover values. This value might be considered a reasonably good indication of central tendency (as cover values are usually log-normally distributed). As the goal of this paper is to distinguish between various forms of diversity effect, we use here only results from the third year of the experiment, i.e. 1998, by which time the communities had stabilised to a certain extent. The detailed description of the changes in productivity and species composition during the whole experiment is presented elsewhere (Van der Putten et al. 2000, Santa Regina et al. unpubl.).

Two parameters characterising ecosystem functioning were analysed: (1) the total above-ground biomass at the peak of the vegetation season (as a measure of the productivity of the plant community) and (2) the number of naturally colonising (weed) species in a plot (as a measure of the competitive strength of the sown mixtures). The sown species were not included in the number of colonisers, even in NC and LD plots where they were not sown. The above-ground biomass was harvested in twelve 0.25-m × 0.25-m quadrats per 10-m × 10-m plot, situated adjacent to the permanent quadrats used for cover assessment. The samples were oven-dried at 80°C, weighed, and values expressed as dry mass per m^2. The number of colonising species is estimated on the basis of analyses of species composition of twelve permanent quadrats in each 10-m × 10-m plot.

The difference between the HD and LD treatments was of primary interest (the fact that sowing grassland plants suppresses arable weeds is trivial). We were interested in three comparisons: the comparison of the average performance of LD and HD treatments, comparison of the best LD and HD replicates, and comparison of the worst LD and HD replicates. For the comparison of average LD and HD performances, ANOVA (with block as a random factor) was first used for each country. ANOVA was then performed with country as a random factor for all the countries together. Because we were interested in general results from our experiment, ‘country’ was taken as a factor with random effect (with country as a fixed factor, the population about which the statistical interference is made would be our experimental plots alone). For comparisons of maximum and minimum performance, we did not have replication within each country, so the countries were taken as replicates (this gives a test with roughly the same power as that with country as a random factor in an ANOVA for the average values). Furthermore, we used maximum, average and minimum performance in each country as a within-country factor and calculated the ANOVA to see whether there is an interaction between treatment and performance.

If there is an important species identity effect, then the performance of a particular LD mixture should be predictable from the species performances in the HD plots. To test this, we calculated the average 'importance' value for each of the sown species (average value of cover abundance) over all the HD plots. For each LD mixture we then calculated the sum of importance values of its constituent species, and used these values as predictors of the performance of each individual LD mixture. The species' importance in HD plots should be proportional to its competitive ability, and, consequently, the total for the particular LD species combination should be a good predictor of its ability to suppress the other species. The ability to predict the performance of a LD mixture reflects the species specific component in differences between LD mixtures.

The difference in the species composition of sown seeds between the HD and any of LD treatments is the same. However, because of competitive interactions (both within the set of sown species and between sown species and non-sown weed species), the resulting species composition of HD characterised by the species cover is expected to be most similar to the LD composed of the strongest competitors. We used PCA (calculated using the CANOCO for Windows package; ter Braak and Šmilauer 1998) to describe similarity of sown species composition in particular plots.

We also tried to explain differences in the responses to the treatments between particular countries. We expected that the effect of competition from sown species on the number of colonising species in the community would be less pronounced if a strong dominant is already present among the colonisers. Consequently, we calculated a simple index of dominance for the NC plots as follows. First, each species was characterised by its average cover value over all the NC plots, C_i. The dominance was then expressed as $D_1 = C_1/C$, where C_1 is the average cover degree of the most abundant species and $C = \Sigma\ C_i$. This dominance value was then correlated with the relative decrease of weed species number in HD plots.

Results

In all countries, there was considerable suppression of the natural colonisers by the sown species. The data for natural colonisation, presented here for illustration only, were not analysed. The average number of colonising species is (with the exception of the Czech site) lower in the high diversity plots than in the low diversity plots (Fig. 1), the difference is significant for the Netherlands ($P = 0.036$) and nearly significant for the UK site ($P = 0.067$). For the complete data set, using country as a random factor, the difference is also marginally significant ($P = 0.088$). Similarly, the worst performing HD replicate is able to suppress the natural colonisers better than the worst performing LD mixture ($P = 0.069$). On the contrary, there is no difference between the best performing LD and HD plots: the averages over countries were exactly the same ($P = 1.00$). Accordingly, when taking each measure (i.e. minimum, average, maximum) as a within-country factor, there is significant interaction ($P = 0.017$) between measure and treatment, signifying that these three measures respond to the diversity treatments in different ways.

The predicted values of 'competitiveness' for the LD mixtures were able to explain 61% of within country variability (measured by the proportion of sum of squares) in the richness of colonising species in the LD plots (Fig. 2). This shows that the species' competitive ability, estimated on the basis of their performance in HD, is a good predictor of their ability to suppress the weed species. Similarly, there are consistent differences between the treatments in minimum biomass (with HD having higher biomass than LD; $P = 0.042$). Differences in the average biomass were less pronounced ($P = 0.380$, but significant in NL, $P = 0.026$) and nearly no differences were found between LD and HD in maximum biomass (Fig. 3).

