Vol. 311: 295–309, 2006 | MARINE ECOLOGY PROGRESS SERIES
Mar Ecol Prog Ser | Published April 13

OPEN ACCESS

Using model systems to address the biodiversity–ecosystem functioning process

M. T. Bulling[1,*], P. C. L. White[1], D. Raffaelli[1], G. J. Pierce[2]

[1]Environment Department, University of York, Heslington, York YO10 5DD, UK
[2]School of Biological Sciences, Zoology Building, University of Aberdeen, Aberdeen AB24 2TZ, UK

ABSTRACT: Declines in biodiversity resulting from anthropogenic disturbance to ecosystems have focused attention on the role of biodiversity in ecosystem functioning. However, the high level of complexity of ecosystems has made this a difficult topic to investigate. Much simpler model systems incorporating small-scale, spatially delimited, artificial assemblages of species have been widely used recently to address the link between biodiversity and ecosystem functioning (BEF). Their simplicity lends tractability to these systems, but has also resulted in much criticism in the literature over their relevance. Here, we examine the strengths and limitations of model systems and examine how useful these systems might be in addressing several issues that are likely to represent future challenges to understanding BEF: spatial scale, multiple trophic levels, variation, environmental stochasticity and the choice of representative combinations of species. We find that model systems have already played an important role in enhancing our understanding of BEF and are likely to continue this role in the future. However, they do have important limitations, and it is essential to take these into account when putting results into the broader context of ecosystems and to improve the level of integration of results with those from other methodologies.

KEY WORDS: Biodiversity · Ecosystem functioning · Model systems · Hidden treatments · Spatial scale · Multiple trophic levels · Mesocosms

INTRODUCTION

There is a long history of the use of model systems to address questions in marine ecology, probably reflecting the more difficult working environment of the sea and seabed compared to more accessible terrestrial systems. Benthic and pelagic micro- and mesocosms have been employed extensively in research, ranging from pollution effects to benthic–pelagic coupling (Lalli 1991), such that they are now a routine tool for marine investigations. In addition, marine ecologists have played a major role in developing appropriate and rigorous experimental designs and subsequent analyses (Gamble 1991, Underwood 1998), and the techniques required to maintain and monitor environmental processes under different experimental treatments are well established.

Few marine ecologists would therefore need convincing of the utility of model systems for exploring relationships between biological and physico-chemical processes. It is somewhat surprising therefore that relatively few attempts (compared to freshwater and terrestrial ecology) have been made to take advantage of such systems for exploring one of the key emerging issues in ecology, the effects of biodiversity on ecosystem functioning. There is a strong tradition in ecology of investigating the abiotic drivers of ecosystem structure and dynamics, whilst biotic factors have been highlighted only comparatively recently (Lawton 1994, Chapin et al. 1997, Loreau 2000). This change in emphasis has been given impetus due to current rates of species loss and concerns as to how biodiversity change will affect the efficiency with which ecosystems will function, especially with regard to the provision of ecological goods and services (Ehrlich & Wilson 1991, Pimm et al. 1995, Costanza et al. 1997, Vitousek et al. 1997, Chapin et al. 2000).

*Email: mtb7@york.ac.uk

Investigating the relationships between biodiversity and ecosystem functioning is extraordinarily challenging given the complexity of real ecosystems. Model systems are a subset of experimental models that have been at the centre of biodiversity–ecosystem functioning (BEF) research (Lawton et al. 1998, Emmerson et al. 2001, Hector et al. 2002, Petchey et al. 2002b, Tilman et al. 2002). These systems use spatially delimited replicates of precisely controlled starting conditions (including identity, number and density of species), and have been at the heart of much of the controversy surrounding BEF research (see 'Combinations of species' below). BEF is a shift from much of the earlier biodiversity research in marine ecosystems, but a field where marine ecologists can play a key role by virtue of their experience and expertise in the use of model systems. Here, we review the strengths and limitations of the model system approach as revealed by studies of BEF issues, mainly conducted in terrestrial and freshwater systems. We then examine how model systems might be used to address future challenges, and how they may be better integrated with other studies in order to obtain a more rigorous understanding of BEF relationships.

MODEL SYSTEMS

The majority of model systems have been based on microcosms or mesocosms, defined as 'spatially delimited artificially constructed model ecosystems' (Petchey et al. 2002b). In the present paper, we use this as the basis of our definition of model systems as 'small-scale, spatially delimited, artificially constructed ecosystems, allowing fine control over initial composition'. The difference between micro- and mesocosms is rather arbitrary (Lawler 1998), with mesocosms tending to be larger in absolute size and, therefore, of greater biological and spatial complexity (Petchey et al. 2002b). In the present paper we do not distinguish between micro- and mesocosms, but differences in scale in the 2 types of systems will affect the strengths of some of the criticisms of model systems in general. Our definition of model systems also embraces spatially larger experimental projects such as Biodepth (Hector et al. 2002) and the biodiversity experiments in grasslands (Tilman et al. 2002). The use of model systems for addressing BEF questions has typically involved experimental designs with biodiversity as the independent (explanatory) variable and the levels of a particular ecosystem function (often productivity) as the dependent (response) variable. The biodiversity metric most often used is species richness (i.e. the number of species).

The main advantage of model systems is the high level of control possible over initial conditions, including the choice of species, numbers of individuals and starting physico-chemical conditions. This and the relatively small size of the individual treatment plots means that configurations can easily be replicated (Lawton 1995, Drake et al. 1996), increasing the level of statistical power in the experimental design and the persuasiveness of findings. Related to this is reproducibility (Lawton 1995, Drake et al. 1996), whereby starting conditions can be set up precisely and repeatedly with minimal variation. This high level of control over initial configurations means that specific factors, such as biodiversity, can be isolated and their role in ecosystem functioning can be assessed unambiguously. Small-scale systems are also easily monitored and maintained. Model systems therefore potentially offer a high level of tractability.

However, these high levels of control and replication mean that model system experiments are very specific to the system, organisms and experimental configuration being used. In turn, this specificity means that results are unlikely to be generally applicable across systems, or indeed between different configurations of the same system. In the same way, model systems are designed to focus on specific independent factors, allowing little understanding of the relative importance of these factors in relation to others in the full and more complex system.

Model systems at the lower end of the spatial scale (meso- and microcosms) involve smaller organisms, with short generation times, making them well suited to addressing questions requiring many generations (Drake et al. 1996, Petchey et al. 2002b). This is an important consideration, particularly when trying to relate results from micro- and mesocosm experiments to those from larger-scale terrestrial plant experiments (Petchey et al. 2002b). If the experimental duration encompasses many generations, the results will be strongly driven by population dynamics. However, terrestrial plant experiments tend to last for only a few generations, and can therefore be more affected by transient effects of initial species composition (Petchey et al. 2002b). Comparison of results between model systems at different spatial scales may thus be inappropriate and misleading.

However, whilst smaller model systems might appear to have advantages over field experiments, these same advantages are seen by some as potential problems. The short time required to run the experiments, the modest cost and the ability to conduct the experiment in the laboratory, rather than having to regularly visit a field site, have been claimed as enticements leading scientists to use these experimental designs at the expense of experimental rigour (Carpenter 1996). A more subtle concern is that of mean field approximations (Petersen & Hastings 2001). In the

tight experimental design of many small-scale systems (cf. more open systems, e.g. Hector et al. 2002, Tilman et al. 2002) levels of environmental variables are approximated and constant, based on average levels of environmental factors (e.g. light, temperature, humidity). There is thus an implicit assumption that average levels of environmental factors, rather than their extreme levels or variability, determine ecosystem dynamics (Petersen & Hastings 2001). This may not be true and could lead to false conclusions being drawn from such experiments. However, mean field approximation can be an advantage in the integration of model systems with mathematical models and theory. Mathematical ecosystem models generally use mean field approximations for environmental variables, and there is therefore a close conceptual link between these types of approaches. This is particularly so for aquatic model systems (Petchey et al. 2002b). Mathematical models and theory often implicitly assume a closed system and a lack of spatial structure, the trademark of many aquatic model systems and quite different from most field experiments done in terrestrial systems. Reduced spatial environmental variation coupled with generally mobile individuals, leads to limited spatial aggregation of individuals and therefore dynamics approximating mean field conditions (Tilman & Kareiva 1997).

