Vol. 311: 263–271, 2006 | MARINE ECOLOGY PROGRESS SERIES
Mar Ecol Prog Ser | Published April 13

OPEN ACCESS

How biodiversity affects ecosystem functioning: roles of infaunal species richness, identity and density in the marine benthos

Elena N. Ieno[1,3], Martin Solan[1,*], Paul Batty[2], Graham J. Pierce[1]

[1]Oceanlab, University of Aberdeen, Main Street, Newburgh, Aberdeenshire AB41 6AA, UK
[2]The Scottish Association for Marine Science, Dunstaffnage Marine Laboratory, Oban, Argyll PA37 1QA, UK

[3]*Present address:* Highland Statistics, 6 Laverock Road, Newburgh, Aberdeenshire AB41 6FN, UK

ABSTRACT: The extent to which changes in biodiversity are causally linked to key ecosystem processes is a primary focus of contemporary ecological research. Highly controlled manipulative experiments have revealed significant and positive effects of increased diversity on ecosystem functioning, but uncertainties in experimental design have made it difficult to determine whether such effects are related to the number of species or to effects associated with species identity and density. Using infaunal marine invertebrates, we established 2 parallel laboratory experiments to examine the hypothesis that changes in the composition of benthic macrofauna alter the biogeochemistry of coastal intertidal mudflats. Our study identified clear effects of increased infaunal species diversity on nutrient generation. However, significant species identity and density effects underpin the observed response, reflecting species-specific traits associated with bioturbation. Post-hoc examination of our conclusions using power analysis revealed that, given our experimental design, the probability of finding a correct significant effect, the minimum detectable difference necessary to detect a significant effect, and the minimum number of replicates necessary in order to achieve an acceptable power, all differed between species. Our study has important implications for the design of biodiversity–ecosystem function experiments because the disparity between the contributions that individual species make to ecosystem function demands the use of different levels of replication for each species within an experiment.

KEY WORDS: Biodiversity · Ecosystem function · Species richness · Experimental design · Nutrients · Bioturbation

INTRODUCTION

Following concerns that declining species diversity may impair the long-term functioning of ecosystems, an extensive body of literature has emerged over the last decade (for review, see Hooper et al. 2005) that focuses on the effects of biodiversity loss on key ecological processes. For the most part, and irrespective of habitat type, these studies have adopted an experimental approach (see Schmid et al. 2002, Raffaelli et al. 2003) that involves randomly assembling simple communities around a basal resource and examining the effect, if any, that varying numbers of species have on direct measures of ecosystem functioning, such as primary productivity, decomposition rates and nutrient cycling (see citations in Hooper et al. 2005). As these kinds of experiments are highly controlled, they offer the opportunity to apply rigorous parametric statistical models (most typically regression and ANOVA) in the absence of confounding variables. This approach has been important in understanding the generality of how biological diversity maintains ecological processes, and it has received widespread attention. Findings from microbial (e.g. Bell et al. 2005), terrestrial plant and soil communities (reviewed in Loreau et al. 2001), as well as freshwater and marine invertebrate commu-

*Corresponding author. Email: m.solan@abdn.ac.uk

nities (reviewed in Covich et al. 2004) indicate that, on average, a reduction in biodiversity is likely to impair the short-term provision and long-term sustainability of ecosystem processes (Hooper et al. 2005).

While the ecological research community broadly agrees that biodiversity is causally related to many ecosystem processes (Schläpfer et al. 1999), and fully endorses many experimental conclusions, some uncertainties in experimental design and caveats to interpretation (e.g. heterogeneity of variance with increasing species richness) cast doubts on the wider applicability of some experimental findings (Hooper et al. 2005). Recognition that there have been difficulties with the appropriateness and discriminatory power of alternative experimental designs (Schmid et al. 2002) led to an extensive debate in the primary literature over whether the findings of first generation biodiversity–ecosystem function experiments (hereafter BEF) support the postulate that biodiversity regulates ecosystem processes (for a summary of debate, see Mooney 2002). Some investigators even questioned whether new BEF research was needed (Flint & Kalke 2005) given the availability and efficacy of using historical data (e.g. Emmerson & Huxham 2002, Statzner & Moss 2004). It is important to realise, however, that such criticism does not undermine the validity of the hypotheses in BEF research, nor does it demand an unequivocal demonstration of the precise mechanisms that underpin the broad conclusions of specific BEF experiments. Rather, it draws attention to the inherent difficulty in identifying and attributing causal mechanisms associated with changes in biodiversity, and the need to explicitly identify and acknowledge the limitations of each study (Rosenfeld 2002). These interpretative difficulties arise because the detection of a significant effect of diversity may indicate (1) a true and intrinsic effect of diversity (= complementarity), (2) the effect of one or more 'hidden treatments' (sensu Huston 1997) caused by concurrent changes in species density, composition or identity when the main explanatory variable (species richness) is manipulated, or (3) a combination of both effects at the same time (e.g. Fargione & Tilman 2005).

Distinguishing between the 'properties of complementarity' and the 'mechanisms of selection' (following Petchey 2003) is vitally important because they relate to how biodiversity regulates ecosystem function in the real world. Indeed, in most real communities, the provision of ecosystem processes will depend not only upon the number of species, but also on the relative contribution of dominant and minor species (i.e. species composition and identity effects; Emmerson et al. 2001, Solan et al. 2004), environmental context (Biles et al. 2003), density dependence (Marinelli & Williams 2003) and how species interact with one another.