The situation can be illustrated by a PCA analysis of the species composition of sown species (Fig. 4, NL and UK analyses are presented as examples). As expected, the HD blocks are very similar to each other, whereas

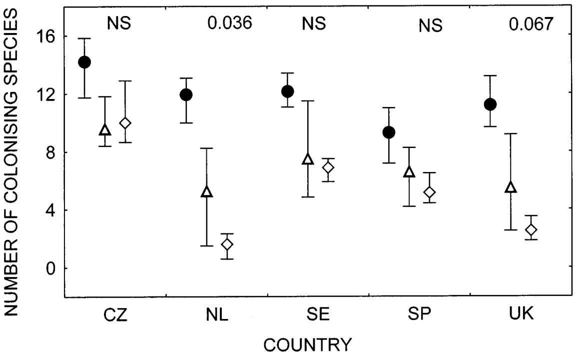

Fig. 1. Maximum, average and minimum number of colonising species in NC (●), LD (△) and HD (◇) treatments. The minimum and maximum (whiskers) are the extreme values among the five blocks, the value for each block is the average for twelve 1-m^2 quadrats. The values shown are the significance values for the comparison of averages derived from a separate ANOVA for each country (with block as a random factor), comparing the LD and HD treatments only, NS means $P > 0.1$.

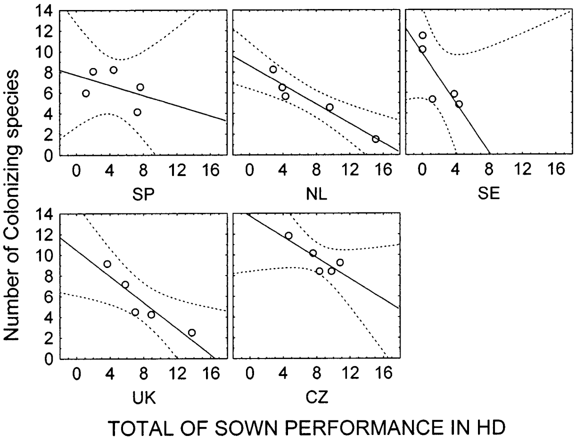

Fig. 2. Relationships between the predicted performance of the LD mixtures, estimated using the performance of their constituent species in the HD mixture, and the number of colonising species in that block. The 95% confidence intervals for particular countries are shown.

there are substantial differences between the LD plots in different blocks. More importantly, the greater the similarity in the species composition of LD to the HD plots, the higher the total community biomass and also the greater the suppression of weeds by that LD mixture (Fig. 4). Regarding both the number of weed species and total biomass, the LD plot with species composition most similar to HD plots fits well into the range of the HD plot values. Clearly, the sown species which are able to suppress other sown species attain high biomass and are able to suppress natural colonisers. Although the diversity profiles of sown species in HD plots differ among countries (Fig. 5), in all of them the HD plots are dominated by a limited number of sown species. In all the countries, four species or fewer attained average cover degree 3 or more (corresponding to 4–10% cover). In the UK, the site with the most equal distribution of sown species performances, only five of the sown species attained average cover scores higher than 2. Although species with cover values of less than 4% affect also the functioning of the community, the basic functions are probably dependent on species with greater cover. Consequently, the functioning of HD seems to be dependent on a very limited number of common species.

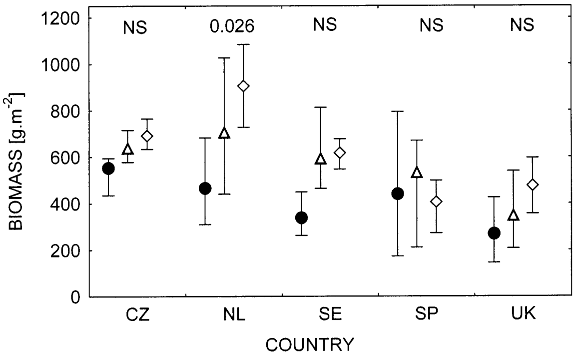

Fig. 3. Minimum, average and maximum biomass in NC (●), LD (△) and HD (◇) treatments. Minimum and maximum (whiskers) are extreme values among the five blocks, value for each block being average based on twelve 0.0625-m^2 plots. The values shown are the significance values for the comparison of averages derived from a separate ANOVA for each country (with block as a random factor), comparing the LD and HD treatments only, NS means $P > 0.1$.

