

## Designing forest biodiversity experiments: general considerations illustrated by a new large experiment in subtropical China

Helge Bruelheide<sup>1,2\*</sup>, Karin Nadrowski<sup>3</sup>, Thorsten Assmann<sup>4</sup>, Jürgen Bauhus<sup>5</sup>, Sabine Both<sup>1</sup>, François Buscot<sup>2,6</sup>, Xiao-Yong Chen<sup>7</sup>, Bingyang Ding<sup>8</sup>, Walter Durka<sup>2,6</sup>, Alexandra Erfmeier<sup>1,2</sup>, Jessica L. M. Gutknecht<sup>6</sup>, Dali Guo<sup>9</sup>, Liang-Dong Guo<sup>10</sup>, Werner Härdtle<sup>4</sup>, Jin-Sheng He<sup>11</sup>, Alexandra-Maria Klein<sup>4,5</sup>, Peter Kühn<sup>12</sup>, Yu Liang<sup>13</sup>, Xiaojuan Liu<sup>13</sup>, Stefan Michalski<sup>6</sup>, Pascal A. Niklaus<sup>14</sup>, Kequan Pei<sup>13</sup>, Michael Scherer-Lorenzen<sup>5</sup>, Thomas Scholten<sup>12</sup>, Andreas Schuldt<sup>4</sup>, Gunnar Seidler<sup>1</sup>, Stefan Trogisch<sup>1,2</sup>, Goddert von Oheimb<sup>4</sup>, Erik Welk<sup>1,2</sup>, Christian Wirth<sup>2,3</sup>, Tesfaye Wubet<sup>2,6</sup>, Xuefei Yang<sup>15</sup>, Mingjian Yu<sup>16</sup>, Shouren Zhang<sup>13</sup>, Hongzhang Zhou<sup>17</sup>, Markus Fischer<sup>18</sup>, Keping Ma<sup>13</sup> and Bernhard Schmid<sup>14</sup>

<sup>1</sup>Martin Luther University Halle Wittenberg, Am Kirchtor 1, 06108 Halle (Saale), Germany; <sup>2</sup>iDiv – German Centre for Integrative Biodiversity Research, Deutscher Platz 5e, D-04103 Leipzig, Germany; <sup>3</sup>University of Leipzig, Johannisallee 21–23, D-04103 Leipzig, Germany; <sup>4</sup>Leuphana University of Lüneburg, Scharnhorststr. 1, D-21332 Lüneburg, Germany; <sup>5</sup>University of Freiburg, Tennenbacherstr. 4, D-79106 Freiburg, Germany; <sup>6</sup>Helmholtz Centre for Environmental Research UFZ, Theodor-Lieser-Str. 4, D-06120 Halle (Saale), Germany; <sup>7</sup>East China Normal University Shanghai, 3663 Zhongshan North Road, Shanghai, 200062, China; <sup>8</sup>Wenzhou University, Zhejiang, China; <sup>9</sup>Institute of Geographic Sciences and Natural Resources Research, CAS, 11A Datun Road, Beijing, 100101, China; <sup>10</sup>Institute of Microbiology, CAS, 3-1 Beichen Western Road, Beijing, 100101, China; <sup>11</sup>Peking University, Beijing, 100871, China; <sup>12</sup>University of Tübingen, Rümelinstraße 19-23, D-72074 Tübingen, Germany; <sup>13</sup>Institute of Botany, CAS, 20 Nanxincun, Xiangshan, Beijing, 100093, China; <sup>14</sup>University of Zurich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland; <sup>15</sup>Kunming Institute of Botany, CAS, 132 Lanhei Road, Kunming, 650204, China; <sup>16</sup>Zhejiang University, Hangzhou, 310058, China; <sup>17</sup>Institute of Zoology, CAS, 1 Beichen West Road, Beijing, 100101, China; and <sup>18</sup>University of Bern, Altenbergrain 21, CH-3013 Bern, Switzerland

### Summary

**1.** Biodiversity–ecosystem functioning (BEF) experiments address ecosystem-level consequences of species loss by comparing communities of high species richness with communities from which species have been gradually eliminated. BEF experiments originally started with microcosms in the laboratory and with grassland ecosystems. A new frontier in experimental BEF research is manipulating tree diversity in forest ecosystems, compelling researchers to think big and comprehensively.

**2.** We present and discuss some of the major issues to be considered in the design of BEF experiments with trees and illustrate these with a new forest biodiversity experiment established in subtropical China (Xingangshan, Jiangxi Province) in 2009/2010. Using a pool of 40 tree species, extinction scenarios were simulated with tree richness levels of 1, 2, 4, 8 and 16 species on a total of 566 plots of 25.8 × 25.8 m each.

**3.** The goal of this experiment is to estimate effects of tree and shrub species richness on carbon storage and soil erosion; therefore, the experiment was established on sloped terrain. The following important design choices were made: (i) establishing many small rather than fewer larger plots, (ii) using high planting density and random mixing of species rather than lower planting density and patchwise mixing of species, (iii) establishing a map of the initial ‘ecoscape’ to characterize site heterogeneity before the onset of biodiversity effects and (iv) manipulating tree species richness not only in random but also in trait-oriented extinction scenarios.