A 'complementarity effect' occurs when a partitioning of resources (i.e. niche differentiation) leads to increased total resource use, or when some other positive inter-specific interaction (e.g. facilitation) leads to an increase in ecosystem functioning (Cardinale et al. 2002). A 'sampling' or 'selection' effect (Huston 1997) may occur in higher diversity treatments because of the increased probability of including species that have a disproportionate effect on ecosystem functioning (which may in itself be a valid biodiversity effect; Tilman et al. 1997).

Several metrics have been put forward that effectively distinguish selection effects from complementarity effects (Loreau 1998, Emmerson & Raffaelli 2000, Loreau & Hector 2001, Špaèková & Lepš 2001, Hector et al. 2002, Petchey 2003). These generally compare the relative performance of a multi-species mixture with those of monocultures comprising species that made up the original mixture (see Loreau 1998 for general framework). This approach allows the expected yield of each species in the mixture (E_i) to be predicted from its observed yield in monoculture (M_i). Any deviation of the observed total yield (O_T) in the mixture from its expected value (E_T = summation of all E_i) then provides a means to unambiguously assert that a mixture is more productive (= overyielding) than the corresponding monocultures. However, further analysis aiming to determine how species interactions give rise to overyielding is possible only if the relative contribution of each species (O_i) to the total observed yield (O_T) can be subsequently partitioned.

For many marine BEF experiments, where cumulative processes are routinely used as surrogates for ecosystem function (e.g. nutrient concentrations, Emmerson et al. 2001, Biles et al. 2003; bioturbation, Solan et al. 2004; oxygen flux, Waldbusser et al. 2004; grazing, O'Connor & Crowe 2005; photosynthesis, Bruno et al. 2005), direct determination of the relative contribution of each species from the size of the observed effect in a mixture is not always possible, and such metrics are of limited use. Instead, the separation of species identity and density effects from those attributable to species richness can be achieved using a series of alternative ANOVA models (e.g. Schmid et al. 2002, O'Connor & Crowe 2005), although problems can be encountered under some circumstances (see Benedetti-Cecchi 2004). While these and other methods (e.g. power analysis) are available, they are seldom used to cross-check the experimental design or the ensuing findings of BEF experiments.

In this study, using infaunal benthic invertebrate assemblages, we present an experimental design and validation procedure appropriate for distinguishing the compositional effects of diversity from those attributable to species richness. Our aim is to encourage the

use of these techniques in BEF experimentation and, in so doing, explicitly recognise the limitations of such studies within a marine context.

MATERIALS AND METHODS

Faunal and sediment collection. Sediment and 3 infaunal invertebrates, the deposit-feeder *Hediste diversicolor* (Polychaeta), the surficial grazer *Hydrobia ulvae* (Gastropoda) and the suspension-feeder *Cerastoderma edule* (Bivalvia) were collected from mud flats in the Ythan Estuary, Aberdeenshire, Scotland (57° 20.085′ N, 02° 0.206′ W). Sediment was sieved (0.5 mm mesh) in a seawater bath to remove macrofauna and then allowed to settle for 24 h to retain the fine fraction (<63 µm). Excess water was removed and the settled sediment (total organic carbon content, 3.84%; mean particle size = 49.79 µm; volume percentile range: d[v, 0.05] = 4.42 µm, d[v, 0.95] = 309.30 µm) was homogenised to a slurry to facilitate distribution between mesocosms. Sediment and seawater (UV-sterilised, 10 µm prefiltered, salinity 33) were added to each mesocosm 36 h prior to addition of invertebrate species. Seawater was siphoned off and replaced after 24 h to allow the removal of excess nutrients associated with sediment disruption during assembly. All mesocosms were continually aerated. Pre-filtered (Nalgene, 0.45 µm) water samples were taken on the final day of each experiment. Ammonium-nitrogen (NH_4-N), nitrate-nitrogen (NO_X-N) and phosphate-phosphorus (PO_4-P) concentrations were determined with a modular flow injection auto-analyser (FIA Star 5010 series) using an artificial seawater carrier solution.

Mesocosms for species diversity and identity. Replicate (n = 5) macrofaunal communities were assembled in monoculture and in mixtures of 2 and 3 species (40 mesocosms, Table 1) to examine whether more diverse communities have a greater effect on sediment nutrient release (NH_4-N, NO_X-N, PO_4-P) than communities containing fewer species. To minimise hidden treatment effects (sensu Huston 1997) and eliminate pseudo-replication, species richness treatments containing 1 and 2 species were replicated using different species permutations (Table 1). This was not possible for the 3-species mixture because of the limited species pool (n = 3). Biomass was fixed at 2.0 g per mesocosm (equivalent to 255 g m^{-2}), a level consistent with that found at the study site (e.g. Biles et al. 2003). Mesocosms were transparent perspex cores (330 mm high, 100 mm internal diameter) containing 10 cm depth of sediment (equivalent to 785 cm^3) and 20 cm of overlying seawater (equivalent to 2.35 l). These were randomly distributed in an environmental chamber (VC 4100, Vötsch Industrietechnik) and maintained at 14.0 ± 0.1°C with a 12 h light – 12 h dark cycle (2 × 36 W fluorescent tube lights, Arcadia, model FO-30) for 21 d.