By definition, model systems are taxonomically and structurally simplistic (Lawton 1995, 1996, Carpenter 1996, Drake et al. 1996). For example, there is little or no spatial or temporal heterogeneity (Lawton 1995), seasonal variation, disturbance, or large environmental perturbation (Lawton 1996). Whilst such simplicity provides for an elegant and tractable experimental design, important characteristics of real communities and ecosystems will be excluded or distorted (Carpenter 1996). Carpenter (1996) cites examples from limnology (Gerhart & Likens 1975, Stephenson et al. 1984, Bloesch et al. 1988) in which container size and experimental duration have affected experimental results.

There are also criticisms of the identities of the species that have been used in these experiments which may contain unnatural assemblages (Lawton 1996) and, hence, be a poor reflection of the real ecosystem. Also, species may be used because they survive well in microcosms, rather than being representative of the wider community (Lawton 1995) in which the species have long interconnected evolutionary histories (Lawton 1996). If species composition in individual-richness treatments is not realistic, between-treatment differences may not represent the differences likely under natural extinction/colonisation scenarios (Solan et al. 2004), leading to difficulties in interpretation.

Given these limitations and abstractions, it is not surprising that model systems have attracted hostility. For example, Carpenter (1996) has argued that the limitations of this experimental approach mean that most important questions in applied ecology cannot be investigated by mesocosms, with statistical advantages not offsetting the problems of limited scale, so that there is considerable risk of obtaining misleading results from these experiments. Despite the strong views expressed by proponents and critics, there are areas of general agreement. All realise that model systems are great simplifications of ecosystems, which are highly complex, making it foolish and misleading to extrapolate results from these experiments directly to ecosystems. Another area of agreement is that these model systems form only one part of a suite of tools for studying ecosystems, including mathematical and conceptual models, as well as field observations, experiments and manipulations. However, views about the contribution that model systems make do vary, from 'supportive' (Carpenter 1996) to 'a unifying thread for the development of a conceptual framework for understanding higher levels of organization' (Drake et al. 1996).

Clearly, model systems have an important role to play in BEF issues. Observations and correlations on their own do not allow us to distinguish between cause and effect (Manly 1992), whilst field studies do not allow the level of control and measurement permitted in model systems, required to develop theories and required to test hypotheses. Model systems therefore provide us with an extremely valuable starting point in our understanding of ecosystems. They should be viewed as one component in an integrated suite of methodologies that we can use to develop theories and test hypotheses of ecosystem processes (Lawton 1995, 1996, Petchey et al. 2002b). They are extremely useful, but must be used with care and in context.

FUTURE CHALLENGES FOR MODEL SYSTEMS

The sheer complexity of ecosystems has meant that most research into understanding BEF has been of the reductionist philosophy. This, combined with initial limitations imposed by the experimental design of model systems, has meant that many of the model systems used to date appear simplistic. Many of the future challenges for model systems stem from the requirement to bridge the gap between these simple systems and real ecosystems by incorporating further key aspects of ecosystems in their design. Incorporating additional ecological complexity in these systems without losing their tractability requires careful thought about which aspects should be included and about experimental design. Below, we review a number of ecosystem features that are likely to be important dri-

vers in a BEF relationship, most of which have so far not been addressed by model systems. We consider why these features are likely to be important, the problems of incorporating them into model system designs, and how these problems may be overcome.

Combinations of species

The majority of model systems have used terrestrial plants (Schlapfer & Schmid 1999, Hector et al. 2002, Tilman et al. 2002), and, as noted below, this is a trophic bias that needs to be corrected. However, the particular combinations of species used in model systems are also open to criticism. The independent variable that is usually adopted in model system experiments is species richness, with different levels of this factor assembled in a controlled fashion, usually species being chosen at random from a species pool. This has generated a debate over whether effects on ecosystem functioning are due to biodiversity through resource partitioning or positive interactions, or are due to 'hidden treatments'. The hidden treatment that has received the most attention has been 'sampling effects' (Huston et al. 2000, Fridley 2001, Loreau & Hector 2001, Huston & McBride 2002, Tilman et al. 2002). A sampling effect occurs when randomly chosen (without replacement) combinations with more species have a greater probability of containing a species or groups of species with strong functional characteristics than those combinations with fewer species, and therefore are likely to demonstrate a greater level of ecosystem functioning through this chance selection. Consequently, it is difficult to determine if an increase in ecosystem functioning is genuinely due to an increase in the number of species or due to this sampling effect.

One approach that allows the distinction of complementarity and species-selection effects on ecosystem functioning has been described (Loreau & Hector 2001, Sala 2001). This requires replications of single-species treatments so that levels of ecosystem functioning in multi-species configurations can be separated according to sampling effects and species complementarity through additive partitioning. Although this methodology is a significant step forward in interpreting results, it is an *a posteriori* approach with limitations and cannot replace experiments that explicitly address actual mechanisms operating between biodiversity and ecosystem functioning (Loreau & Hector 2001). An important limitation of this methodology is that it allows separation of the 'complementarity effect' and the 'selection effect', but that an underlying mechanism resulting from complementarity can influence both effects (Petchey 2003). Thus, it is not possible to link a single mechanism or class of mechanisms to the 'complementarity effect', an error that has been made frequently (Petchey 2003). This technique of separating effects has been used as part of an explicit ANOVA (analysis of variance) design in conjunction with separate ANOVAs, with each species as a dummy variable to assess the level of species-selection effects with changing soil fertility in a terrestrial plant experiment (Fridley 2002).

The methodology of Loreau & Hector (2001) for separating complementarity and species-selection effects relies on a replacement series experimental design that assumes no density-dependent effects (Benedetti-Cecchi 2004). These designs require compensatory reductions in abundance of species with increasing diversity to keep total density constant. This allows confounding of complementarity and species-selection effects with density-dependent effects. Many of these issues of separating the different effects and their interactions have been examined through the use of Monte Carlo simulations (Benedetti-Cecchi 2004). This work investigated a new design, with biodiversity (species richness) and density of species as fixed, crossed treatments. Species assemblage was treated as a random factor, which was nested within richness treatments and crossed with density of species treatments. This allows the identification of species-richness effects even with the added complication of density-dependent effects, whilst controlling for species-identity effects, which can also be identified (Benedetti-Cecchi 2004).

Views of the relevance of these sampling effects fall into 2 camps. Some authors consider such effects as experimental artefacts that interfere with the correct interpretation of biodiversity effects on ecosystem functioning (Huston 1999, Wardle 1999, Huston et al. 2000, Huston & McBride 2002). Others see them as a natural consequence of the inherent differences between species and therefore as a mechanism through which species richness may affect ecosystem functioning (Lawton et al. 1998, Chapin et al. 2000, Purvis & Hector 2000). However, the validity of the sampling effect must rest with how well random species addition/deletion in the experiment matches the processes of species addition and deletion in real ecosystems (Fridley 2001), and this is likely to vary considerably. It is probable that the environment will strongly affect the level of this matching, with resulting immigration and emigration processes that are highly stochastic likely to lead to strong matching (Fridley 2001). The significance of sampling effects in model systems must therefore be considered on a case-by-case basis in relation to the ecosystem being modelled, rather than assigning it the constant value of a true effect of biodiversity or a statistical artefact. The sampling effect is a mechanism through which levels of ecosystem func-

tioning can be affected, and therefore the differences are largely semantic (Cottingham et al. 2001).

Two other potential hidden treatments are the 'quasi-replication effect' and the 'variance-reduction effect' (Huston & McBride 2002). These are subtler than the sampling effect. The term 'quasi-replication' (Huston & McBride 2002) refers to the fact that drawing species randomly from a species pool for each replicate at a particular species richness means that replicates will not be true replicates, as they will not contain exactly the same species.

With a given pool of species to randomly draw from (without replacement), treatments with intermediate numbers of species have a much larger number of possible combinations of species than treatments with low or high numbers of species. This means that with the same number of replicates for each treatment the species combinations that are used for intermediate biodiversities will represent only a small proportion of the total number of potential combinations. Replicates of lowest and highest biodiversity treatments will represent much greater proportions of the potential species combinations at these species richnesses. This underrepresentation of species combinations at the intermediate diversity levels means that with the usual small number of replications in each richness treatment, it is unlikely that resulting statistical properties are representative of the population of species combinations (Huston & McBride 2002). This is the quasi-replication effect (Huston & McBride 2002). This implies that the level of variance will change with richness treatment, introducing heterogeneity of variance and therefore breaking one of the assumptions made by ANOVA and linear regression models, both used extensively to analyse BEF model systems.