**4.** Data management and analysis are particularly challenging in BEF experiments with their hierarchical designs nesting individuals within-species populations within plots within-species compositions. Statistical analysis best proceeds by partitioning these random terms into fixed-term contrasts, for example, species composition into contrasts for species richness and the presence of particular functional groups, which can then be tested against the remaining random variation among compositions.

**5.** We conclude that forest BEF experiments provide exciting and timely research options. They especially require careful thinking to allow multiple disciplines to measure and analyse data jointly and effectively. Achiev-

\*Correspondence author. E-mail: helge.bruehlheide@botanik.uni-halle.de

ing specific research goals and synergy with previous experiments involves trade-offs between different designs and requires manifold design decisions.

**Key-words:** BEF-China, ecoscape, genetic diversity, planting pattern, random partitions design, species richness, trait-oriented extinction sequence

## Introduction

Positive effects of biodiversity on the functioning of ecosystems have been observed in numerous experiments (Loreau, Naeem & Inchausti 2002; Hooper *et al.* 2005; Balvanera *et al.* 2006; Worm *et al.* 2006; Duffy 2009). Starting from the first experiments in climate chambers (Naeem *et al.* 1994) and on grassland field plots (Leadley & Körner 1996; Tilman, Wedin & Knops 1996), experiments have become more and more sophisticated, often in constructive response to criticism (e.g. Grime 1997; Huston 1997; Schmid *et al.* 2002). Most biodiversity–ecosystem functioning (BEF) experiments have employed small model systems with fast-growing primary producers, in particular herbaceous plants (for reviews see Loreau, Naeem & Inchausti 2002; Hooper *et al.* 2005; Scherer-Lorenzen *et al.* 2007; Cardinale *et al.* 2011). However, considering the large contribution of forests to ecosystem services such as carbon storage, climate regulation, water filtration or erosion control at the global scale (Durieux, Machado & Laurent 2003; Bala *et al.* 2007; Quijas *et al.* 2012), it is important to test whether the results obtained for the simpler systems of smaller and short-lived organisms can be extrapolated to forest ecosystems harbouring the largest and longest lived plant species on land.

Recent studies and meta-analyses using forest plots from sample surveys in established forests indeed found significant correlations between tree species richness and ecosystem properties (e.g. standing biomass and associated diversity of faunistic groups) and processes (e.g. litter decomposition, herbivory, productivity; Gamfeldt *et al.* 2013; Scherer-Lorenzen 2013; Chisholm *et al.* 2013). Despite a potential publication bias, tree diversity might thus indeed play a critical role for ecosystem functioning. However, as with all observational studies, it is not clear whether these correlations reflect causal relationships, in which direction causality works, or whether additional ‘third’ variables are involved, for example, stand age or tree density (Marquard *et al.* 2009a). In recent years, a number of new BEF experiments have, therefore, been initiated with plots deliberately planted with different tree species richness and composition (Scherer-Lorenzen *et al.* 2005a; Nadrowski, Wirth & Scherer-Lorenzen 2010; Verheyen & Scherer-Lorenzen 2012). Here, we discuss the major issues that we encountered when designing a new forest BEF experiment in subtropical China (referred to as BEF-China). We draw from experience with previous experiments and explain our own design choices to provide a case-study example. While we are not striving for a complete review of forest BEF experiments, we aim to provide guidelines to assist others in designing and establishing further forest BEF experiments.

## MAJOR QUESTIONS ADDRESSED BY BEF EXPERIMENTS

Major questions addressed by BEF experiments so far have been whether random species loss, mostly of plants, can negatively affect ecosystem functioning, in particular primary productivity, nutrient cycling and the diversity and abundance of other trophic groups (Balvanera *et al.* 2006). This has most often been addressed by assembling experimental communities with different species numbers. In other words, the scenario simulates the random extinction of species from a local species pool. Extinction is simulated by leaving species out when sowing or planting the experimental communities, usually at constant total densities of individuals per community. From these major BEF questions, numerous additional questions can be derived, which will influence the particular choice of the experimental design (Schmid *et al.* 2002).

These additional questions include mechanisms through which biodiversity may enhance ecosystem functioning, in particular whether average performance across all species is higher at higher diversity (the complementarity effect, Loreau & Hector 2001) or whether diverse mixtures have a higher chance to contain high-performing species (sampling or selection effect, Huston 1997; Loreau & Hector 2001). Examining this question requires growing all species in monocultures, an aspect not considered in early biodiversity experiments. When all species are available in monoculture, the additive partitioning method of Loreau & Hector (2001) provides a statistical method to separate complementarity and selection effects, but further design and measurement aspects need to be considered to analyse the underlying mechanisms of these statistical effects, for example, resource partitioning between species (for example see von Felten *et al.* 2009).