Table 1. Species combinations used in the assembled macrofaunal communities for species richness and identity manipulations (n = 5 in all cases). Realised biomass accuracy (mean ± SE): 2.0082 ± 0.0196 g; n = 35). *H diversicolor*: *Hediste diversicolor*; *C. edule*: *Cerastoderma edule*; *H. ulvae*: *Hydrobia ulvae*

Species richness	Biomass (g mesocosm^{-1})		
	H. diversicolor	*C. edule*	*H. ulvae*
0	–	–	–
1	2.00	0	0
1	0	2.00	0
1	0	0	2.00
2	1.0	1.0	0
2	1.0	0	1.0
2	0	1.0	1.0
3	0.67	0.67	0.67

Mesocosms for species density. We assembled 30 additional mesocosms and drew on previous data (Solan & Ford 2003) obtained using the same experimental procedure and sample location to examine the effects of species density on sediment nutrient release (NH_4-N, NO_X-N, PO_4-P). Mesocosms consisted of non-transparent plastic aquaria (210 × 150 × 150 mm) containing 1.0 l of homogenised sediment and 3.0 l of seawater. Macrofaunal biomass was fixed across 5 levels (*Hediste diversicolor* and *Cerastoderma edule*: 0, 0.5, 1.0, 1.5 and 2.0 g) or 6 levels (*Hydrobia ulvae*: 0, 1.0, 3.0, 5.0, 7.0 and 9.0 g; Solan & Ford 2003) and replicated within each biomass treatment (n = 3, n = 3 and n = 5, respectively). Mesocosms were maintained in the dark to prevent microphytobenthic activity. The experiment ran for 15 d at ambient temperature (ca. 8°C).

Data analysis. The relationships between species richness and nutrient concentration, species identity and nutrient concentration, and between species density and nutrient concentration were defined using standard regression and ANOVA procedures followed by post hoc analyses where appropriate. Prior to analysis, graphical exploratory techniques were used to check for outliers and, where appropriate, a data transformation was applied to reduce their effect and also to linearise the relationships. Where outliers were identified, we removed them from the analysis. Where this was necessary (2 occasions), ≥94% of the data remained and a reanalysis with 100% of the data did not alter the conclusions. Following analysis, a model validation was applied to verify that underlying statistical assumptions were not violated; normality was assessed by plotting theoretical quantiles versus standardised residuals (Q-Q plots), homogeneity of variance was evaluated by plotting residuals versus fitted values, and influential datapoints were identified using Cook's distance (Quinn & Keough 2002). Statistical tests were

performed using the software package Brodgar 2.4.8 (www.brodgar.com). Post hoc analyses were performed with the software package SPSS 13.0.1 (SPSS).

Two separate ANOVAs tested for the effects of species richness on nutrient concentration. In the first model, we used a 1-way ANOVA with each of the 3 nutrients (n = 3; NH_4-N, NO_X-N and PO_4-P) treated as separate response variables and species richness (n = 4; 0, 1, 2 and 3) as the nominal explanatory variable. In the second model, we used a 2-way ANOVA where nutrient concentration, irrespective of nutrient identity, was considered as the response variable with species richness (n = 4) and nutrient identity (n = 3) as nominal explanatory variables. For the latter, significant differences ($p < 0.05$) were identified with Tukey's and Sidak's (more conservative) post-hoc comparison tests. Use of the 2 separate ANOVA approaches (multiple response variables treated singularly versus grouped) allows predictions to be made regarding the importance of species richness for both the generation of individual nutrients and for nutrient generation per se.

To identify species identity effects on nutrient generation per se, we applied a 2-way ANOVA with nutrient concentration as the response variable and species identity (n = 8; representing each unique combination in Table 1) and nutrient identity (n = 3) as nominal explanatory variables. We assumed that the behaviour of each species was functionally different and that species effects were not additive (i.e. inter-specific interactions were present; see Emmerson et al. 2001, Biles et al. 2003), such that each unique species combination could be treated as a unique 'species'. Multiple comparisons to identify the source of any significant differences were performed using Tukey's and Sidak's post-hoc tests.

We tested species density effects for each of the 3 species using a 2-way ANOVA with nutrient concentration as the response variable and density (for *Hediste diversicolor* and *Cerastoderma edule*, n = 5; for *Hydrobia ulvae*, n = 6) and nutrient identity (n = 3) as nominal explanatory variables.

Validation of experimental approach. We used a power analysis to test 3 related quantities: (1) the probability of finding a correct significant effect, i.e. the power of the performed analyses; (2) the minimum detectable difference required to detect a significant effect between the mean values for nutrient concentration in the monocultures and those of the multispecies assemblages or, for our density manipulations, between the lowest and largest mean values of nutrient concentration across biomass treatments; and (3) the minimum number of replicates necessary in order to achieve an acceptable power (>80%) given our experimental design. We considered ecosystem function to be the sum of all 3 investigated nutrients because the minimum replication required for an experiment in which 3 responses are being tested simultaneously is likely to be greater than that predicted from the behaviour of only 1 nutrient (i.e. we sought a conservative recommendation). As ecosystem performance may depend on the presence of specific functional groups and/or certain faunal traits more than it does on species richness per se (Solan et al. 2004), we reapplied the 2-way ANOVA, with species identity and nutrient identity as nominal explanatory variables, in the absence of the controls (no macrofauna). This yields a more appropriate fit of residuals versus predicted values and prevents species presence (non-controls) versus absence (controls) effects from distorting the analysis. We determined power function and sample size using the statistical graphs of Pearson & Hartly (1951) at a significance level of $\alpha = 0.05$. All power analyses were performed using an add-on software package for Brodgar 2.4.8.