The variance-reduction effect (Huston 1997, Huston & McBride 2002) results from the similarity of replicates within a species-richness treatment changing with species richness (Fig. 1). The similarity of quasi-replicates within a richness treatment level increases at an increasing rate with species richness, assuming random selection without replacement. This asymmetry in similarity results in high-richness treatments containing almost true replicates, and those treatments with low species numbers being very far from true replicates. As a result, statistical behaviour of replicates is likely to change with species richness. The range of 'real' (i.e. statistical population parameter) levels of ecosystem functioning from the quasi-replicates in high species treatments is likely to be smaller than the range at low species richness due to the different levels in similarity of replicates within a treatment. A consequence of this is that, in high-richness treatments, levels of ecosystem functioning are more likely to be proportionately more strongly affected by experimental error than by biodiversity effects. Due to the parallel sampling effect making it more likely that these high-richness treatments will contain the more productive species, the effect of experimental error is likely to be in the direction of increased ecosystem functioning. This increase in ecosystem functioning due to experimental error cannot be separated from increases due to species richness and therefore may lead to a misrepresentation of the meaning of experimental results.

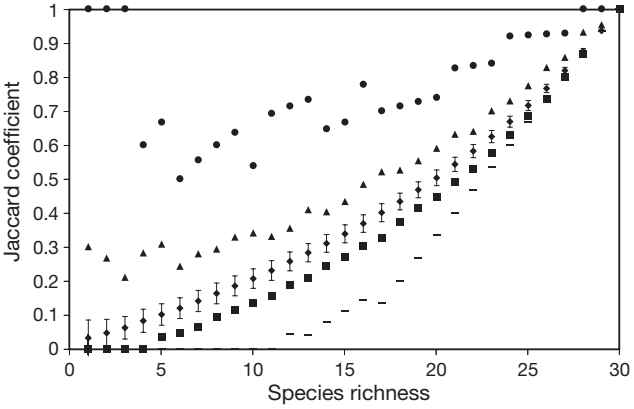

Fig. 1. Levels of similarity between 'quasi-replicates' (Huston & McBride 2002) with increasing species richness. Similarity between 2 replicates is measured with the Jaccard coefficient, which is the number of species in both replicates divided by the number of species in just one or the other replicate. A coefficient of 0 indicates no similarity, and a coefficient of 1 indicates the replicates contain the same species. Simulations were run using a pool of 30 species from which species were selected at random without replacement and placed in a replicate. An experiment with 5 replicates per species-richness treatment was simulated. The mean coefficient over all unique pairings of replicates was calculated for each set of 5 replicates to get an 'overall similarity'; 1000 simulations were run, and the mean level of overall similarity at each species richness is shown (◆) with the standard deviation. The minimum (■) and maximum (▲) levels of 'overall similarity' in the 1000 simulations are also shown. The minimum (-) and maximum (●) individual Jaccard coefficients that occurred between pairs of replicates are shown to indicate the likely range of similarity in different species-richness treatments

The variance-reduction effect means that there will be a breakdown of the assumption of homogeneity of variance across treatments, an assumption made by ANOVA, which has been used extensively to analyse the results from BEF experiments. The variance-reduction effect could be avoided through the use of true replicates in the experimental design. If this is done, then the experimental design advocated by Benedetti-Cecchi (2004) will allow a formal test of the variance-reduction effect.

Such concerns have led some to call for conclusions based on results from experiments involving random sampling from a species pool to be re-evaluated and

for future experimental designs to focus on the mechanistic drivers of BEF (Huston & McBride 2002). The level of error due to environmental variation can be accounted for in 2 main ways (Huston & McBride 2002). Firstly, using the usual experimental design (biodiversity as an independent factor and ecosystem functioning as a dependent variable), environmental parameters can be measured and incorporated into the statistical analysis as covariates. Secondly, a more controlled approach involves the deliberate manipulation of environmental conditions and for these levels to be included as factors in the experimental design. The levels of interactions between these environmental factors with diversity (as a fixed factor) in an ANOVA experimental design can also be obtained. This approach has been successfully used in a grassland model system experiment, which demonstrated that the effect of diversity on productivity significantly increased with soil fertility (the environmental variable) (Fridley 2002).

The need to address BEF via a mechanistic framework has been recognised, and several experimental designs have focussed on functional diversity rather than species richness (McGrady-Steed et al. 1997, Naeem & Li 1997, Mulder et al. 1999, Paine 2002). However, this change in focus has its own associated problems. The functional approach has not been particularly successful in achieving an understanding of how taxonomic or functional diversity can affect ecosystem functioning, because the functional groups used have been subjective or defined via statistical segregation based on ecophysiological data (Naeem & Wright 2003). A general framework for identifying appropriate functional groups has been proposed that splits response traits from effect traits, and selects the relevant functional traits based on the driver and ecosystem function of interest (Naeem & Wright 2003). A more appropriate measure of functional diversity than functional richness has been proposed (Petchey & Gaston 2002). FD is a functional diversity index that is a continuous and simultaneous estimate of the dispersion of species in trait space at all hierarchical scales. Another proposed measure is functional attribute diversity (FAD) (Walker et al. 1999), which uses an estimate of the total distance between species in trait space. These 2 measures explained a greater level of variation in above-ground biomass production than species richness and functional group richness in a comparison of functional diversity measures (Petchey et al. 2004). This greater explanatory power was partly explained by FD and FAD not arbitrarily assigning species to functional groups, enabling continuous measures of biodiversity and the greater amount of biological information that the 2 measures contain (Petchey et al. 2004). An alternative approach is to use a test of statistical significance that accounts for the effects of grouping per se (Petchey 2004). This is done through a randomisation test that compares the observed statistic against a distribution determined using repeated random assignment of species to functional groups, allowing the testing of the null hypothesis that functional group richness has no effect on ecosystem functioning.

Whether experimental design focuses on species richness or functional richness, the choice of species or functional group treatments has still been primarily based on random selection (Naeem & Wright 2003). This is equivalent to random extinction, but, in the real world, extinction tends to be non-random (Pauly et al. 1998, Srivastava 2002, Smith & Knapp 2003). Therefore random selection of species or functional groups in experimental design is unlikely to produce combinations representative of the real world (Wardle 1999, Schwartz et al. 2000), although this is possible in systems with strong stochastic drivers (Fridley 2001). The possible significance of non-random extinctions on BEF has recently been demonstrated using data from marine invertebrate communities to parameterise models predicting the effects of extinctions on sediment bioturbation (Solan et al. 2004). The models demonstrated that the magnitude of the effect of species loss on bioturbation depends on how the functional traits of species co-vary with their risk of extinction. Similarly, loss of rare (and therefore more prone to extinction) species had no effect on above-ground net primary productivity in plots on a prairie grassland over 2 growing seasons, whereas loss of dominant species led to decreased productivity (Smith & Knapp 2003). These results contrast with decreases in ecosystem functioning with decreasing species richness found in similar studies, but with random species assemblages (Hector et al. 2002, Tilman et al. 2002).

Acknowledging this in experimental designs could be done by selecting combinations of species as before, but with weighting for a species to be included based on a measure of the relative likelihood of that species going extinct (Srivastava 2002). However, this would mean that a general effect of biodiversity on ecosystem functioning would be confounded by systematic changes in species composition (Lawton et al. 1998, Hector et al. 2002). A possible solution to this would be to use both approaches in the experimental design (Srivastava 2002). The replicates based on randomly selected combinations of species would act as a null-model, and comparison with replicates with weighted selection of species would allow the separation of the effect of losing biodiversity per se from the effect of losing biodiversity selectively (Srivastava 2002). A further step would be to allow the experiment to continue over many generations and so allow extinction events to occur naturally (Srivastava 2002). This type of

approach combined with some form of perturbation has been applied (Petchey et al. 1999, Griffiths et al. 2000), but without the null-model treatments.

Concern over the consequences of how species have been selected for replicates in BEF experiments has led to much strong debate over the role of 'hidden treatments'. Overall, this debate has been very healthy, resulting in a greater rigour in BEF experimental designs. There is still much to debate, and methods of choosing which species to include in replicates in model system experiments will remain critical in interpreting the results and placing them in their correct context.