In general, species that were successful in HD plots also attained high cover in LD plots (Fig. 6 with UK as example). The cover of each sown species was usually higher in the LD plots, which is not surprising because the individual species were sown there in higher densities (in order to maintain the same sowing densities of plant functional groups). Moreover, in the LD plots, levels of interspecific competition were probably lower. The performance in LD was also influenced by the composition of the sown mixture (which other species were present in a particular mixture). For example, the extremely high value attained by *Galium verum* in LD plots (Fig. 6) results from this species being sown in LD mixture with weak competitors (*Cynosurus cristatus*, *Holcus lanatus*, *Trifolium dubium*). For dominants that attained high cover in HD plots the five-fold increase in sowing densities did not cause a five-fold increase in cover (with higher sowing density, the effect of intraspecific competition increased and cover cannot exceed 100%). However, in Sweden, where a limited number of species attained high cover, the increased densities of dominants in LD plots in comparison to HD plots is probably the reason why the performance of the best LD mixture exceeded that of the best HD replicate.

There is no consistency between the countries in which of the sown species became important in the HD plots, nor in which functional group was the important one. HD plots in two countries were legume-dominated (SE: *Trifolium pratense*, *Lotus corniculatus*, NL: *Lotus corniculatus*, *Vicia cracca*, but with high representation of *Festuca rubra*). In the Czech Republic, the grass *Trisetum flavescens* attained the highest cover (often exceeding 50%), with strong representation of the legumes *Lotus corniculatus* and *Lathyrus pratensis*. In the UK and Spain, the forb *Plantago lanceolata* was the most successful plant, followed by grasses.

The Czech and UK sites used the same grass species; however, the relative importance differs considerably between countries: the correlation of average cover values in HD plots is only 0.38. Some generalities, however, can be found. The legume *Lotus corniculatus* and forb *Plantago lanceolata* were used in all the sites, and in all of the sites they are the best (exceptionally, the second best) in the corresponding functional group. In contrast, *Trifolium dubium*, the annual legume, managed to survive but with very low cover in all the sites

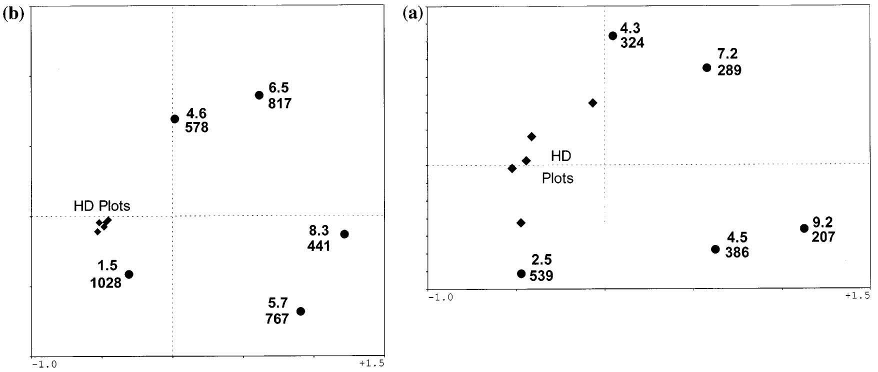

Fig. 4. PCA of species composition of sown species in (a) UK and (b) NL. First, a PCA of data from all the LD and HD permanent plots was calculated for the sown species. Centroids are shown for the twelve quadrats in each plot. The HD plots are very similar to each other and are displayed by diamonds. The LD plots are shown with circles and labelled: the upper figure shows the average number of weed species per m^2, and the lower figure the average biomass [g dry mass m^{-2}] in a given block. The average number of weed species in HD is 1.6 (range 0.6–2.3) in NL and 2.5 (range 1.8–3.5) in UK, average biomass is 905 (728–1084) g m^{-2} in NL and 476 (range 355–595) g m^{-2} in UK.

it was sown. Similarly, *Cynosurus cristatus* represented a minor part of the community in all the plots where it was sown (CZ, SE, UK).

The differences between countries in weed species suppression can be partly explained by the structure of the non-suppressed communities of natural colonisers in NC plots. In countries where a strong dominant was present in NC plots (i.e. countries characterised by high value of D_1: SE and CZ), the relative decrease in number of colonising species was lower than in countries where the NC community was without an apparent dominant weed species (NL, UK, Fig. 7).

Discussion

Our field experiments have shown that, on average, the performances (productivity, ability to suppress the weed species) of high diversity mixtures were usually better than those of the low diversity mixtures. However the performance of the best low diversity mixture was usually as good as the best of the high diversity replicates. On the contrary, the worst performing low diversity mixture always produced less biomass and had poorer weed suppression than the worst performing high diversity plot. The results of the third year of the experiment confirm results of the first two years described by Van der Putten et al. (2000) and suggest that this diversity effect was the result of the 'chance effect', i.e. with a higher number of species, the chance that influential species are present increases. The results correspond to the outcome of the 'competition for one resource' model of Tilman et al. (1997a, b), particularly to their Fig. 1, which is based completely on the 'chance effect' (see also Fig. 3a in Huston 1997). Similar results have also been obtained in other experiments (e.g. Naeem et al. 1995).