Other additional questions concern the particular aspect of biodiversity causing an effect, including whether it is species richness *per se* or the functional diversity of a community that matters (Hooper *et al.* 2005). The latter would be expected if biodiversity effects are related to functional differences among species, for example, regarding resource uptake, as has been shown for the combination of legumes and grasses in grassland systems (Spehn *et al.* 2002). To disentangle species richness and functional diversity effects, special designs can be envisaged that vary functional diversity within species richness levels and *vice versa*, yet this is difficult to achieve in a balanced way (see e.g. Le Roux *et al.* 2013; Table A1). Furthermore, differences among plant species in attracting pests and pathogens may cause positive biodiversity effects (Petermann *et al.* 2008; Schnitzer *et al.* 2011). In the German BIOTREE experiment, Hantsch *et al.* (2013) demonstrated a negative effect of tree richness on the pathogen load of common powdery mildew species on *Quercus petraea*. Finally, a rarely tested question is

whether genetic diversity within species affects ecosystem functioning (e.g. Crutsinger *et al.* 2006; Moreira & Mooney 2013).

Further questions address the type of species loss simulated in biodiversity experiments. Whereas initial BEF experiments (Naeem *et al.* 1994; Leadley & Körner 1996) simulated random extinction scenarios by nesting all species compositions of lower diversity within compositions of the next higher diversity, most new experiments do not nest compositions anymore [but see Bell *et al.* (2005) with a new random partitions design (Bell *et al.* 2009)]. This unfortunately reduces the power to disentangle effects of diversity and composition. In any case, random extinction scenarios are the best choice in the absence of information about drivers of extinction, or if drivers do not lead to biased extinction of species with particular contributions to ecosystem functioning. However, if such biased extinction does occur, appropriate nonrandom extinction scenarios should be used (see e.g. Schlöpfer, Pfisterer & Schmid 2005). These may reflect the preferential loss of species with particular traits and thus particular contributions to ecosystem functioning. As a cautionary note, it should be mentioned that when sufficient information about extinction drivers is lacking or several, perhaps diffuse, extinction drivers interact with each other, nonrandom extinction scenarios focusing on a wrong or a single major cause of extinction may be even more unrealistic than random extinction scenarios.

Finally, there are questions concerning the context in which BEF experiments are conducted. Biodiversity effects may differ between ecosystem types such as aquatic vs. terrestrial (Cardinale *et al.* 2006) or grasslands vs. forests (for which we address design questions here), between homogeneous and heterogeneous environmental conditions (Wacker *et al.* 2008), between low- and high-nutrient environments (Weigelt *et al.* 2009), and between different ecosystem functions such as productivity vs. erosion control. Different contexts may require additional treatments or measurements in BEF experiments. In particular, it is increasingly recognized that biodiversity effects may be most relevant when multiple ecosystem functions are considered (e.g. Hector & Bagchi 2007; Zavaleta *et al.* 2010; Isbell *et al.* 2011; Pasari *et al.* 2013). This finding might be explained by trade-offs between ecosystem functions that may be more severe in low- than in high-diversity ecosystems.

#### USING FOREST EXPERIMENTS IN BEF RESEARCH

Forests are globally important ecosystems because of their wide geographical cover and their prominent standing biomass, thus provisioning unique ecosystem goods, such as timber, food, fuel or medicinal plants, and delivering services for humans, such as carbon sequestration, soil erosion control, water retention and purification, nutrient provision, local climate regulation, global climate change mitigation and cultural services (Quijas *et al.* 2012). Understanding the role of biodiversity for forest ecosystems renders forest BEF experiments highly relevant.

Despite a number of disadvantages compared with smaller, short-lived organisms such as microbes and annual or perennial plants (the most commonly used test organisms in BEF

experiments), trees also have some advantages when used in BEF experiments. First, and as already discussed in detail by Scherer-Lorenzen, Körner & Schulze (2005b) and Scherer-Lorenzen *et al.* (2007), tree diversity experiments allow for working at the level of single individuals to quantify mechanisms underlying BEF relationships, such as the role of demographic processes in plant–plant interactions, albeit at a constant and thus somewhat unnatural spacing distance between individuals (in the absence of mortality). For example, net biodiversity effects observed at the community level could be traced back to enhanced *per capita* growth of single individuals (Potvin & Gotelli 2008; Potvin & Dutilleul 2009). Individual-based approaches also allow for determining the role of intraspecific trait plasticity for complementarity effects, or elucidating the role of individual vs. population fluctuations for ecosystem stability. Second, as species have different growth rates, interactions between different species are likely to change with age. Because biodiversity effects become more pronounced with time in grassland experiments (Marquard *et al.* 2009b; Reich *et al.* 2012), we may expect even stronger effects with trees. However, BEF experiments with trees contain a single cohort of individuals, whereas age-structured populations and communities can develop in BEF experiments with smaller organisms, including grassland perennials. Third, the large sizes of tree individuals allow for a more detailed study of interactions with the local topographic, microclimatic and edaphic environment, for which we here use the term ‘ecoscape’, and which can potentially be both co-driver and response of biodiversity effects in forest BEF experiments.

In addition to some of the above-mentioned disadvantages of using trees and shrubs in BEF experiments, they pose the following more specific challenges. Most important of all, they need large areas and long experimental time spans. Related to the latter are the potentially later onset of direct tree–tree interactions (whereas indirect interactions, e.g. via pathogens, may start early) between individuals due to large planting distance and slow growth compared with other plants and smaller, short-lived organisms.