Spatial scale

Here, spatial scale is defined as the dimension of observed entities and phenomena in space (O'Neill & King 1998, Schneider 2001). Most research into BEF has focussed on the small spatial scale, and experimental design has required these systems to be closed to the surroundings (Bengtsson et al. 2002). It could be misleading to simply extrapolate results from small-scale experiments to a larger scale, assuming that processes and mechanisms do not change, an assumption unlikely to be true for ecosystems. There have been 3 main hypotheses proposed to explain increased ecosystem functioning with increased species richness at the small spatial scales used in model systems: species complementarity, positive interactions between species and hidden treatments. These mechanisms work at the small, local scale, assuming that species interactions and their functional traits drive the BEF relationship (Bengtsson et al. 2002). However, at larger spatial scales, it is likely to be variation in resources and abiotic factors that are the main drivers (Huston 1994, Anderson 1995). It has been suggested that these scale differences may lead to differences in the main effects of changing diversity, with changes at the small scale mainly affecting rates of ecosystem processes, whereas, at the larger scale, changing diversity will tend to alter levels of ecosystem resilience and stability (Bengtsson et al. 2002).

Not only may scale change the type of relationship between biodiversity and ecosystem functioning, it may also alter the level of effect. As one moves towards community-wide scales in heterogeneous systems, effects of biodiversity could be overwhelmed by other factors (Levine et al. 2002), such as abiotic factors. Thus, at increasing scales, effects of biodiversity may become harder to detect against a background of other effects.

Integration of scale as a factor in model system design will be intimately related to, and complicated by, the integration of multiple trophic levels (see 'Multiple trophic levels' section). Species in the higher trophic levels tend to have larger home ranges, and the scales at which their dynamics operate will be greater. This implies that studies conducted at one trophic level may not be directly relevant to studies at another trophic level, even if the studies have been conducted at the same spatial resolution (Bengtsson et al. 2002). A further implication is that including multiple trophic levels will change the model system from working at a single spatial scale, to working at multiple interacting spatial scales. If changes in biodiversity at different scales have different effects (Bengtsson et al. 2002), interaction of mechanisms between scales could be significant in determining which are the main effects of biodiversity loss in an ecosystem.

One way of allowing a range of spatial scales in model systems would be to simply expand the size of enclosure or plot used in the experiments. This approach has been utilised (Petchey et al. 1999, Hulot et al. 2000), but as part of research of multi-trophic interactions and not to specifically examine biodiversity–ecosystem functioning. A similar approach might be to connect blocks of micro- or mesocosms together, allowing interactions over a larger spatial scale (Bengtsson et al. 2002). These approaches allow only a limited increase in the range of scales for model ecosystems, but offer an important next step in addressing the role of scale in BEF relations.

Other methodologies such as utilising large-scale enclosures of ecosystems, either using natural boundaries such as lakes and islands (Carpenter et al. 2001), or artificial boundaries (Krebs et al. 1995), have been used. These designs would be better for continuing to address the role of scale in BEF at larger spatial scales, but offer much less control of starting conditions, fewer replicates (often only 1), and are far more difficult and costly to monitor. Realistically, model systems will be limited in the scales that they can incorporate. However, if they can be expanded in size, albeit to a limited degree, from the current small spatial scales, it may be possible to integrate results from these experiments with experiments based on different experimental designs and incorporating larger spatial scales. Simply having multiple spatial scales will allow a start to be made in addressing the role of scale in BEF.

Working at increasing spatial scales has the associated problems of increased requirements of space, time and financial costs, with a possible compromise being obtained by a reduction in the number of replicates. The great need to address the role of scale in BEF with the associated likelihood of decreasing the number of experimental replicates has led to the suggestion that we may have to be prepared to accept a lower level of statistical rigour in experimental design and analysis (Bengtsson et al. 2002).

Dimensional analysis offers potential for designing model systems that will allow results to be extrapolated to larger spatial scales (Petersen & Hastings 2001, Englund & Cooper 2003). The technique involves determining dimensionless variables by rearranging different measurements of the systems so that their units cancel each other out. The relative values of the component variables can be adjusted to maintain the same value of the dimensionless variable in the model system as in the ecosystem. The lack of flexibility in controlling organism characteristics for scaling in this way strongly limits this approach (Englund & Cooper 2003). However, the greatest limitation is due to our lack of knowledge of the scaling relationships that drive ecosystem dynamics (Petersen & Hastings 2001, Englund & Cooper 2003). These limitations will be exacerbated in multi-trophic studies, since the level and dynamics of scaling will vary between species, making a common scaling relationship for all species involved unlikely.

An alternative approach is the translation of results across scales using scaling models, incorporating known empirical relationships or mechanistic models (Schneider et al. 1997, Englund & Cooper 2003). Application of this methodology to open predation experiments (Englund et al. 2001) has been used to identify scale domains (Wiens 1989), ranges in scale over which there is no or only weak scale dependence, but between which there is strong scale dependence or changes in the drivers of the process of interest. The defining of scaling relationships in the open predation experiments was accomplished via a meta-analysis of past experiments performed at different scales (Englund et al. 2001). Identifying such scale domains explicitly shows the scales over which we need to be particularly careful about scaling effects. Again the major limitation in translating results between spatial scales is our lack of knowledge of scaling processes; complex multi-scale experiments or meta-analyses are required to develop empirical scaling relationships, and scaling mechanisms are poorly understood. Multi-scale experiments in which the same treatment is applied at different spatial scales as part of an integrated experimental design will generate less ambiguous results than meta-analyses (Chen et al. 1997).

As part of an assessment of the various research paradigms for addressing multi-scale processes, strategic cyclical scaling was considered to be the most appropriate paradigm to use (Root & Schneider 1995). This paradigm involves using observations at large scales to direct the testing of mechanistic processes in small-scale experiments, the results of which are used to make predictions at larger scales. This cyclic process builds up knowledge of processes and mechanisms across a spectrum of scales. Gaps in our knowledge of scaling processes for a particular system can be explicitly visualised with scope diagrams (Schneider et al. 1997), and these represent a possible guiding methodology in the recommended strategic cyclical scaling paradigm.

Multiple trophic levels

Ecosystems are composed of several trophic levels with interactions between these trophic levels. Therefore, any theory of BEF that does not take this dimension of ecosystems into account will necessarily be incomplete. This raises concerns about current work using model systems. There has been an overwhelming concentration on experimentation with plants (a single trophic level), and multiple trophic levels have been explicitly incorporated into only a few model systems addressing BEF (Naeem & Li 1997, Petchey et al. 1999, Naeem et al. 2000, Duffy et al. 2001, 2003, 2005).

Effects of biodiversity are just as likely, or more likely, to be driven by the consumer level as the level of plants (Petchey et al. 1999, Duffy 2002, 2003), and these effects will lead to further effects via trophic interactions (Naeem et al. 2000, Worm & Duffy 2003, Duffy et al. 2005). It has been suggested that changes in ecosystem processes resulting from changes in biodiversity at the consumer level are likely to be more idiosyncratic than would be the case with changes in plant biodiversity (Duffy 2002). So, not only is there the complication of added complexity due to interaction between individuals in the different trophic levels, but changes in biodiversity at these different levels could result in very different ecosystem responses (Petchey et al. 1999, Duffy 2002, 2003, Worm & Duffy 2003). This has experimental support. Aquatic microcosms have shown different patterns of species loss at different trophic levels, resulting in complex responses in ecosystem functioning (Petchey et al. 1999). In estuarine mesocosms increased grazer diversity led to increased seagrass diversity only when a predator species was present (Duffy et al. 2005).

Many of the problems of extending model systems to incorporate multiple trophic levels are in common with those of extending spatial and/or temporal scale, these 3 aspects of ecosystems being intimately linked. The inclusion of multiple trophic levels requires systems of larger spatial scales and therefore has higher financial and physical costs, resulting in a trade-off with replication. A further complication is that the effects of different trophic levels can operate at different temporal and spatial scales (Bengtsson et al. 2002, Raffaelli et al. 2002, Raffaelli 2006, in this Theme Section). For example, the soil community appears to respond with a marked time lag behind plants, indicating a decou-

pling of temporal dynamics between these 2 trophic compartments (Van der Putten et al. 2000).

Organisms at some trophic levels tend to be more easily identified, and by more people, than those at other trophic levels, so that taxonomic resolution declines at lower trophic levels, especially in invertebrate (but not plant-based) webs (Raffaelli et al. 2002). Although model systems can overcome this problem in principle by using clearly separate species in the system, biases will inevitably result, whilst the systematics of many lower taxa, such as those in sediments (meiofauna) and soils (mites), are extremely challenging or even unresolved.

Despite these issues, multi-trophic model systems have been constructed, although they have rarely addressed BEF directly and have incorporated limited numbers of species (Naeem & Li 1997, Petchey et al. 1999, Naeem et al. 2000, Duffy et al. 2001, 2003, 2005). As with incorporating larger spatial scales, there is potential here also for using naturally enclosed ecosystems such as lakes and islands (Carpenter et al. 1995, 2001, Krebs et al. 1995). However, the level of control, the number of replicates and the level of information on the states of the system with time in these experiments will be inherently much lower.