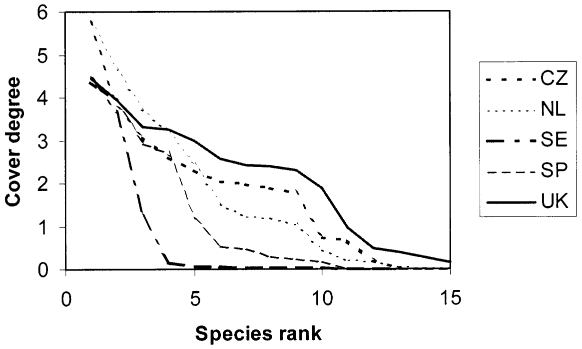

Fig. 5. Diversity profiles of the sown species in HD plots in the five sites. For each country, the species' average cover degree is plotted against its rank within the sown species.

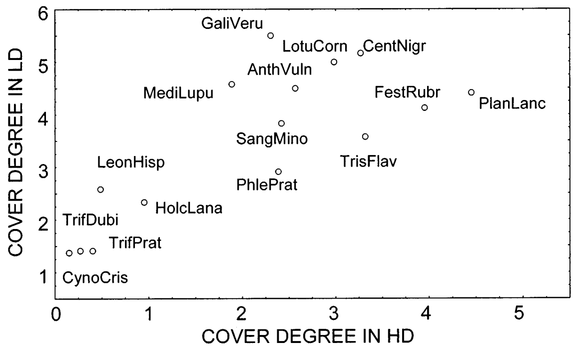

Fig. 6. Relationship between performance of the species in LD plots and HD plots. UK is used as an example.

If there was some complementarity, then it was impossible to distinguish it from the chance effect (see Hector 1998). However, our ability to predict the performance of particular low diversity mixtures on the basis of the performance of the constituent species in the high diversity plots also supports the importance of the chance effect. In interpreting our comparisons of HD and LD treatments, it should, however, be noted that only the species diversity within functional groups differed between treatments, the composition of functional groups was the same in HD and LD plots. It is expected that the differences would be more pronounced if the number of functional groups was varied (e.g. Hooper and Vitousek 1997). In particular, higher complementarity may be expected between species belonging to different functional groups than between species of the same group.

Our data suggest that the identity of the strongest competitor amongst the sown species depends considerably on local conditions. Moreover, the establishment success of particular species in early phases of succession could affect considerably the further development of a plant community (Lepš 1999, unpubl., Van der Putten et al. 2000) Very probably, the dominance of legumes in some sites is probably caused by their early establishment success (all of the dominant legumes had large seed mass), rather than to their nitrogen fixing abilities (neither the SE, nor the NL site are deficient in nitrogen; see Table 1). In contrast to the poor predictability amongst dominants, there are some species (*Trifolium dubium*, *Cynosurus cristatus*) that attained high cover at none of the sites. These species usually belong to the subordinate part (sensu Grime 1998) of the natural grassland communities in the area. Our results (Fig. 5) have shown that, although the species were sown in equal densities (within a functional group), they developed a diversity profile corresponding to the distinction between dominants and subordinate species of Grime (1998). In concordance with Grime, we expect that the effects measured (weed suppression, total biomass) were mostly influenced by the dominant species. The equitability of the final community is dependent on the equitability of establishment success, and is further reduced by competitive suppression of subordinate species. The UK site, exhibiting the highest equitability among all the sites (Fig. 5) was characterised by successful establishment of a vast majority of the species, and relatively low biomass (Fig. 3) resulting in low competitive suppression of subordinate species. The UK site is the only one located on rendzinas, which are generally less productive than brown soils. Accordingly, the surrounding (seminatural) communities, chalk grasslands, are characterised by low productivity and high species diversity.

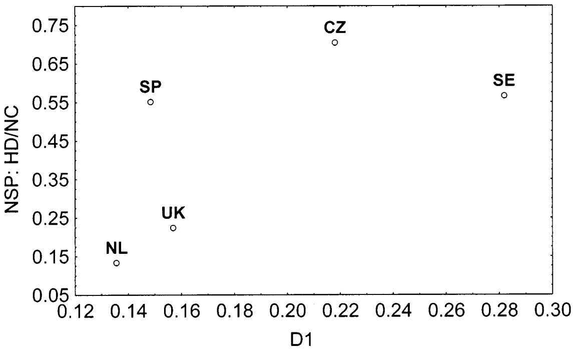

Fig. 7. Relationship between the suppression of colonising species, expressed as ratio of the average number of colonising species in HD to that in NC (NSP: HD/NC), and index of dominance D_1.

An increase in average yield (or any ecosystem function) for a mixture compared to the average of the monocultures could be caused by either the chance effect or the complementarity effect. It has been shown by Hector (1998) and Loreau (1998b) that only overyielding, i.e. mixtures of species producing higher yield than the best monoculture, shows unequivocally that some mechanism other than the chance effect plays a role (e.g. complementarity, facilitation, etc.). As the chance effect is much simpler and obvious, it is reasonable to consider it as a parsimonious explanation for cases where there is no overyielding. However, the complementarity effect cannot be excluded. Several analyses show that our results can be explained simply by the chance effect; on the contrary, if there was some complementarity effect, we were not able to reveal it. As we have no monoculture treatments, but only LD plots, we cannot rule out the occurrence of complementarity in the LD treatment and also (but not stronger) in the HD treatments. Similar results are not an exception. For example, in a system similar to ours (fallow field, four species in monocultures and their mixtures), Kroh and Stephenson (1980) concluded that mixture yields were less than the highest yielding monoculture but greater than the lowest monoculture yield. Similar results were obtained by, e.g. Haizel (1972).