RESULTS

Species richness effects

The concentration of NH_4-N and PO_4-P increased with species richness, while the concentration of NO_X-N decreased with species richness levels (Fig. 1). ANOVA confirmed significant effects of species richness on $\log_{10}$ transformed nutrient concentration for NH_4-N ($F = 3.65$, df = 3, $p < 0.05$) and PO_4-P ($F = 5.44$, df = 3, $p < 0.01$), but not for NO_X-N ($F = 2.12$, df = 3, $p = 0.12$). For the 2-way ANOVA of nutrient concentration ($\log_{10}$ transformed), where nutrient identity was treated as an additional explanatory variable to species richness, we found significant effects of species richness ($F = 7.00$, df = 3, $p < 0.001$), nutrient identity ($F = 192.39$, df = 2, $p < 0.001$) and the interaction of species richness × nutrient identity ($F = 4.43$, df = 6, $p < 0.001$).

Post hoc analyses of the 1-way ANOVA for NH_4-N revealed that the controls were not significantly different from the monocultures (Tukey's, $p = 0.57$; Sidak's, $p = 0.74$) but there were marginal differences between the controls and the multiple species richness levels (Tukey's, $p \approx 0.05$; Sidak's, $p \approx 0.07$). The highest species richness treatment (= 3) was not significantly different to the intermediate (1 and 2) species richness levels (Tukey's, $p \geq 0.22$; Sidak's, $p \geq 0.30$). For PO_4-P, controls were not significantly different from the monocultures (Tukey's, $p = 0.34$; Sidak's, $p = 0.46$) but there were significant differences between the controls and the multiple species richness levels (Tukey's and Sidak's, $p \leq 0.01$). The highest species richness treatment (= 3) was not significantly different to the intermediate species richness levels (Tukey's, $p \geq 0.12$; Sidak's, $p \geq 0.16$).

Post hoc analyses for the 2-way ANOVA confirmed that nutrient concentrations were dependent on nutri-

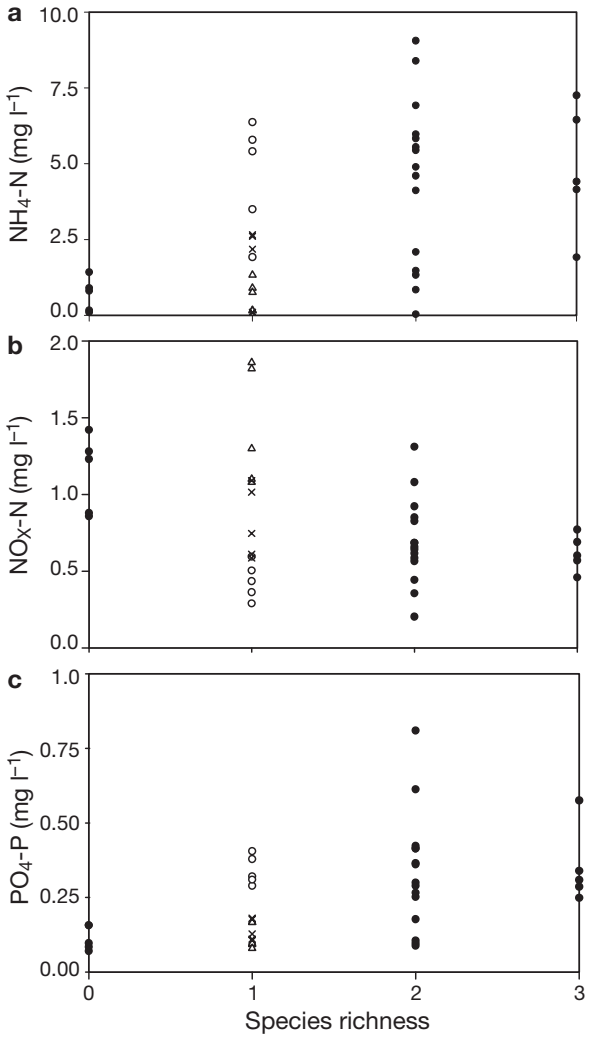

Fig. 1. Replicated (n = 5) accumulated nutrient concentrations of (a) NH_4-N (b) NO_X -N and (c) PO_4-P with increasing species richness. Species identity in the single species treatments are denoted by ○: *Hediste diversicolor*; △: *Cerastoderma edule*; x: *Hydrobia ulvae*

ent identity (Tukey's and Sidak's, $p < 0.001$ for every nutrient identity comparison), and that there were significant species richness effects irrespective of nutrient identity. Although the controls were not significantly different from the monocultures (Tukey's, $p = 0.28$; Sidak's, $p = 0.38$), the 2-species combinations were significantly different to the controls (Tukey's and Sidak's, $p < 0.01$) and the monocultures (Tukey's and Sidak's, $p < 0.05$), but not from the 3-species combinations (Tukey's and Sidak's, $p \geq 0.89$); the 3-species combinations were significantly different from the controls (Tukey's and Sidak's, $p < 0.01$) and monocultures (Tukey's and Sidak's, $p \leq 0.05$), but not from the 2-species combinations (Tukey's and Sidak's, $p \geq 0.87$).