#### A short manual for forest BEF experiments

The slow growth, large size and longevity of trees affect the appropriate plot size, planting density and mixing pattern for the observation of processes and functions that are close to those of natural mature forest stands. Furthermore, the large spatial scale causes high variation in initial microclimatic and edaphic conditions within plots and across the whole experimental site. This variation needs to be considered to obtain snap-shots of the ecoscape against which later changes in soil microclimate, soil properties and matter fluxes at the plant–soil boundary can be compared.

Biodiversity experiments rely on a species pool from which the different species combinations are then sampled to create communities of different species richness. This sampling can be performed in many ways. Here, we focus on the distinction between simulated random extinction scenarios, which at the same time ascertain that all species occur at all diversity levels

in the same number of plots, and nonrandom extinction scenarios that would occur if known extinction drivers differentially affected species, for example rare species or species with particular traits.

#### ILLUSTRATIVE EXAMPLE OF A FOREST BEF EXPERIMENT

We illustrate the design issues related to the different aspects marked above in bold with examples from the BEF-China experiment established in subtropical China in 2009/2010. Using a pool of 40 tree species, planted on 566 plots over a net area of 38.4 ha, we manipulated species richness and composition as well as genetic diversity to study their effects on a range of ecosystem functions, including primary productivity, carbon and nutrient cycling, soil processes and abundance and biodiversity of other trophic groups. A sloped experimental site was chosen to assess biodiversity effects on soil erosion as a particularly relevant ecosystem function provided by forests in subtropical China and elsewhere (Wang *et al.* 2005; Geißler *et al.* 2013). Table 1 provides a decision table to illustrate the decisions to be made when designing a BEF experiment with trees as well as the decisions taken in the specific case of BEF-China. The next sections follow the order of rows in Table 1. As specific questions in other forest, BEF experiments are best met with particular designs, the decisions to be made may follow different routes.

#### PLOT SIZE

The decision on plot size can be considered crucial as the relationships between tree diversity and ecosystem functions such as productivity in non-experimental studies have been found to be scale-dependent (Chisholm *et al.* 2013). As a rule of thumb, to prevent edge effects, Scherer-Lorenzen *et al.* (2005a) suggested a side length of a plot of about twice the maximum final tree height, usually corresponding to a plot area of 0.5–1 ha. However, the only two forest BEF experiments using plots of this size or even larger – the Borneo experiment (Hector *et al.* 2011) and the BIOTREE experiment in Germany (Scherer-Lorenzen *et al.* 2007; Fig. 1) – could only be set up in the context of large afforestation projects. Such contexts may not always be available or impose restrictions on other design choices. While such large plot sizes help to reduce edge effects when plots are surrounded by open land, they may not be required in a matrix of established forest or of adjacent experimental forest plots. However, arranging plots of different species composition next to each other will create more subtle between-treatments edge effects. Thus, trees along the plot margins should be treated differently from central trees and ecosystem functions should be measured in the centre of plots, as we do in our BEF-China experiment. Interestingly, in the only grassland biodiversity experiment addressing plot sizes, biodiversity effects did not differ at all between grassland plots of 3.5 × 3.5 and 20 × 20 m, and the ecosystem function of plant productivity could be measured successfully on even smaller scales (Roscher *et al.* 2005).

For statistical reasons, we strongly advocate minimizing plot size in favour of greater number and improved management (e.g. weeding) of plots. More replicates increase the statistical power to detect plot-level biodiversity effects (Schmid *et al.* 2002; Table 1, no. 1). The power analysis in Fig. 2 shows the probability of detecting the biodiversity effect on basal area we measured in our comparative study plots (CSPs) in a natural forest near to the experimental site (Gutianshan National Nature Reserve, hereafter GNRR; Bruelheide *et al.* 2011). A statistical power of 0.8 is achieved with 35 plots. Thus, 35 plots (e.g. of one of our random extinction scenarios, Table 2, second line) provide us with an 80% chance to detect a biodiversity effect comparable with that found in our CSPs.

It is desirable that the plot size in forest BEF experiments is at least as large as needed to potentially accommodate all species at the highest diversity level with at least one mature individual. Whether this will eventually be achieved of course not a choice left to the researcher, because differential mortality and potential extinction of species from plots is a normal and accepted feature in BEF experiments (e.g. with a strong selection effect few species may eventually remain in an originally diverse mixture, what matters is that the few species ‘were selected’ from an originally large number of species). Nevertheless, extinctions will be less likely if initial population sizes of species are not too small (Outbor 1993). The desirable minimum plot size thus increases linearly with the highest diversity level included in a BEF experiment and with the reciprocal of planting density, that is, the square of distance between trees.