There is also a statistical problem in comparing the best LD with the best HD (this problem is equivalent to the use of Max (M_i) (maximum yield in monoculture) by Loreau (1998b) in calculating his D_{Max} index). In our case, we compared the best LD with the best HD. For simplicity, we will ignore the block design of the experiment. We will use the classical ANOVA model notation. Taking biomass Y as an example (the same can be applied to any other measured response), consider n independent replications in each country, with the same species composition for HD, and n replications with different species composition for LD. In each country, the biomass in a particular replicate can be expressed as

$$Y = mean + diversity + specific + random$$

where Y is the biomass of a particular replicate, *mean* is the mean biomass over all the replicates of all (both) diversity levels, *diversity* is the effect of the diversity level, *specific* is the effect of particular species combination and *random* is the random variability. Then we test whether the values for *diversity* effect differ systematically from zero. When comparing the best replications, the values of *diversity* are estimated as the mean of differences between the best of the LD replications and the best of the HD replications in each country. In our case, we are not able to distinguish *specific* and *random* effects; nevertheless, we know that *specific* is the same for all the HD replications (so that it may be parameterised as zero), and it differs for particular LDs. We can reasonably expect that the variability in *specific* is much higher than the *random* variability. Consequently, by selecting the best HD we select a plot with the highest value of *random* whereas by selecting the best LD we select the plot with the highest *specific* regardless of the *random* variability. In other words, we compare the best of n replications of HD with a non-replicated value of the best LD species combination. Provided that the average response (biomass) for HD and for the best LD is the same, this procedure will give higher estimate for HD. Because we were not able to find any difference between the best HD and LD, even with this bias 'favouring' HD, it is very likely that there is no difference in our plots. This bias could be prevented by taking the average value of HD. However, in that case the results will be biased in favour of LD if *random* variability is comparable or higher to variability in *specific* (if variability in *specific* is small, the LD with high value of *random* is selected).

Unlike Huston (1997), we do not consider the chance effect to be a pure artefact. For example, the processes leading to selection of realized dominants from the pool of potential dominants after environmental change (see Grime 1998, point 5 of his summary) are probably similar to those engaged in the sampling effect. This phenomenon might be important in stability effects in natural situations. As a result of changes in environmental conditions (e.g. anthropogenic or natural variation in weather), species-poor communities are more likely to be prone to functional failure than species-rich communities because they do not contain species able to thrive in the new environmental conditions (the so-called 'portfolio effect', Tilman 1999). A similar example was reported by Rychnovská (1993: 145), in which the relative biomass of particular species in floodplain meadows changed considerably according to the floods in particular years, whereas the total biomass remained fairly constant. On the contrary, it is probable that this phenomenon manifests itself in nature mostly with respect to environmental fluctuations. Low diversity communities are usually formed of strong dominants, i.e. species able 'to do a job' *under the given conditions*. Communities dominated by a single keystone species are prone to functional failure when the conditions change and become unfavourable for the keystone species. For example, spruce forests in Central Europe, both natural and planted, collapsed due to the sensitivity of the spruce population to the air-pollution load (Kubíková 1991). This was accompanied by a loss in their hydrological function. On the contrary, outbreaks of Dutch elm disease, leading to the nearly complete loss of elm in vast areas, did not change ecosystem function to the same extent, because of functional redundancy in the group of canopy trees. In mixed forests, the gaps were quickly filled by competing species of similar function and the functioning of the ecosystem did not change much. Similarly, Chapin et al. (1998) suggested that high species diversity reduces the probability of large changes in ecosystem processes in response to invasions of pathogens and other species. However, similarly to Huston (1997), we believe that in experimental studies, the sampling effect must be distinguished from the other diversity effects because their consequences in real nature are probably very different.

Wardle (1999) suggested that 'if we accept sampling effect as a mechanism by which biodiversity-ecosystem function relationships could express themselves, then this would require us to assume that communities of species were randomly assembled with regard to their relative effect on the ecosystem function being investigated … and that whatever species were lost from an ecosystem were lost at random with respect to these effects'. Whereas we agree with Wardle (1999) that communities are usually not random assemblages of species and species are usually not lost at random, we do not think that complete randomness is necessary for the sampling effect to manifest itself. In our view, it is sufficient that the traits differ among the species in a community, and this assumption is a feasible one. However, the greater the dissimilarity between species, the higher chance that the sampling effect will manifest itself, and we agree with Wardle (1999) that the dissimilarity between species in experiments as compared to the dissimilarity in real communities has to be considered before claiming sampling effect being a legitimate diversity effect.