Species identity effects

Our data reveal clear differences in the relative contribution of individual species to nutrient generation, irrespective of nutrient identity (*Hediste diversicolor* > *Hydrobia ulvae* > *Cerastoderma edule*; Fig. 1). Two-way ANOVA of nutrient concentration ($\log_{10}$ transformed), with species identity and nutrient identity treated as explanatory variables, confirmed significant effects of species identity ($F = 12.39$, df = 7, $p < 0.001$), nutrient identity ($F = 335.43$, df = 2, $p < 0.001$) and a significant interaction of species identity × nutrient identity ($F = 13.55$, df = 14, $p < 0.001$). Post hoc analyses on all treatments revealed that nutrient concentrations in monocultures of *H. diversicolor* were marginally greater than those in *H. ulvae* (Tukey's, $p = 0.05$; Sidak's, $p = 0.07$) and significantly greater than those containing monocultures of *C. edule* (Tukey's and Sidak's, $p < 0.001$). When comparisons were made between all species identity levels, post hoc analysis revealed that the presence of *H. diversicolor* in a multispecies combination, irrespective of the composition of the remaining species, led to a significant increase in nutrient concentration relative to those combinations where *H. diversicolor* was absent (Tukey's and Sidak's, $p \leq 0.05$; dark grey shaded entries, Table 2). When species combinations that contained *H. diversicolor* are compared to other multi-species combinations that also contained *H. diversicolor*, irrespective of proportional representation, differences in nutrient concentrations are non-significant (Tukey's and Sidak's, $p > 0.05$; non-shaded areas in HD column, Table 2). For treatments that contained *H. ulvae*, *C. edule* or *H. ulvae* and *C. edule*, nutrient concentrations

Table 2. Tukey's pairwise comparisons of species identity for controls (CNTRL) and all species combinations in the species richness manipulations; HD: *Hediste diversicolor*; CE: *Cerastoderma edule*; HU: *Hydrobia ulvae*. Light grey: HD absent in both pairwise combinations; dark grey: HD in one of the pairwise combinations. ***p < 0.001, **p < 0.01, *p < 0.05, ns = not significant. Sidak's pairwise comparisons had identical results

	CNTRL	HD	CE	HU	CEHU	CEHD	HDHU	HDCEHU
CNTRL	–							
HD	**	–						
CE	ns	***	–					
HU	ns	*	ns	–				
CEHU	ns	ns	ns	ns	–			
CEHD	***	ns	***	**	**	–		
HDHU	***	ns	***	***	**	ns	–	
HDCEHU	***	ns	***	**	*	ns	ns	–

were not significantly different from those generated in the controls or any other treatment where *H. diversicolor* was absent (Tukey's and Sidak's, $p > 0.05$; light grey shaded areas, Table 2). Thus, mesocosms that contained *H. diversicolor* had high mean NH_4-N and PO_4-P concentrations, and low NO_X-N concentrations.

Examination of our biomass treatments confirmed that inclusion of biomass as a covariate, or as an independent explanatory variable, was unnecessary in any of our models examining species richness or species identity effects, because the realised (mean ± SE = 2.0082 ± 0.0196 g, n = 35) and target (2.0 g) biomass were sufficiently similar (CV < 0.01 g, n = 35).

Species density effects

Two-way ANOVA of nutrient concentration, with species density and nutrient identity treated as nominal explanatory variables, revealed significant effects of density (Fig. 2) relative to the controls for *Hediste diversicolor* ($F = 5.20$, df = 4, $p < 0.01$), but not for *Cerastoderma edule* ($F = 1.12$, df = 4, $p = 0.37$) or *Hydrobia ulvae* ($F = 1.86$, df = 5, $p = 0.11$). The same analysis revealed significant effects for nutrient identity ($p < 0.001$) and density × nutrient identity interactions ($p < 0.01$) for all 3 species. These results are consistent with the findings of the single species mesocosms in the species richness manipulations, although the magnitude of response is comparatively higher for *H. ulvae* (reflecting seasonal differences in sediment nutrient content).

Validation of effects

Application of a power analysis to the species richness and species identity data revealed that the minimal difference required for the detection of a significant species richness effect for nutrient concentration per se was 1.79 mg l^{-1}. By comparing the nutrient concentrations for individual species in monocultures with that observed in the 3-species combinations, results showed that both *Hydrobia ulvae* and *Cerastoderma edule* exhibited sufficiently large between-mean concentration values (1.85 mg l^{-1} and 2.30 mg l^{-1}, respectively) in order to achieve a power ≥80% at the level of replication used in our study (minimum replication required, n = 5 and n = 3, respectively; Fig. 3). Thus, for a given density, nutrient concentrations were higher in a 3-species combination than would be expected for corresponding monocultures containing only *H. ulvae* or *C. edule* (i.e. evidence of a significant diversity effect). For *Hediste diversicolor*, however, the same comparison revealed an insufficiently large between-mean concentration value (1.14 mg l^{-1}) to detect a significant diversity effect. When treatments contain *H. diversicolor*, >100 replicates at each density level are necessary to detect a significant effect of diversity over that of species identity at a desirable power (Fig. 3).