In BEF-China, we calculated the desirable minimum plot size such that 130 years after establishment a single plot could still hold one mature individual for each of 16 species in the most extreme case. With a mature average canopy size of 36 m<sup>2</sup>, as expected for 130-year-old trees in subtropical forests of the study region (Bruelheide *et al.* 2011 and A.C. Lang, pers. comm.), the minimum plot size would then be 576 m<sup>2</sup>, which we increased to 667 m<sup>2</sup>, that is, 25.82 × 25.82 m, the traditional Chinese areal unit of 1 mu (Table 1, no. 1). This size will allow some species in diverse plots to have clearly above-average canopy sizes (whereas others will have below-average canopy sizes). Our CSPs mentioned previously also have a size of 1 mu appropriate to record plot-level forest functions such as tree growth (Bruelheide *et al.* 2011), rainfall characteristics (Geißler *et al.* 2013) and herbivory (Schuldt *et al.* 2012). However, because we also aimed at higher population sizes per tree species (a minimum of four individuals per species per plot) and to allow for a nested shrub diversity treatment, we assembled some of our 1-mu plots with identical tree species composition into parcels of 2 × 2 plots, yielding 4-mu (0.267 ha) ‘super-plots’ (see first line in Table 2). Given a total area of 50 ha available at our two experimental sites, a plot size of 1 mu allowed us to establish 566 plots (Table 1, no. 2).

#### PLANTING DENSITY

It is common practice in forestry to plant higher densities of saplings than the desired final tree density, anticipating losses and allowing for selection (Nyland 2002; Scherer-Lorenzen



**Table 1.** Key considerations for designing a biodiversity-ecosystem functioning (BEF) experiment with trees

No.	Design issue	Considerations	Decision in BEF-China	Pros	Cons
1	Large vs. small plot size	Has to be large enough to accommodate mature trees of each species	Small: 1-mu plots = 666.66 m <sup>2</sup> In addition: medium-sized 'super-plots': 4-mu plots = 2666.66 m <sup>2</sup>	Small plots may decrease within and between-plot environmental heterogeneity	Small plots allow for fewer replicates of individuals of the tree species within a plot (less insurance for random losses)
2	Large vs. small size of the experimental site	Impact on plot size and required number of plant individuals	Large: 50 ha, can accommodate 566 plots of 1 mu	Large size allows for larger plots and/or a higher number of plots	Large size implicates large establishment and management efforts
3	High vs. low planting density	Impact on establishment rates	High: 400 trees per 1 mu-plot (=0.6 trees per m <sup>2</sup> )	High density increases opportunities for early interactions	Increasing establishment effort with density, but probably decreasing weeding effort
4	Random individual planting pattern vs. regular individual or patch wise mixing pattern	Impact on opportunities for interspecific vs. intraspecific interactions	Random individual planting pattern	Random individual planting pattern enables interactions at early stages of the experiment	Regular individual planting pattern involves a higher risk of losing species at early stages of the experiment
5	Complex vs. simple site topography	Impact on slope and aspect of plots	Average plot slopes of site A and B are 27.5° and 31°, respectively	Slopes offer more realistic growth sites for forests than flat sites in many parts of the world	Increasing within-plot and among-plot heterogeneity with increasing slopes
6	Completely randomized design vs. blocking	Diversity effects might be confounded by between-plot environmental heterogeneity	No blocking, instead at each site a minimum distance rule was applied: replicates of a particular species mixture had to have a larger distance than 100 m	Useful when the direction of environmental gradients is unknown	Blocking is preferable, when there are few and clear gradients across a site
7	Ecoscape (measurement of microclimatic and edaphic conditions across the experimental site: yes or no?)	Assessment of between- and within-plot environmental heterogeneity	Repeated measurements of environmental variables such as soil erosion, evaporational demand, plant water availability in a grid across the study sites	<i>A posteriori</i> control for environmental heterogeneity	Laborious measurements of a multitude of environmental variables required
8	Multi- vs. single-site experiment	Extrapolation of the results beyond the single-site/single-species pool	Two sites were established 5 km apart, at each there are three different species starting compositions reflecting different natural ecosystems	Higher generality	Increasing establishment effort
9	Large vs. small species pool	Impact on design	Large pool with 40 tree species, 24 at each of the two sites, with an overlap of 8 species, plus 2 commercial plantation species	The more species, the more different compositions, the larger the gradient in species traits	Increasing investment of resources to monocultures with increasing pool size
10	High vs. low maximum plot richness	Interdependence with plot size	High maximum plot richness of 24 tree species, but extinction scenarios were chosen from sets of 16 species	The higher the maximum richness, the higher the variation in species traits in a plot, the more opportunities for neighbour interactions	Decreasing fraction of mixtures that can be realized with increasing maximum richness
11	Random vs. trait-oriented extinction scenarios	Degree of approximation of natural processes	Combination of random and trait-oriented extinction scenarios, based on rarity and SLA	Nonrandom extinction scenarios might be closer to extinction processes in reality	Nonrandom extinction scenarios require precise predictions and knowledge of the species involved
12	Choosing species combinations per richness level by a random partitions design vs. fully randomly drawn design	Balance of sampling intensity of every species at each diversity level	Random partition of different sets of 16 species, resulting in nested 16, 8, 4, 2, 1 species communities within each set	Avoids selection effects that cannot be accounted for	–

Table 1. (continued)