Understanding BEF dynamics in the context of multiple-trophic levels will be essential to making properly informed predictions of the consequences of decreasing biodiversity. As a result, the need for incorporating multiple-trophic levels in model systems is probably the most pressing (Raffaelli 2006). However, achieving this represents a considerable challenge.

Variation

There are 3 principal sources of variation which need to be considered when addressing BEF issues. Firstly, biodiversity, viewed as a constant and independent factor within each biodiversity treatment in the experimental design, will be variable over space and time in real ecosystems. Secondly, there will be variation in the level of ecosystem functioning under consideration due to biodiversity and, thirdly, residual variation due to other independent factors.

Ecosystem functioning will vary in space and time, and these dynamics need to be understood. Model systems, with high levels of replication, and strong control over initial conditions are ideal for measuring this inherent variation in functioning at constant biodiversity, at least at small spatial scales. However, current experimental designs are not conducive to measuring the effect of variation (temporal or spatial) in biodiversity on ecosystem functioning (Benedetti-Cecchi 2003).

The majority of ecological experiments attempt to relate the effect of mean intensity of the independent variable (biodiversity) to the resulting mean intensity and variation in the dependent variable (ecosystem functioning). Variation in the independent variable tends to be ignored, with experimental design incorporating the independent variable as a constant factor (Benedetti-Cecchi 2003). However, experimental designs that begin to address the role of temporal and spatial variation in the independent variable are starting to appear (e.g. Navarrete 1996, Benedetti-Cecchi 2000). For example, variance in the spatial distribution of resource populations can be sensitive to changes in the variance of the consumer–resource trophic interaction (Benedetti-Cecchi 2000), in addition to the mean effect of the consumers. A consumer–resource model incorporating this effect was able to replicate patterns in empirical data, which previous models had not been able to produce (Benedetti-Cecchi 2000).

Such considerations have not yet been applied to most BEF experiments. The general experimental design has been addressed recently (Benedetti-Cecchi 2003), allowing the separation of effects between the intensity and variability (temporal or spatial) of an ecosystem process driver. The design must be structured so that the intensity and variability are arranged as independent factors (i.e. orthogonal). It should also be noted that this independence does not hinder examining relationships between the 2 factors via tests on interaction terms (Benedetti-Cecchi 2003).

Although these experimental designs separating intensity and variability of independent factors were constructed with consumer–resource interactions in mind, it should be possible to adapt them for addressing BEF questions, and model systems offer great potential in this respect. The strong control in determining starting conditions and the relative ease of producing replicates should provide for experimental designs in which biodiversity level and variation are orthogonal, therefore allowing for the separation of effects due to variation in biodiversity. However, since heterogeneity is linked intricately to scale, the small spatial and short temporal scales of most model systems impose a limitation on the levels of variation in biodiversity that can be addressed using this methodology.

Environmental stochasticity

Related to the role of the environment in BEF is the issue of environmental stochasticity. In the real world, environmental variables change with time, and these changes incorporate a stochastic component. A stochastic variable has a random probability distribution that can be analysed statistically, but not predicted precisely. Demographic stochasticity describes un-

certainties relating to characteristics of individuals, whereas environmental stochasticity describes uncertainties imposed at the population level by the environment (Bonsall & Hastings 2004). Stochasticity in models has often been treated as random fluctuations, with no temporal correlation (Halley 1996), also known as white noise. However, so-called 'reddened' spectra, with positive temporal autocorrelation between fluctuations (Halley 1996), may be more appropriate for environmental fluctuations (Mandelbrot & Wallis 1969), particularly in marine systems (Steele 1985). Environmental stochasticity is a well-known determinant of population dynamics (May 1973, Lande 1993), and long-term population records appear to show reddened spectra (Pimm & Redfearn 1988), suggesting a connection between reddening of environmental drivers and reddening of population dynamics. If this connection is real, the level of temporal autocorrelation in environmental drivers is likely to play a role in BEF dynamics, possibly causing BEF relationships to change independently of changes in species composition (Johnson 2000).

Aquatic microbial systems are excellent for testing connections between characteristics of environmental variation and BEF, providing high levels of control over environmental conditions and the capacity to run experiments over many generations (ensuring that the results are driven by population dynamics). They have already been used to demonstrate that single-species microbial population dynamics can be sensitive to the colour of temperature fluctuation with time (constant, white, or reddened) (Petchey 2000). In a more recent experiment, similar temperature fluctuations have been applied in conjunction with varying species richness to test for effects on changing total community biomass (Petchey et al. 2002a). Results indicated that change in total biomass was unaffected by the temperature regime. However, the authors point out that biomass in each microcosm was estimated only twice and therefore measured long-term change in biomass and not fluctuations in biomass, and that this restriction meant they were unable to separate change due to stochastic fluctuations and change due to directional changes (Gaston & McArdle 1994). To do this requires experimental designs with a number of samples and a frequency of sampling appropriate for obtaining an estimate of the general trend and therefore enabling the partitioning of change due to stochastic environmental fluctuations (Gaston & McArdle 1994). Future experiments aiming to relate environmental stochasticity to variation in ecosystem functioning through time will also need to ensure that a number of samples is taken from each microcosm at each time. This will enable an estimate of sampling error for each ecosystem functioning estimate and therefore enable an estimate of the true but unknown level of ecosystem functioning (Gaston & McArdle 1994). Any measure of variation of ecosystem functioning across time will then be an estimate of the variation in ecosystem functioning, rather than an estimate of the variation in ecosystem functioning estimates (Gaston & McArdle 1994).

CONCLUSIONS AND RECOMMENDATIONS

This review has shown that model systems have been much used recently in trying to gain an understanding of BEF. Generally using isolated, small-scale collections of species, they are simplistic systems that are designed to be tractable whilst distilling out essential characteristics of BEF. The high level of control over starting and environmental conditions, along with good levels of reproducibility and replication make model systems a potentially very powerful tool for understanding BEF.

These systems have been heavily criticised in the literature because of their artificiality and their simplicity compared to real ecosystems. But this is to view model systems in the wrong context. The criticisms are really limitations to the methodology; they do not make the methodology invalid or less valid than alternative approaches which have their own faults. Model systems need to be viewed as part of a holistic approach to understanding BEF that utilises a range of methodologies. In this context model systems are a tool with strengths and weaknesses, but with strengths that complement those from other methodologies. The complexity of ecosystems is enormous, and model systems are a reductionist methodology designed to address possible links between biodiversity and ecosystem functioning in a simplistic and therefore understandable context. They are tools for examining these links in isolated fine detail; they are not designed to address how ecosystems function as a whole. The high level of control and monitoring that is possible makes these systems particularly suitable for investigating mechanistic processes. We therefore endorse the strategic cyclic paradigm (Root & Schneider 1995) with an emphasis on the use of model systems to investigate possible mechanisms underlying BEF, evidence of which can then be tested for in larger-scale experiments and field observations (Fig. 2). Although this paradigm was suggested originally to deal with issues of different scales between the different methodologies used to investigate BEF, the cyclic procedure is just as valid for dealing with other issues, such as multiple trophic levels.