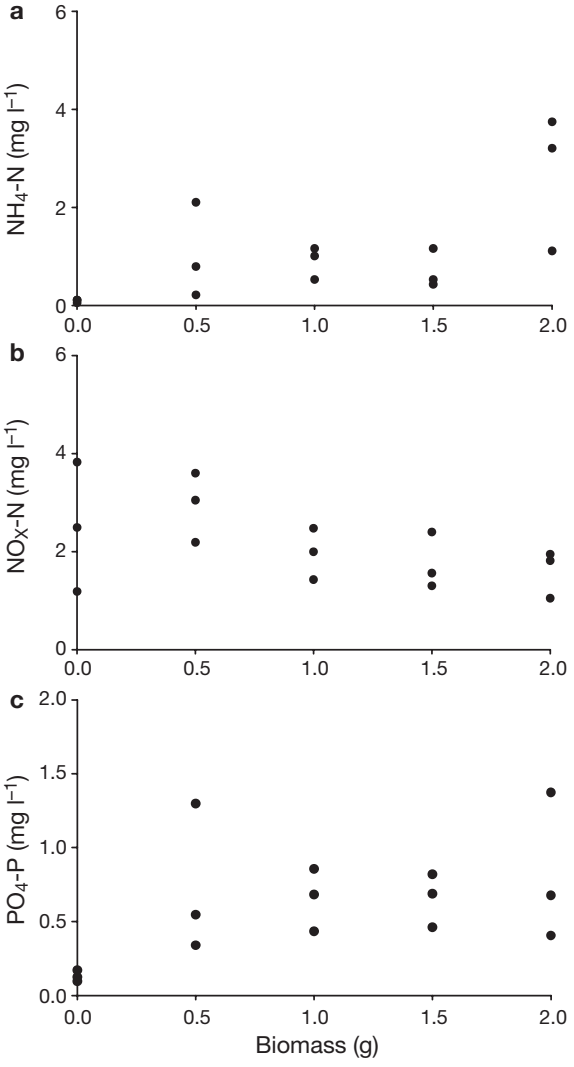

Fig. 2. Relationship between species biomass and (a) NH_4-N, (b) NO_X-N and (c) PO_4-P nutrient concentration for *Hediste diversicolor*

When applied to the species density data, power analysis revealed that our design (replication n = 3 or n = 5 per density level) was not sufficient in all cases to detect true density effects for the biomass levels used in our experiments. The minimum number of replicates per density interval required to detect a density effect with sufficient power ($1 - \beta \geq 80\%$) was ≥3 for *Hediste diversicolor* (Fig. 4a), ≥9 for *Hydrobia ulvae* (Fig. 4b) and ≥80 for *Cerastoderma edule* (Fig. 4c). Thus, only for *H. diversicolor* had we achieved the appropriate level of replication necessary to detect density effects in our experiments.

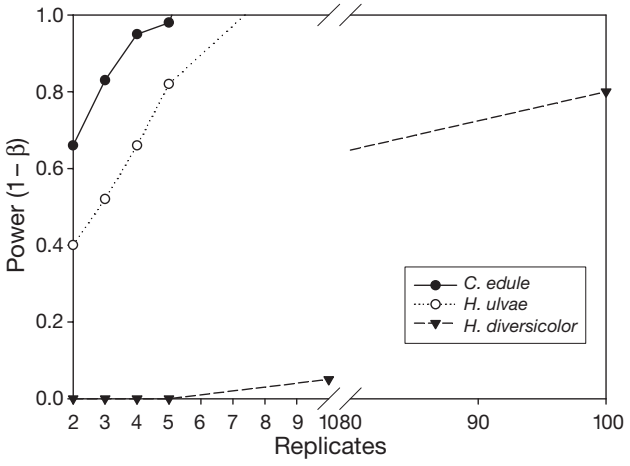

Fig. 3. Sample size (replicates) required to distinguish species richness from species identity effects for *Cerastoderma edule*, *Hydrobia ulvae* and *Hediste diversicolor* at α = 0.05 for the present experimental set-up. The minimum level of replication occurs when power (1– β) ≥ 80%. Large values of replication when 1– β = 80% indicate strong effects of individual species

DISCUSSION

Our study identifies clear effects of increased infaunal species diversity on nutrient generation in the marine benthos, although significant species identity and density effects underpin the observed response. These findings support those of previous studies on marine benthic invertebrates where consistent diversity effects have been shown to develop from the idiosyncratic contributions of individual species (Emmerson et al. 2001, Biles et al. 2003, Waldbusser et al. 2004). Several studies have reported the relative importance of complementarity by using comparisons between the performances of species combinations and monocultures (Loreau 1998, Emmerson & Raffaelli 2000, Loreau & Hector 2001, Špaèková & Lepš 2001, Hector et al. 2002, Petchey 2003), though it has always been a difficult task to select the minimum subset of complementary species within the regional species pool in order to explain diversity effects. In the present study, nutrient concentrations varied between diversity treatments and undoubtedly correlate with the role that the component species play within the ecosystem. The bioturbation activities of the deposit-feeding polychaete *Hediste diversicolor* were of particular importance to nutrient generation, more so than either *Cerastoderma edule* (suspension feeder) or *Hydrobia ulvae* (surficial grazer), whose bioturbatory activities were masked when *H. diversicolor* was present (consistent with Mermillod-Blondin 2005). The decreasing patterns of NO_X-N with increasing NH_4-N observed here indicate that macrofaunal bioturbation stimulated microbial denitrification, leading to a reduction of the internal pool of dissolved and particulate nutrients in the sediment and the associated increase in nutrients in the water column (Hansen & Kristensen 1997).