No.	Design issue	Considerations	Decision in BEF-China	Pros	Cons
13	Nested vs. non-nested extinction scenarios	Simulating extinction scenarios	Nested partitions according to a 'broken-stick' design with 6 subpools	The nested diversity levels correspond to a single extinction scenario	Independence between diversity levels
14	Replicated vs. single species compositions	Separating between-plot variation from variation between species compositions	Species compositions of the random extinction scenario replicated on 4-mu and 1-mu plots, those of the trait-oriented scenario on two 1-mu plots	Allows for testing effects of particular species compositions with unique effects on ecosystem functioning	Replication at higher levels (i.e. sites) instead of at the level of species compositions might result in higher generality of the results
15	Considering vs. ignoring genetic variation within species	Assessment of the interaction of species and genetic diversity	Seed families were tracked for 13 tree species; each of them was represented by up to eight seed families per 1-mu plot at site B	Allows for the separation of genetic variation from random variation among individuals	Increased effort in tracking seed sources

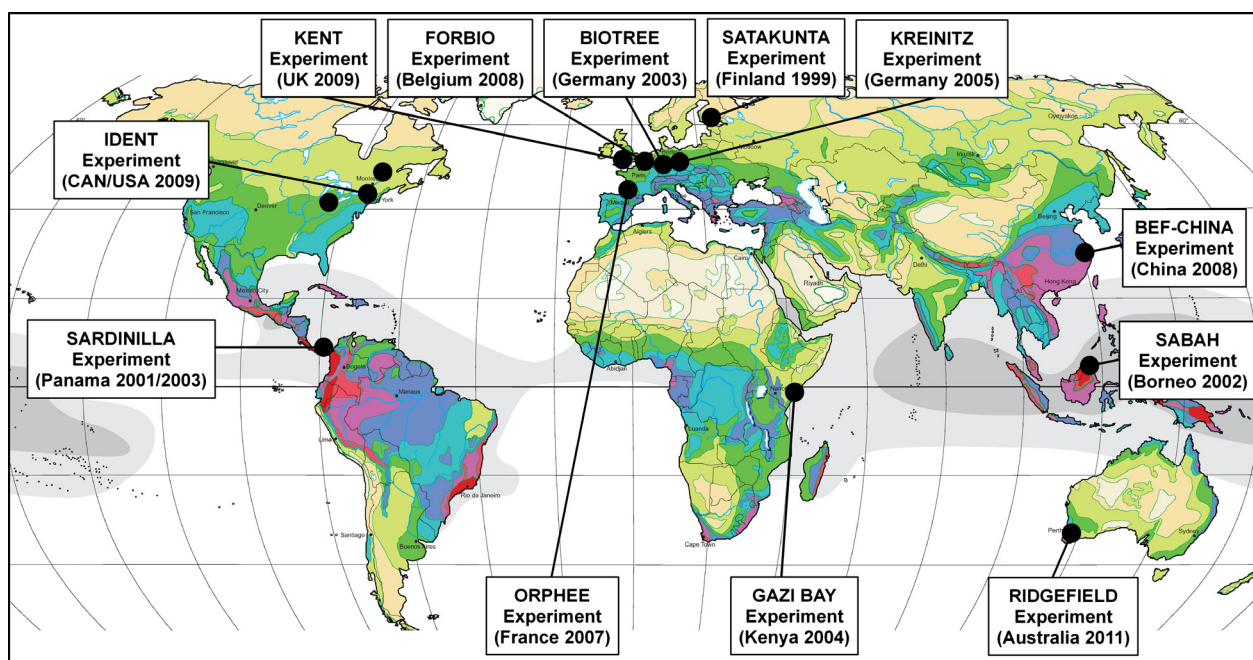
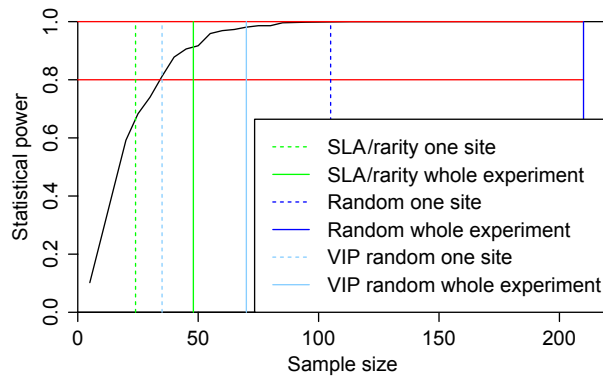


Fig. 1. Location of the BEF-China sites and of all other established forest experiments worldwide with tree diversity manipulations. Note that further experiments are in the planning stage. The map shows species richness of vascular plants, with blue colours 2000–3000, magenta 3000–4000 and red >5000 species per 10,000 km<sup>2</sup>. With permission from Barthlott *et al.* (2005). ([http://www.uni-bonn.de/~unb147/biodiversity/Plantdiversitymap\\_Barthlott\\_etal2005.pdf](http://www.uni-bonn.de/~unb147/biodiversity/Plantdiversitymap_Barthlott_etal2005.pdf)).