Model systems, particularly the microbial microcosms, offer a strong connection with theoretical mod-

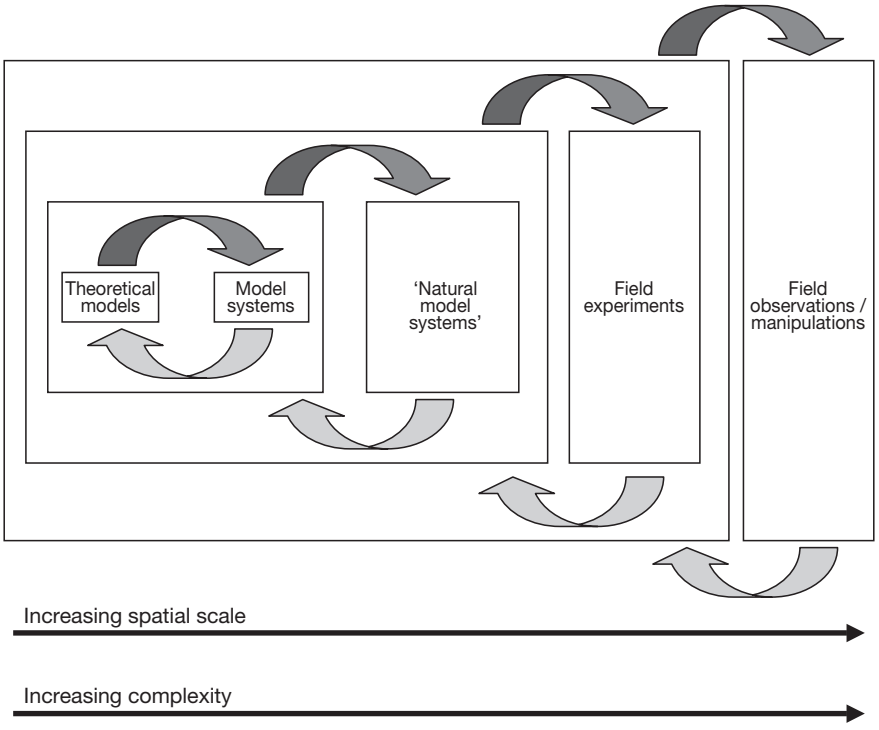

Fig. 2. Suggested nested hierarchical integration of different methodologies in understanding mechanisms driving biodiversity–ecosystem functioning (BEF) based on the strategic cyclical scaling paradigm (Root & Schneider 1995). Light grey arrows indicate verification processes, dark grey arrows indicate the generation of mechanistic hypotheses and predictions. Connections between any 2 methodologies are not precluded but would represent cases where extra caution in interpretation may be required

elling. This is of great potential benefit as predictions from theoretical models (including many of the simplifying assumptions) can be rigorously tested under highly controlled conditions. We suggest that particular use of this strong link be made to generate theoretical models of BEF that can be strengthened by the validation of matching of model predictions from theoretical models with model system behaviour.

If there is to be effective integration of the different methodologies, there will need to be a common language. Although dimensional analysis (Petersen & Hastings 2001, Englund & Cooper 2003) can be difficult and care must be taken to provide ecologically meaningful dimensions, it provides a framework for translating results and concepts across different methodologies and between different systems. We therefore advocate its regular use in model systems and suggest that it be part of an increase in meta-analyses trying to identify commonalities and differences across systems as well as scales, providing a more complete view of mechanisms underlying BEF.

A greater effort needs to be made to put the results of model systems into proper ecological context. This will enhance coherency in the general BEF debate by highlighting common factors and differences between different model systems, and between model systems and other methodologies. For example, some model systems will be driven by population dynamics, whereas others will be strongly driven by initial starting conditions, making it difficult to integrate the results from the 2 groups. Similarly some model systems will be based on ecosystems strongly influenced by stochastic factors and others on ecosystems where this is not the case. BEF dynamics will alter according to such considerations, and results from model systems need to be put into context to facilitate an understanding of how these considerations can alter BEF. This needs to be recognised so that debate is not over differences due to a lack of understanding of relevant contexts, but over true differences and similarities in different situations.

A second vigorous debate in the literature has been over the implications of hidden treatments with species selection conducted randomly, leading to the questioning of the significance of results from BEF model systems. The debate over the sampling effect is illustrative of what can be gained from the use of model systems. It has led to a clarification and increased understanding of possible mechanisms linking biodiversity to ecosystem functioning, and to much more robust and rigorous experimental designs for the future.

There are 2 main concerns over hidden treatments: the match between the random selection of species method and reality and the resulting effects of the statistical sample distributions on the validity of the statistical analyses. The random-selection method can reflect reality in ecosystems in which population dynamics are strongly driven by stochastic processes or by a number of interacting processes (Fridley 2001). The relevance of randomisation must therefore be

judged in the context of the ecosystem being addressed. If random selection is not realistic, species-richness treatments could be chosen using expert knowledge or using weighted choices based on likelihood of extinction. Microcosm experiments in which many generations will elapse could be allowed to progress for different amounts of time, allowing extinction of species to occur 'naturally' and therefore determine realistic species combinations to be used in BEF model systems.

ANOVA has been used extensively in the analysis of BEF model systems. Hidden treatment effects can lead to heterogeneity of variance, breaking one of the assumptions underlying this method. However, ANOVA is robust to this assumption under certain circumstances. Problems are much more pronounced when sample sizes differ (Box 1953). However, if samples are balanced, heterogeneity of variance is usually only a serious problem when one of the samples has a larger variance than the others (Underwood 1998). The effects of heterogeneity of variance are also significantly reduced as the number of treatments increases and when sample sizes are large (usually taken to be >6) (Underwood 1998). If heterogeneity of variance is judged to be likely, these considerations in experimental design should be regarded as a priority.

A second concern about the effects of hidden treatments on the interpretation of results from ANOVA is asymmetry in similarity of quasi-replicates between species-rich and species-poor treatments, leading to differences in the interpretation of experimental error. This can be minimised, as well as gaining more information about the mechanisms involved, by explicitly incorporating the likely environmental sources of experimental error as factors in the ANOVA design.

Ecosystem functioning has usually been measured by a single metric (e.g. primary production). This collapsing of a multi-dimensional characteristic into a single dimension offers tractability, but decreases the level of information gained, as well as leading to possible misinterpretations due to a lack of perspective. We suggest an expansion in the number of metrics used to measure ecosystem functioning in a model system as a way of increasing the level of understanding of the mechanistic processes driving BEF.

Similarly, 'biodiversity' is a multi-dimensional concept that is collapsed into 1 dimension for BEF model-system studies, usually species richness, since this is a metric that is both intuitive and easily measured and manipulated. Future experiments will need to address other components of biodiversity (e.g. evenness) to gain a more complete understanding of BEF. This could be done by repeating experiments using a measure of biodiversity other than species richness or with the new measure plus richness. Dealing with the multiple dimensions of biodiversity is likely to be harder than expanding the dimensionality of ecosystem functioning, as biodiversity is a treatment rather than a response, and the possible permutations within and amongst the newly applied metrics of biodiversity will increase the already large number of replicates required for BEF model system studies. An ANOVA design with crossed biodiversity factors would be appropriate and allow for the measurement of interaction terms between the different biodiversity dimensions.

The focus on species richness as the metric for biodiversity has started to change towards functional richness. This is a trend that should continue, but it will not be without difficulties. Functional richness is a measure based on ecological behaviour and is therefore much more suitable than species richness for generating a mechanistic understanding of BEF. Where functional richness has been compared to species richness, it has shown greater explanatory power for levels of ecosystem functioning (Petchey et al. 2004). The main difficulty is generating sensible functional groupings that relate properly to the ecosystem function of interest, although progress in this has already been made (Petchey & Gaston 2002, Petchey 2004).

Although model systems are by design simplistic, it is clear that they can be extended or experimental designs can be altered, in order to investigate issues of BEF not yet addressed. Complexity can be increased, whilst keeping the systems tractable, so that the role of vital components of BEF can be investigated, such as spatial scale and multiple trophic levels. The key is to ensure that these extensions integrate well with other experimental methodologies.

Natural microcosms are small, contained habitats that are naturally populated by minute organisms (Srivastava et al. 2004). Although they offer a valuable link between the simpler model systems and extremely complex ecosystems (Srivastava et al. 2004), they have been used relatively little and not to address BEF. Examples of such systems include aquatic rock pools (Romanuk & Kolasa 2002), marine pen shells (Munguia 2004), pitcher plants (Kneitel & Miller 2002) and tree holes (Fincke et al. 1997, Srivastava & Lawton 1998). These systems offer several benefits: real multiple trophic combinations of species with a shared evolutionary past and a level of openness of individual microcosms arranged in a natural hierarchical spatial structure (Srivastava et al. 2004). These benefits result in a reduction in the level of control and tractability, but not to the levels of whole ecosystems. The decrease in tractability and increase in generality mean that use of model systems and natural microcosms can be complementary methodologies—model systems testing if hypothesised effects occur, whilst natural microcosms

test if these effects are important in the real world (Srivastava et al. 2004).

Model systems have already proved to be of value in BEF research and are likely to continue to be useful. They will not provide us with an overall understanding of how ecosystems function, but they are capable of providing valuable insights into how small, isolated components of ecosystems work. Model systems have great strengths, but also strong limitations that must be recognised when trying to put experimental results into context. They offer an extremely valuable methodology that must be integrated with other methodologies, such as observation and manipulation of ecosystems, in order to provide a more complete understanding of BEF. This integration of methodologies will be challenging, but is a necessity if we are to understand BEF. Marine ecology has a long history in the experimental design of, and interpretation of results from, model systems. This knowledge and experience can play a valuable role in enhancing our understanding of BEF in marine ecosystems and more generally.