The detection of strong species identity effects give credence to the view that functional diversity (groups of ecologically equivalent species) may be equally, if not more, important for the delivery of ecosystem services than species diversity (Solan et al. 2004). For certain ecosystem processes, some species within an assemblage will therefore be functionally less important than other species (in line with insurance; Yachi & Loreau 1999), albeit for a given set of circumstances in time and/or space (i.e. context, e.g. Biles et al. 2003). This property of biological communities has important implications for the design of BEF experiments. If we are to distinguish identity effects from diversity effects, the disparity between the contributions that individual species make to ecosystem function demands the use of different levels of replication for each species within an experiment. Where species have a particularly strong and disproportionate effect on ecosystem func-

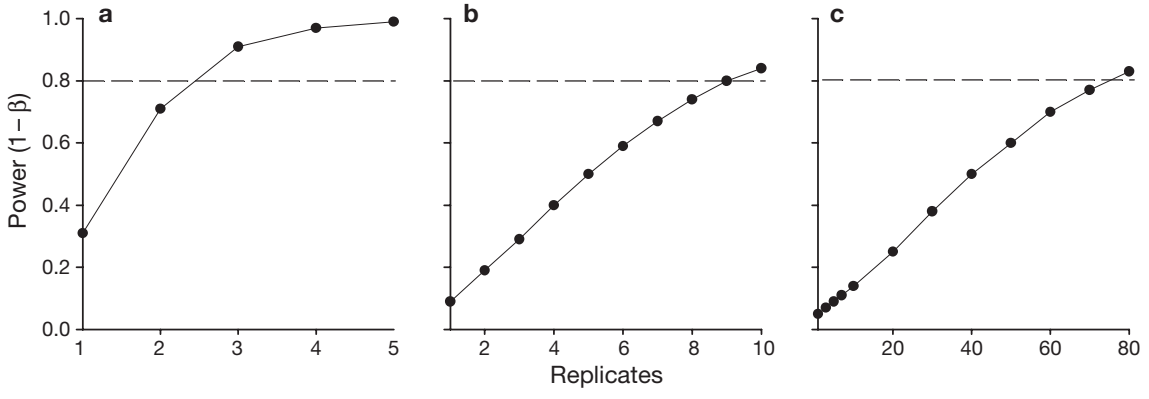

Fig. 4. Estimation of sample size (replicates) required to detect a significant effect of species density between the control and highest density level at a significance level of α = 0.05 for (a) *Hediste diversicolor*, (b) *Hydrobia ulvae* and (c) *Cerastoderma edule*. The minimum level of replication occurs when power (1 – β) ≥ 80% (indicated by the dashed line). Large values of replication when 1– β = 80% indicate weak density effects of individual species

tion, such as *Hediste diversicolor*, a high number of replicates are required to statistically rule out the possibility that the observed response is a true diversity effect (complementarity) and not a product of species identity (sampling or selection). The converse is true when species have a comparatively less marked contribution to ecosystem function, such as *Cerastoderma edule* and *Hydrobia ulvae*. Thus, both types of factorial balanced designs used in the present study were sufficient to identify a diversity effect, but they were unable to eliminate the selection effect for treatments that included *H. diversicolor*.

A further complication is that a similar problem arises with the choice of density within each mesocosm (Marinelli & Williams 2003). Although we independently found evidence for a significant effect of species density, the level of replication per density level to detect a true density effect is markedly different between species. These differences arise because species-specific traits mediate the functional contribution of each species, such that mesocosms containing species with strong per capita effects (e.g. *Hediste diversicolor*) require fewer replicates per density level to detect a true density effect from that of a presence versus absence effect. That said, there is some evidence that adopting a larger biomass interval, rather than increasing the minimum replication required, might solve the problem of insufficient discriminatory power between density treatments.

An appropriate solution may be a factorial balanced design in which species richness and biomass are manipulated in concert (Benedetti-Cecchi 2004) at appropriate replication levels that account for density and identity effects of the species under study. Adoption of such a design may also relieve some of the constraints imposed by inequality of variance with increasing levels of species richness, although when no significant differences among treatments are found in balanced samples, heterogeneity of variance is of no concern (p. 194, Underwood 1997). Nevertheless, whichever design is ultimately chosen it is clear that the density, identity and number of species richness levels within an experiment need to be chosen with care because they greatly influence the scope of inference permitted. This point is not trivial, because it raises concerns about experimentally manipulating unknown or uncharacterised species, whose contribution to ecosystem function is either more difficult to determine or unknown. Use of species that are functionally equivalent (i.e. selectively choosing species from within a functional group to standardise the comparative weighting of per capita effects between species) in BEF experiments would be unrepresentative of real biological communities and could misrepresent the relative importance of contrasting functional groups. We contend that, in naturally occurring communities, the individual contribution of species to ecosystem function is a complex product of niche complementarity, species density, sampling and selection effects, although the relative contribution of each of these mechanisms most likely alters in time and space according to context (e.g. Biles et al. 2003). If we are to fully understand the causal relationship between biodiversity and the delivery of ecosystem processes, future experimental analyses will need to adopt a more holistic approach aimed at distinguishing the relative contribution of multiple mechanisms to ecosystem functioning.

Acknowledgements. We thank D. MacKinnon of Plant and Soil Sciences, University of Aberdeen for analysis of water samples. O. McPherson and T. Doherty-Bone assisted with fieldwork. The support and co-operation received from Forvie National Nature Reserve, Scottish Natural Heritage is gratefully acknowledged. All work was supported by NERC.