Many functional characteristics of ecosystems are in some way connected to productivity and, because of the strong effect of productivity on diversity (Al-Mufti et al. 1977), it is extremely difficult to interpret correlations between diversity and the functioning of ecosystems. For example, early- and late-successional communities differ not only in their stability characteristics, but also in their species richness. Nevertheless, both differences are dependent on the composition of prevailing life-history strategies in a community (Lepš et al. 1982, Lepš 1990). This might suggest that an experimental approach would be more profitable. However, the experimental approach also has its limitations. Probably the most important is the difficulty in separating the effects of species composition and species diversity (Aarssen 1997, Huston 1997, Tilman 1997). According to Tilman (1997), the separation is theoretically possible when '..., given a group of all potential species, called the "species pool", for effects to be attributed to species diversity, they must occur in comparison of the average response of two or more levels of diversity'. In our experiment, total number of seeds of each species sown in all the replications was the same in high and low diversity (Table 2). However, the average cover of the strong competitive species was higher in HD plots than their average over all the LD plots. For example, in the Czech plots, the strongest grass dominant, *Trisetum flavescens*, reached a cover of more than 60% in all HD plots (so the average was over 60%), which is more than the maximum attainable average in the LD plots, where it was present in two plots out of five only). Consequently, the average species composition in HD plots is biased toward the most competitive species (this effect is another manifestation of the sampling or chance effect). This makes it important to compare not only the average responses, but also the extremes. For the sampling effect to appear, it is important that the species differ in their efficiency. Consequently, the results of biodiversity experiments are dependent also on the species pool (i.e. on the set of species the researcher selects for the experiment).

Diversity might be important for an ecosystem's ability to buffer environmental fluctuations. These fluctuations, and particularly the occurrence of extreme conditions, are often eliminated or appear with low probability in short-term ecological experiments. Because of the necessity for a large number of replications, experiments can be executed on a limited spatial scale (and also over limited time). Peterson et al. (1998) have shown that the interaction of spatial and temporal scales can be important for the manifestation of diversity effects on ecosystem functioning (see also Huston 1999). Lawton et al. (1998) claim that 'massive replication of diversity treatments is not a problem'. In our view, it was serious practical problem, which depends on the scale of observation. As not only the development of vegetation, but also that of other trophic levels, especially including soil organisms, was the target of this study (see e.g. Korthals et al. unpubl.), the minimum plot size considered necessary to avoid the massive edge effect was 10 by 10 m; this clearly limited the possible number of replications.

The species composition of natural communities is a result of competitive forces acting over time. The species that are present are there because they were able to survive in the competitive struggle, potentially the result of niche complementarity. It is not clear whether diversity will have the same effect in randomly assembled mixtures of species, such as those used in manipulative experiments, as in a community in which species composition has developed and stabilised over a long period. Nevertheless, with all the constraints imposed by the experimental design, the species in the present study were selected so that they usually grow together in meadow communities in corresponding regions. This problem highlights the need for extreme caution in the interpretation of the results of simple ecological experiments. It is probable that simple effects caused by chance will be revealed, whereas the more complicated effects connected with interaction of scales and/or buffering of environmental fluctuations, or with other biota may remain hidden. Moreover, in this paper, we considered sampling effect and complementarity only. In more complicated natural ecosystems, there might be diversity effects connected with positive interactions among plants (e.g. facilitation, Holmgren et al. 1997).

Our ability to predict the performance of particular LD mixtures based on independent observations of their competitive ability (in our case their performance in the HD mixture) agrees well with Walker's (1992, 1995) 'drivers and passengers' model (see also Grime 1998, Peterson et al. 1998, Walker et al. 1999). Some species (the drivers) play an important role in ecosystem function, whilst the absence of others (the passengers) does not affect ecosystem function at all, at least in the short term and small spatial scale (Nijs and Roy 2000). The role of the latter can be 'to wait for their chance' to appear somewhere in space or time (Grime 1998, Walker et al. 1999). Furthermore, whilst these species may not be important for ecological functions such as carbon fluxes, they may be essential for others. For example, Lepš et al. (1998) have shown that plant species with a low average biomass (0.27 g dry weight m^{-2}) were able to support large populations of monophagous insects.

In our experiments, we did not find any effect of diversity beyond the chance effect; complementarity, if it was at play, did not have a strong enough effect to be detected. However, we do not consider the sampling effect a mere artefact. Similar mechanisms, i.e. selection of species from a pool of species available at a site, can probably cause higher adaptability of ecological communities to a changing environment. With respect to temporal and spatial scales important for the manifestation of diversity effects within ecological communities, it is very likely that many of these diversity effects might be missed by spatially and temporally limited ecological experiments.