Acknowledgements. This work was funded by the Natural Environment Research Council (NERC) (Ref.: NER/A/S/2003/00577). We thank L. Benedetti-Cecchi and 3 anonymous reviewers for helpful comments and criticism on the manuscript. M. Termansen and J. Chamberlain provided helpful comments on an earlier draft of this paper, and J. Smart and J. Minx were always available for helpful discussions.

LITERATURE CITED

Anderson J (1995) Soil organisms as engineers: microsite modulation of macroscale processes. In: Jones CG, Lawton JH (eds) Linking species and ecosystems. Chapman & Hall, New York, p 94–106

Benedetti-Cecchi L (2000) Variance in ecological consumer–resource interactions. Nature 407:370–374

Benedetti-Cecchi L (2003) The importance of the variance around the mean effect size of ecological processes. Ecology 84:2335–2346

Benedetti-Cecchi L (2004) Increasing accuracy of causal inference in experimental analyses of biodiversity. Funct Ecol 18:761–768

Bengtsson J, Engelhardt K, Giller P, Hobbie S, Lawrence D, Levine J, Vilà M, Wolters V (2002) Slippin' and slidin' between the scales: the scaling components of biodiversity–ecosystem functioning relations. In: Loreau M, Naeem S, Inchausti P (eds) Biodiversity and ecosystem functioning: synthesis and perspectives. Oxford University Press, Oxford, p 209–220

Bloesch J, Bossard P, Buhrer H, Burgi HR, Uehlinger U (1988) Can results from limnocorral experiments be transferred to *in situ* conditions? Hydrobiologia 159:297–308

Bonsall MB, Hastings A (2004) Demographic and environmental stochasticity in predator–prey metapopulation dynamics. J Appl Ecol 73:1043–1055

Box GEP (1953) Non-normality and tests on variances. Biometrika 40:318–335

Carpenter SR (1996) Microcosm experiments have limited relevance for community and ecosystem ecology. Ecology 77:677–680

Carpenter SR, Chisholm SW, Krebs CJ, Schindler DW, Wright RF (1995) Ecosystem experiments. Nature 269:324–327

Carpenter SR, Cole JJ, Hodgson JR, Kitchell JF and 6 others (2001) Trophic cascades, nutrients, and lake productivity: whole-lake experiments. Ecol Monogr 71:163–186

Chapin FS III, Walker BH, Hobbs RJ, Hooper DU, Lawton JH, Sala OE, Tilman D (1997) Biotic control over the functioning of ecosystems. Science 277:500–504

Chapin FS III, Zavaleta ES, Eviner VT, Naylor RL and 8 others (2000) Consequences of changing biodiversity. Nature 405:234–242

Chen CC, Petersen JE, Kemp WM (1997) Spatial and temporal scaling of periphyton growth on walls of estuarine mesocosms. Mar Ecol Prog Ser 155:1–15

Costanza R, dArge R, deGroot R, Farber S and 9 others (1997) The value of the world's ecosystem services and natural capital. Nature 387:253–260

Cottingham KL, Brown BL, Lennon JT (2001) Biodiversity may regulate the temporal variability of ecological systems. Ecol Lett 4:72–85

Drake JA, Huxel GR, Hewitt CL (1996) Microcosms as models for generating and testing community theory. Ecology 77: 670–677

Duffy JE (2002) Biodiversity and ecosystem function: the consumer connection. Oikos 99:201–219

Duffy JE (2003) Biodiversity loss, trophic skew and ecosystem functioning. Ecol Lett 6:680–687

Duffy JE, Macdonald KS, Rhode JM, Parker JD (2001) Grazer diversity, functional redundancy, and productivity in seagrass beds: an experimental test. Ecology 82:2417–2434

Duffy JE, Richardson JP, Canuel EA (2003) Grazer diversity effects on ecosystem functioning in seagrass beds. Ecol Lett 6:637–645

Duffy JE, Richardson JP, France KE (2005) Ecosystem consequences of diversity depend on food chain length in estuarine vegetation. Ecol Lett 8:301–309

Ehrlich PR, Wilson EO (1991) Biodiversity studies: science and policy. Science 253:758–762

Emmerson MC, Solan M, Emes C, Paterson DM, Raffaelli D (2001) Consistent patterns and the idiosyncratic effects of biodiversity in marine ecosystems. Nature 411:73–77

Englund G, Cooper SD (2003) Scale effects and extrapolation in ecological experiments. Adv Ecol Res 33:161–202

Englund G, Cooper SD, Sarnelle O (2001) Application of a model of scale dependence to quantify scale domains in open predation experiments. Oikos 92:501–514

Fincke OM, Yanoviak SP, Hanschu RD (1997) Predation by odonates depresses mosquito abundance in water-filled tree holes in Panama. Oecologia 112:244–253

Fridley JD (2001) The influence of species diversity on ecosystem productivity: How, where, and why? Oikos 93: 514–526

Fridley JD (2002) Resource availability dominates and alters the relationship between species diversity and ecosystem productivity in experimental plant communities. Oecologia 132:271–277

Gamble JC (1991) Mesocosms, statistical and experimental design considerations. In: Lalli CM (ed) Experimental marine ecosystems: a review and recommendations. Springer-Verlag, Heidelberg, p 188–196

Gaston KJ, McArdle BH (1994) The temporal variability of animal abundances: measures, methods and patterns. Philos Trans R Soc Lond B 345:335–358

Gerhart DZ, Likens GE (1975) Enrichment experiments for determining nutrient limitation: four methods compared.