LITERATURE CITED

Bell T, Newman JA, Silverman BW, Turner SL, Lilley AK (2005) The contribution of species richness and composition to bacterial services. Nature 436:1157–1160

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

Biles C, Solan M, Isaksson I, Paterson DM, Emes C, Raffaelli DG (2003) Flow modifies the effect of biodiversity on ecosystem functioning: an in situ study of estuarine sediments. J Exp Mar Biol Ecol 285/286:165–177

Bruno JF, Boyer KE, Duffy JE, Lee SC, Kertesz JS (2005) Effects of macroalgal species identity and richness on primary production in benthic marine communities. Ecol Lett 8:1165–1174

Cardinale BJ, Palmer MA, Collins SL (2002) Species diversity enhances ecosystem functioning through interspecific facilitation. Nature 415:426–429

Covich AP, Austen M, Bärlocher F, Chauvet E and 7 others (2004) The role of biodiversity in the functioning of freshwater and marine benthic ecosystems. BioScience 54(8): 767–775

Emmerson M, Huxham M (2002) How can marine ecology contribute to the biodiversity-ecosystem functioning debate? In: Loreau M, Naeem S, Inchausti P (eds) Biodiversity and ecosystem functioning: synthesis and perspectives. Oxford University Press, Oxford, p 139–146

Emmerson MC, Raffaelli DG (2000) Detecing the effects of diversity on measures of ecosystem function: experimental design, null models and empirical observations. Oikos 91(1):195–203

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

Fargione JE, Tilman D (2005) Diversity decreases invasion via both sampling and complementarity effects. Ecol Lett 8:604–611

Flint RW, Kalke RD (2005) Reinventing the wheel in ecology research? Science 307:1875

Hansen K, Kristensen E (1997) Impact of macrofaunal recolo-

nization on benthic metabolism and nutrient fluxes in a shallow marine sediment previously overgrown with macroalgal mats. Estuar Coast Shelf Sci 45(5):613–628

Hector A, Bazeley-White E, Loreau M, Otway S, Schmid B (2002) Overyielding in grassland communities: testing the sampling effect hypothesis with replicated biodiversity experiments. Ecol Lett 5(4):502–511

Hooper DU, Chapin FS, Ewel JJ, Hector A and 11 others (2005) Effects of biodiversity on ecosystem functioning: a consensus of current knowledge. Ecol Monogr 75(1):3–35

Huston MA (1997) Hidden treatments in ecological experiments: re-evaluating the ecosystem function of biodiversity. Oecologia 110:449–460

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

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

Loreau M, Naeem S, Inchausti P, Bengtsson J and 8 others (2001) Biodiversity and ecosystem functioning: current knowledge and future challenges. Science 294:804–808

Marinelli RL, Williams TJ (2003) Evidence for density-dependent effects of infauna on sediment biogeochemistry and benthic-pelagic coupling in nearshore systems. Estuar Coast Shelf Sci 57:179–192

Mermillod-Blondin F, François-Carcaillet F, Rosenberg R (2005) Biodiversity of benthic invertebrates and organic matter processing in shallow marine sediments: an experimental study. J Exp Mar Biol Ecol 315:187–209

Mooney HA (2002) The debate on the role of biodiversity in ecosystem functioning. In: Loreau M, Naeem S, Inchausti P (eds) Biodiversity and ecosystem functioning: synthesis and perspectives. Oxford University Press, Oxford, p 12–17

O'Connor NE, Crowe TP (2005) Biodiversity loss and ecosystem functioning: distinguishing between number and identity of species. Ecology 86(7):1783–1796

Pearson ES, Hartly HO (1951) Charts for the power function for analysis of variance tests, derived from the non-central F-distribution. Biometrika 38:112–130

Petchey O (2003) Integrating methods that investigate how complementarity influences ecosystem functioning. Oikos 101:323–330

Quinn GP, Keough MJ (2002) Experimental design and data analysis for biologists. Cambridge University Press, Cambridge

Raffaelli D, Emmerson M, Solan M, Biles C, Paterson D (2003) Biodiversity and ecosystem functioning in shallow coastal waters: an experimental approach. J Sea Res 49(2): 133–141

Rosenfeld JS (2002) Functional redundancy in ecology and conservation. Oikos 98(1):156–162

Schläpfer F, Schmid B, Seidl I (1999) Expert estimates about effects of biodiversity on ecosystem processes and services. Oikos 84(2):346–352

Schmid B, Hector A, Huston MA, Inchausti P, Nijs I, Leadley PW, Tilman D (2002) The design and analysis of biodiversity experiments. In: Loreau M, Naeem S, Inchausti P (eds) Biodiversity and ecosystem functioning: synthesis and perspectives. Oxford University Press, Oxford, p 61–75

Solan M, Ford R (2003) Macroalgal-induced changes on the distribution, composition and diversity of benthic macrofauna: implications for ecosystem functioning. In: Raffaelli D, Solan M, Paterson D, Buck AL, Pomfret JR (eds) The estuaries and coasts of north-east Scotland. Coastal zone topics 5. Proc ECSA local meet, Aberdeen, 2001. Estuarine and Coastal Sciences Association, p 77–87

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

Špaèková I, Lepš J (2001) Procedure for separating the selection effect from other effects in diversity-productivity relationship. Ecol Lett 4:585–594

Statzner B, Moss B (2004) Linking ecological function, biodiversity and habitat: a mini-review focusing on older ecological literature. Basic Appl Ecol 5:97–106

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

Underwood AJ (1997) Experiments in ecology: their logical design and interpretation using analysis of variance. Cambridge University Press, Cambridge

Waldbusser GG, Marinelli RL, Whitlatch RB, Visscher PT (2004) The effects of infaunal biodiversity on biogeochemistry of coastal marine sediments. Limnol Oceanogr 49(5):1482–1492

Yachi S, Loreau M (1999) Biodiversity and ecosystem productivity in a fluctuating environment: the insurance hypothesis. Proc Nat Acad Sci USA 96:1463–1468

Editorial responsibility: David Raffaelli (Guest Editor), Heslington, UK

Submitted: April 12, 2005; Accepted: October 21, 2005
Proofs received from author(s): March 13, 2006