*et al.* 2005a). Moreover, higher initial densities mimic naturally regenerating forests and allow for early plant–plant interactions (Both *et al.* 2012; Lang *et al.* 2012). While higher initial densities imply higher costs for raising and planting more seedlings, they increase the flexibility to maintain desired species richness and evenness, in particular in the case of irregular natural mortality of seedlings. Higher initial densities also allow for more rapid canopy closure, which facilitates suppression of weeds, commonly the largest management effort in biodiversity experiments (Marquard *et al.* 2009b; see also Weiner *et al.* 2010).

Planting distance between individual saplings in the BEF-China experiment was 1.29 m in horizontal projection

(Fig. 3a) which is rather low compared with other forest BEF experiments (Scherer-Lorenzen *et al.* 2005a). This corresponds to 400 seedlings per 1 mu or c. 6000 seedlings per ha (Table 1, no. 3). As all plots in our experiment were planted at the same community density, population sizes per species decrease inversely with increasing species richness. This substitutive design (de Wit 1960) is commonly used in BEF research and avoids a problem encountered with additive designs, which lead to larger apparent ecosystem functioning effects of higher diversity than substitutive designs do, which, however, are due to confounding community density and diversity (Balvanera *et al.* 2006). To maintain our desired plot compositions during the phase of early establishment, herbaceous and nonplanted



**Fig. 2.** Power analysis for detecting the biodiversity effect on basal area we measured in our Comparative Study Plots. The analysis was based on the 12 youngest plots with a stand age of <60 years since the last logging event in the Gutianshan National Nature Reserve (Bruelheide *et al.* 2011) close to the experimental site. Using the effect size and variance of these plots, we simulated 1000 data sets for sample sizes from 5 to 200 plots. Each of these data sets was then analysed for a biodiversity effect. We counted the number of analyses, in which the biodiversity effect given in the simulation was significantly different from zero at a 5% level. The statistical power (black line) shows the proportion of the significant analyses. The horizontal red lines indicate the thresholds of 80% and 100% statistical power. The vertical lines indicate the sample sizes of three main aspects of the experiment: the trait-oriented extinction scenarios using specific leaf area (SLA) and rarity (SLA/rarity; lines 5 and 6 in Table 2), the first random extinction scenario on large plots (VIP, Very Intensively Studied Plots, line 1 in Table 2), and the complete sample size of the random extinction scenarios, including replicates and the 2nd and 3rd random extinction scenarios (Random; lines 1–4 in Table 2). Broken lines indicate the sample size of one of the experimental sites (A or B), solid lines the sample size for both experimental sites combined. The intersection of the power curve in black and the sample sizes for the different aspects of the experiment give the statistical power of this part of the experiment. The trait-oriented scenarios SLA/rarity exceed a statistical power of 80% when the two sites A and B are combined, while in the VIP random extinction scenario on the large plots, the statistical power of 80% is already achieved using the sample size of one of the sites alone.

woody species have been removed twice per year from all plots in weeding campaigns. Originally, the planned duration of the experiment over 100 years was considered appropriate to allow all species to equilibrate with associated soil microbial communities (e.g. Visser 1995), herbivores and predators (e.g. Watt, Hunter & Stork 1997) and to capture the most relevant forest dynamics (e.g. Pretzsch & Schütze 2009). However, based on our experiment, we can now testify that interactions between tree species start even in the first years after planting. For example, crowns of the fast-growing species began to touch each other in the third year. This is also supported by observations on crown interactions between the initially fast-growing deciduous and some of the initially slow-growing evergreen broadleaved species in native forest remnants of the study region (von Oheimb *et al.* 2011).

#### MIXING PATTERN

The spatial arrangement of individuals and species may affect species interactions (Pacala & Deutschman 1995; Stoll

& Prati 2001; Potvin & Gotelli 2008). While regular mixing of species maximizes interspecific interactions, patchwise mixing minimizes them, at least during early stages of stand development (Scherer-Lorenzen *et al.* 2005a). These two approaches define the endpoints of a design trade-off between optimizing early detection of biodiversity effects and long-term maintenance of biodiversity gradients. Weak competitors among the species may quickly be out-competed by strong ones in regularly mixed plots (e.g. ABABABAB, where each letter represents an individual of species A or B), reducing 'realized' species richness. This may be prevented by patchwise planting (e.g. AAAABBBB) to ensure survival of weak competitors with minimum tending efforts (Saha *et al.* 2012). For example, patchwise mixing of species was used in the BIOTREE experiment in Germany, where patch size was defined as the mature crown area of the species with the largest crown requirement (Scherer-Lorenzen *et al.* 2005a). However, patchwise mixing has the disadvantage that most neighbour interactions between trees are initially intraspecific, even in diverse plots.