Acknowledgements – The experimental work described was supported by the Framework IV Environment and Climate Programme of the European Commission, contract number ENV4-CT95-0002 and ERBIC20-CT96-0025. We are indebted to many people for their help with field-work: Liz Allchin, Helen Burman, Blanka Divišová, Tereza Francírková, Michal Hejcman, Anders Johansson, Frank Menting, Alvaro Peix, Luis Felipe Rodriguez, Petr Šmilauer, Majka Šmilauerová, Miroslav Šrůtek, David Stahlberg, Camilla Svegborn, and Dana Vašková. We thank Michael Huston for valuable comments on the manuscript.

References

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

Al-Mufti, M. M., Sydes, C. L., Furness, S. B. et al. 1977. A quantitative analysis of shoot phenology and dominance in herbaceous vegetation. – J. Ecol. 65: 759–791.

Chapin, F. S., Sala, O. E., Burke, I. C. et al. 1998. Ecosystem consequences of changing biodiversity. – BioScience 48: 45–52.

Darwin, C. 1859. The origin of species by means of natural selection. – John Murray.

de Wit, C. T. 1960. On competition. – Versl. landbouwk. Onderz. 66: 1–82.

Diamond, J. 1986. Overview: laboratory experiments, field experiments, and natural experiments. – In: Diamond, J. and Case, T. J. (eds), Community ecology. Harper & Row, 3 p.

Doak, D. F., Bigger, D., Hardin, E. K. et al. 1998. The statistical inevitability of stability-diversity relationship in community ecology. – Am. Nat. 151: 264–276.

Elton, C. 1958. The ecology of invasions by animals and plants. – Methuen.

Grime, J. P. 1979. Plant strategies and vegetation processes. – Wiley.

Grime, J. P. 1997. Biodiversity and ecosystem function: the debate deepens. – Science 277: 1260–1261.

Grime, J. P. 1998. Benefits of plant diversity to ecosystems: immediate, filter and founder effects. – J. Ecol. 86: 902–910.

Haizel, K. A. 1972. The productivity of mixtures of two and three species. – J. Appl. Ecol. 9: 601–608.

Hector, A. 1998. The effect of diversity on productivity: detecting the role of species complementarity. – Oikos 82: 597–599.

Hector, A., Schmid, B., Beierkuhnlein, C. et al. 1999. Plant diversity and productivity in European grasslands. – Science 286: 1123–1127.

Hodgson, J. G., Thompson, K., Bogaard, A. and Wilson, P. J. 1998. Does biodiversity determine ecosystem function? The Ecotron experiment reconsidered. – Funct. Ecol. 12: 843–848.

Holmgren, M., Scheffer, M. and Huston, M. A. 1997. The interplay of facilitation and competition in plant communities. – Ecology 78: 1966–1975.

Hooper, D. U. and Vitousek, P. M. 1997. The effects of plant composition and diversity on ecosystem processes. – Science 277: 1302–1304.

Hooper, D. U. and Vitousek, P. M. 1998. Effects of plant composition and diversity on nutrient cycling. – Ecol. Monogr. 68: 121–149.

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

Huston, M. A. 1999. Local processes and regional patterns: appropriate scales for understanding variation in the diversity of plants and animals. – Oikos 86: 393–401.

Johnson, K. H., Vogt, K. A., Clark, H. J. et al. 1996. Biodiversity and the productivity and stability of ecosystems. – Trends Ecol. Evol. 11: 372–377.

Kroh, G. C. and Stephenson, S. N. 1980. Effect of diversity and pattern on relative yields of four Michigan first year fallow field plant species. – Oecologia 45: 366–371.

Kubíková, J. 1991. Forest dieback in Czechoslovakia. – Vegetatio 93: 101–108.

Lawton, J. H., Naeem, S., Thompson, L. J. et al. 1998. Biodiversity and ecosystem function: getting the Ecotron experiment in its correct context. – Funct. Ecol. 12: 848–852.

Lepš, J. 1990. Stability. – In: Osbornová, J., Kovářová, M., Lepš, J. and Prach, K. (eds), Succession in abandoned fields. Studies in Central Bohemia, Czechoslovakia. Kluwer, pp. 143–147.

Lepš, J. 1999. Chance effects in early stages of succession cause divergence. Abstracts. – British Ecological Society Winter and Annual General Meeting, University of Leeds, 20–22 December 1999: 15.

Lepš, J., Osbornová-Kosinová, J. and Rejmánek, M. 1982. Community stability, complexity and species life history strategies. – Vegetatio 50: 53–63.

Lepš, J., Spitzer, K. and Jaroš, J. 1998. Food plants species composition and variability of the moth community in undisturbed forest. – Oikos 81: 538–548.

Lepš, J., Goldberg, D. E., Herben, T. and Palmer, M. 1999. Mechanistic explanations of community structure: Introduction. – J. Veg. Sci. 10: 147–150.

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

Loreau, M. 1998b. Separating sampling and other effects in biodiversity experiments. – Oikos 82: 600–602.

MacArthur, R. 1955. Fluctuations of animal populations and a measure of community stability. – Ecology 36: 533–536.

MacGillivray, G. W., Grime, J. P. and the integrated screening programme (ISP) team. 1995. Testing predictions of the resistance and resilience of vegetation subjected to extreme events. – Funct. Ecol. 9: 640-649.