Limnol Oceanogr 20:649–653
Griffiths BS, Ritz K, Bardgett RD, Cook R and 8 others (2000) Ecosystem response of pasture soil communities to fumigation-induced microbial diversity reductions: an examination of the biodiversity–ecosystem function relationship. Oikos 90:279–294
Halley JM (1996) Ecology, evolution and 1/f-noise. Trends Ecol Evol 11:33–37
Hector A, Loreau M, Schmid B, BIODEPTH project (2002) Biodiversity manipulation experiments: studies replicated at multiple sites. In: Loreau M, Naeem S, Inchausti P (eds) Biodiversity and ecosystem functioning: synthesis and perspectives. Oxford University Press, Oxford, p 36–46
Hulot FD, Lacroix G, Lescher-Moutoue F, Loreau M (2000) Functional diversity governs ecosystem response to nutrient enrichment. Nature 405:340–344
Huston MA (1994) Biological diversity. The coexistence of species on changing landscapes. Cambridge University Press, Cambridge
Huston MA (1997) Hidden treatments in ecological experiments: re-evaluating the ecosystem function of biodiversity. Oecologia 110:449–460
Huston MA (1999) Microcosm experiments have limited relevance for community and ecosystem ecology: synthesis of comments. Ecology 80:1088–1089
Huston MA, McBride AC (2002) Evaluating the relative strengths of biotic versus abiotic controls on ecosystem processes. In: Loreau M, Naeem S, Inchausti P (eds) Biodiversity and ecosystem functioning: synthesis and perspectives. Oxford University Press, Oxford, p 47–60
Huston MA, Aarssen LW, Austin MP, Cade BS and 8 others (2000) No consistent effect of plant diversity on productivity. Science 289:1255 (abstract)
Johnson KH (2000) Trophic–dynamic considerations in relating species diversity to ecosystem resilience. Biol Rev Camb Philos Soc 75:347–376
Kneitel JM, Miller TE (2002) Resource and top-predator regulation in the pitcher plant (*Sarracenia purpurea*) inquiline community. Ecology 83:680–688
Krebs CJ, Boutin S, Boonstra R, Sinclair ARE, Smith JNM, Dale MRT, Martin K, Turkington R (1995) Impact of food and predation on the snowshoe hare cycle. Nature 269: 1112–1115
Lalli CM (ed) (1991) Experimental marine ecosystems: a review and recommendations. Springer-Verlag, Heidelberg
Lande R (1993) Risks of population extinction from demographic and environmental stochasticity and random catastrophes. Am Nat 142:911–927
Lawler SP (1998) Ecology in a bottle: using microcosms to test theory. In: Bernardo J (ed) Experimental ecology: issues and perspectives. Oxford University Press, Oxford, p 236–253
Lawton JH (1994) What do species do in ecosystems? Oikos 71:367–374
Lawton JH (1995) Ecological experiments with model systems. Science 269:328–331
Lawton JH (1996) The Ecotron facility at Silwood Park: the value of ''big bottle'' experiments. Ecology 77:665–669
Lawton JH, Naeem S, Thompson LJ, Hector A, Crawley MJ (1998) Biodiversity and ecosystem function: getting the Ecotron experiment in its correct context. Funct Ecol 12: 848–852
Levine JM, Kennedy T, Naeem S (2002) Neighbourhood scale effects of species diversity on biological invasions and their relationship to community patterns. In: Loreau M, Naeem S, Inchausti P (eds) Biodiversity and ecosystem functioning: synthesis and perspectives. Oxford University Press, Oxford, p 114–124
Loreau M (2000) Biodiversity and ecosystem functioning: recent theoretical advances. Oikos 91:3–17
Loreau M, Hector A (2001) Partitioning selection and complementarity in biodiversity experiments. Nature 413:548–548
Mandelbrot B, Wallis J (1969) Some long-run properties of geophysical records. Water Resour Res 5:321–340
Manly B (1992) The design and analysis of research studies. Cambridge University Press, Cambridge
May RM (1973) Stability in randomly fluctuating versus deterministic environments. Am Nat 107:621–650
McGrady-Steed J, Harris PM, Morin PJ (1997) Biodiversity regulates ecosystem predictability. Nature 390:162–165
Mulder CPH, Koricheva J, Huss-Danell K, Högberg P, Joshi J (1999) Insects affect relationships between plant species richness and ecosystem processes. Ecol Lett 2:237–246
Munguia P (2004) Successional patterns on pen shell communities at local and regional scales. J Anim Ecol 73:64–74
Naeem S, Li S (1997) Biodiversity enhances ecosystem reliability. Nature 390:507–509
Naeem S, Wright JP (2003) Disentangling biodiversity effects on ecosystem functioning: deriving solutions to a seemingly insurmountable problem. Ecol Lett 6:567–579
Naeem S, Hahn DR, Schuurman G (2000) Producer–decomposer co-dependency influences biodiversity effects. Nature 403:762–764
Navarrete SA (1996) Variable predation: effects of whelks on a mid-intertidal successional community. Ecol Monogr 66: 301–321
O'Neill RV, King AW (1998) Homage to St. Michael; or, why are there so many books on scale. In: Peterson DL, Parker VT (eds) Ecological scale: theory and applications. Columbia University Press, New York, p 3–15
Paine RT (2002) Trophic control of production in a rocky intertidal community. Science 296:736–739
Pauly D, Christensen V, Dalsgaard J, Froese R, Torres F Jr (1998) Fishing down marine food webs. Science 279: 860–863
Petchey OL (2000) Environmental colour affects aspects of single-species population dynamics. Proc R Soc Lond B 267:747–754
Petchey OL (2003) Integrating methods that investigate how complementarity influences ecosystem functioning. Oikos 101:323–330
Petchey OL (2004) On the statistical significance of functional diversity effects. Funct Ecol 18:297–303
Petchey OL, Gaston KJ (2002) Functional diversity (FD), species richness and community composition. Ecol Lett 5: 402–411
Petchey OL, McPhearson PT, Casey TM, Morin PJ (1999) Environmental warming alters food-web structure and ecosystem function. Nature 402:69–72
Petchey OL, Casey T, Jiang L, McPhearson PT, Price J (2002a) Species richness, environmental fluctuations, and temporal change in total community biomass. Oikos 99:231–240
Petchey OL, Morin PJ, Hulot FD, Loreau M, McGrady-Steed J, Naeem S (2002b) Contributions of aquatic model systems to our understanding of biodiversity and ecosystem functioning. In: Loreau M, Naeem S, Inchausti P (eds) Biodiversity and ecosystem functioning: synthesis and perspectives. Oxford University Press, Oxford, p 127–138
Petchey OL, Hector A, Gaston KJ (2004) How do different measures of functional diversity perform? Ecology 85:847–857
Petersen JE, Hastings A (2001) Dimensional approaches to scaling experimental ecosystems: designing mousetraps to catch elephants. Am Nat 157:324–333

Pimm SL, Redfearn A (1988) The variability of population densities. Nature 334:613–614
Pimm SL, Russell GJ, Gittleman JL, Brooks TM (1995) The future of biodiversity. Science 269:347–350
Purvis A, Hector A (2000) Getting the measure of biodiversity. Nature 405:212–219
Raffaelli D (2006) Biodiversity and ecosystem functioning: issues of scale and trophic complexity. Mar Ecol Prog Ser 311:285–294 (in this Theme Section)
Raffaelli D, van der Putten WH, Persson L, Wardle DA and 5 others (2002) Multi-trophic dynamics and ecosystem processes. In: Loreau M, Naeem S, Inchausti P (eds) Biodiversity and ecosystem functioning: synthesis and perspectives. Oxford University Press, Oxford, p 147–154
Romanuk TN, Kolasa J (2002) Environmental variability alters the relationship between richness and variability of community abundances in aquatic rock pool microcosms. Ecoscience 9:55–62
Root TL, Schneider SH (1995) Ecology and climate: research strategies and implications. Science 269:334–341
Sala OE (2001) Ecology–Price put on biodiversity. Nature 412: 34–36
Schlapfer F, Schmid B (1999) Ecosystem effects of biodiversity: a classification of hypotheses and exploration of empirical results. Ecol Appl 9:893–912
Schneider DC (2001) The rise of the concept of scale in ecology. BioScience 51:545–553
Schneider DC, Walters R, Thrush S, Dayton P (1997) Scale-up of ecological experiments: density variation in the mobile bivalve *Macomona liliana*. J Exp Mar Biol Ecol 216: 129–152
Schwartz MW, Brigham CA, Hoeksema JD, Lyons KG, Mills MH, van Mantgem PJ (2000) Linking biodiversity to ecosystem function: implications for conservation ecology. Oecologia 122:297–305
Smith MD, Knapp AK (2003) Dominant species maintain ecosystem function with non-random species loss. Ecol Lett 6:509–517
Solan M, Cardinale BJ, Downing AL, Engelhardt KAM, Ruesink JL, Srivastava DS (2004) Extinction and ecosystem function in the marine benthos. Science 306: 1177–1180
Srivastava DS (2002) The role of conservation in expanding biodiversity research. Oikos 98:351–360
Srivastava DS, Lawton JH (1998) Why more productive sites have more species: an experimental test of theory using tree-hole communities. Am Nat 152:510–529
Srivastava DS, Kolasa J, Bengtsson J, Gonzalez A and 6 others (2004) Are natural microcosms useful model systems for ecology? Trends Ecol Evol 19:379–384
Steele JH (1985) A comparison of terrestrial and marine ecological systems. Nature 313:355–358
Stephenson GL, Hamilton P, Kaushik NK, Robinson JB, Solomon KR (1984) Spatial distribution of plankton in enclosures of three sizes. Can J Fish Aquat Sci 41: 1048–1054
Tilman D, Kareiva P (1997) Spatial ecology. The role of space in population dynamics and interspecific interactions. Princeton University Press, Princeton, NJ
Tilman D, Knops J, Wedin D, Reich P (2002) Plant diversity and composition: effects on productivity and nutrient dynamics of experimental grasslands. In: Loreau M, Naeem S, Inchausti P (eds) Biodiversity and ecosystem functioning: synthesis and perspectives. Oxford University Press, Oxford, p 21–35
Underwood AJ (1998) Experiments in ecology. Their logical design and interpretation using analysis of variance. Cambridge University Press, Cambridge
Van der Putten WH, Mortimer SR, Hedlund K, Van Dijk C and 10 others (2000) Plant species diversity as a driver of early succession in abandoned fields: a multi-site approach. Oecologia 124:91–99
Vitousek PM, Mooney HA, Lubchenco J, Melillo JM (1997) Human domination of Earth's ecosystems. Science 277: 494–499
Walker B, Kinzig A, Langridge J (1999) Plant attribute diversity, resilience, and ecosystem function: the nature and significance of dominant and minor species. Ecosystems 2: 95–113
Wardle DA (1999) Is "sampling effect" a problem for experiments investigating biodiversity–ecosystem function relationships? Oikos 87:403–407
Wiens JA (1989) Spatial scaling in ecology. Funct Ecol 3: 385–397
Worm B, Duffy JE (2003) Biodiversity, productivity and stability in real food webs. Trends Ecol Evol 18:628–632

Editorial responsibility: Martin Solan (Guest Editor), Newburgh, UK

Submitted: April 12, 2005; Accepted: November 10, 2005
Proofs received from author(s): March 6, 2006