May, R. 1973. Stability and complexity in model ecosystems. – Princeton Univ. Press.

McNaughton, S. J. 1967. Relationship among functional properties of Californian grassland. – Nature 216: 168–169.

Mooney, H. A. 1997. Ecosystem function of biodiversity: the basis of the viewpoint. – In: Smith, T. M., Shugart, H. H. and Woodward, F. I. (eds), Plant functional types: their relevance to ecosystem properties and global changes. Cambridge Univ. Press, pp. 341–354.

Naeem, S. 2000. Reply to Wardle et al. – ESA Bull. 81: 241–246.

Naeem, S., Thompson, L. J., Lawler, S. P. et al. 1995. Empirical evidence that declining species diversity may alter the performance of terrestrial ecosystems. – Proc. R. Soc. Lond. B 347: 249–262.

Naeem, S., Håkansson, K., Lawton, J. H. et al. 1996. Biodiversity and plant productivity in a model assemblage of plant species. – Oikos 76: 259–264.

Naeem, S., Chapin III, F. S., Costanza, R. et al. 1999. Biodiversity and ecosystem functioning: maintaining natural life support processes. – Issues in Ecology 4: 1–14.

Nijs, I. and Roy, J. 2001. How important are species richness, species evenness and interspecific differences to productivity? A mathematical model. – Oikos 87: 57–66.

Odum, E. P. 1971. Fundamentals of ecology. – Saunders.

Peterson, G., Allen, C. R. and Holling, C. S. 1998. Ecological resilience, biodiversity, and scale. – Ecosystems 1: 6–18.

Pimm, S. L. 1984. The complexity and stability of ecosystems. – Nature 307: 321–326.

Rejmánek, M. 1996. Species richness and resistance to invasions. – In: Orians, G. H., Dirzo, R. and Cushman, J. H. (eds), Biodiversity and ecosystem processes in tropical forests. Springer-Verlag, pp. 153–172.

Roy, J. 2001. How does biodiversity control primary productivity? – In: Roy, J., Saugier, B. and Mooney, H. A. (eds), Global terrestrial productivity. Academic Press, pp. 169–186.

Rychnovská, M. 1993. Temperate semi-natural grasslands of Eurasia. – In: Coupland, R. T. (ed.), Natural grasslands. Elsevier, pp. 125–166.

Schläpfer, F, Schmidt, B. and Seidl, I. 1999. Expert estimates about effects of biodiversity on ecosystem processes and services. – Oikos 84: 346–352.

Singh, J. S. and Misra, R. 1969. Diversity, dominance, stability, and net production in the grasslans at Varanasi, India. – Can. J. Bot. 47: 425–427.

ter Braak, C. J. F. and Šmilauer, P. 1998. CANOCO Release 4. Reference manual and user's guide to Canoco for Windows: Software for Canonical Community Ordination. – Microcomputer Power, Ithaca, NY.

Tilman, D. 1997. Distinguishing between the effects of species diversity and species composition. – Oikos 80: 185.

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

Tilman, D., Knops, J., Wedin, D. et al. 1997a. The influence of functional diversity and composition on ecosystem processes. – Science 277: 1300–1302.

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

Tilman, D., Lehman, C. L. and Bristow, C. E. 1998. Diversity-stability relationship: statistical inevitability or ecological consequence. – Am. Nat. 151: 277–282.

Trenbath, B. R. 1974. Biomass productivity of mixtures. – Adv. Agron. 26: 177–210.

van den Bergh, J. P. and Elberse, W. T. 1970. Yields of monocultures and mixtures of two grass species differing in growth habit. – J. Appl. Ecol. 7: 311–320.

van der Putten, W. H., Mortimer, S. R., Hedlund, K. et al. 2000. Plant species diversity as a driver of early succession in abandoned fields: a multi-site approach. – Oecologia 124: 91–99.

Vandermeer, J. 1989. Ecology of intercropping. – Cambridge Univ. Press.

Van Voris, P., O'Neil, R., Emanuel, W. R. and Shugart, H. H. 1980. Functional complexity and ecosystem stability. – Ecology 61: 1352–1360.

Walker, B. 1992. Biological diversity and ecological redundancy. – Conserv. Biol. 6: 18–23.

Walker, B. 1995. Conserving biological diversity through ecosystem resilience. – Conserv. Biol. 9: 747–752.

Walker, B., Kinzig, A. and Langridge, J. 1999. Plant attribute diversity, resilience, and ecosystem function: the nature and significance of dominant and minor species. – Ecosystems 2: 95–113.

Wardle, D. A. 1999. Is 'sampling effect' a problem for experiments investigating biodiversity-ecosystem function relationships? – Oikos 87: 403–407.

Wardle, D. A., Huston, M. A., Grime, J. P. et al. 2000. Biodiversity and ecosystem functioning: an issue in ecology. – ESA Bull. 81: 235–239.