In BEF-China, we first considered to use patchwise vs. individual mixing of species as an additional split-plot treatment but then decided against to reduce the complexity of the design. Instead, we assigned species randomly to individual planting positions (e.g. ABBABAAB) on a regular quadratic grid within plots (Table 1, no. 4). With this, we aimed to avoid the potential disadvantages of either regular individual or patchwise mixing. In addition, random mixing has the advantage that local neighbourhood diversity will vary to some degree, which is more natural and allows a richer analysis of the response of individual trees to diversity than do the other two planting schemes. For example, the regular mixing of six species in Potvin & Gotelli (2008) led to a maximum local diversity of only three species. With random mixing, local neighbourhoods vary from clumped, that is, monospecific patches, to very diverse patches. In our case, the most diverse possible neighbourhood would include eight different species in the four closest and four diagonal neighbours around a single target tree.

#### TOPOGRAPHY

Generally, BEF experiments, including those involving trees, have been established on flat land (Fig. 1, but see Paul *et al.* 2012). This facilitates manual work. Moreover, ecosystem variables expressed per horizontal area do not need to be corrected for slope. In many regions around the world, flat land is more fertile than sloped land, due to alluvial deposits, and is thus used for agriculture, whereas forests are allowed to grow on slopes, where in addition to providing timber they help to prevent landslides, rockslides, avalanches and erosion. However, establishing forest BEF experiments on slopes poses additional challenges such as dealing with topographic heterogeneity within and among plots, substantially more manual work (Fig. 3b) and issues of slope correction, that is, whether ecosystem variables – including the density of individuals and of species, standing crop, litter decomposition, stocks of elements,

**Table 2.** Overview of all extinction scenarios, size and number of trees planted at site A (top) and B (below) of the tree biodiversity-ecosystem functioning experiment BEF-China (Xingangshan, Jiangxi Province)

	0	1	2	4	8	16	24	Number	Plot size (m <sup>2</sup> )	mu	Size (ha)	Number of trees	Number of shrubs	
Site A Tree species richness														
1st Random extinction scenario	3	16	8	4	2	1	1	35	51.64 × 51.64	4	9.33	56,000	42,000	
1st Random ext. seen. replicate		16	8	4	2	1	1	32	25.82 × 25.82	1	2.13	12,800	0	
2nd Random ext. seen.		8	8	4	2	1		23	25.82 × 25.82	1	1.53	9200	0	
3rd Random ext. seen.		0	8	4	2	1		15	25.82 × 25.82	1	1.00	6000	0	
Non-random extinction scenario		0	6	6	6	6		24	25.82 × 25.82	1	1.60	9600	0	
Non-random ext. scen. replicate		0	6	6	6	6		24	25.82 × 25.82	1	1.60	9600	0	
Economically important species		10						10	25.82 × 25.82	1	0.67	4000	0	
Free succession plots								3	25.82 × 25.82	1	0.20			
Shrub monocultures		20						20	12.91 × 12.91	0.25	0.33		2000	
Total	(in terms of 1 mu-plots)							271				18.40	107,200	44,000
Site B Tree species richness														
1st Random extinction scenario	3	16	8	4	2	1	1	35	51.64 × 51.64	4	9.33	51,200	42,000	
1st Random ext. scen. replicate		16	8	4	2	1	1	32	25.82 × 25.82	1	2.13	12,800	0	
2nd Random ext. scen.		8	8	4	2	1		23	25.82 × 25.82	1	1.53	9200	0	
3rd Random ext. scen.		0	8	4	2	1		15	25.82 × 25.82	1	1.00	6000	0	
Non-random extinction scenario		0	6	6	6	6		24	25.82 × 25.82	1	1.60	9600	0	
Non-random ext. scen. replicate		0	6	6	6	6		24	25.82 × 25.82	1	1.60	9600	0	
Economically important species		10						10	25.82 × 25.82	1	0.67	4000	0	
Genetic diversity scenarios		10		14				24	25.82 × 25.82	1	1.60	9600	0	
Free succession plots								3	25.82 × 25.82	1	0.20			
Shrub monocultures								20	12.91 × 12.91	0.25	0.33		2000	
Total	(in terms of 1 mu-plots)							295				20.00	112,000	44,000

The total number of 566 (=271 + 295) refers to 1-mu plots. For the selection of species compositions and monocultures (i.e. the first random extinction scenario referred to as species pool 1 at each site), four such plots were arranged in a 2 × 2 plot super-plot (first line in each of the tables). Within these super-plots, the four plots ranged in shrub richness from 0 to 2 to 4 to 8 species. The second line shows the replicates of the species compositions of these scenarios. The third and fourth lines refer to the random extinction scenarios starting from species pools 2 and 3, respectively, at each site (see Table A1). The fifth and sixth lines show the trait-oriented (nonrandom) extinction scenarios for the specific leaf area (SLA) and rarity scenario, respectively (same at the two sites). The seventh line shows the plots used for commercial plantation species, with five plots each per site for *Cunninghamia lanceolata* and *Pinus massoniana*. Finally, the eighth line in the table for site B shows the plots for the genetic diversity scenarios, which were only established at site B. Finally, there are 3 free succession plots at each site. In consequence, the total numbers of 1-mu plots are 271 and 295 for sites A and B, respectively, not considering 0.25-mu plots for shrub monocultures.

nutrient turnover and erosion – should be measured per surface area or per area of horizontal projection.

In BEF-China, we established our plots on sloped terrain (average slope 27.5° for site A and 31° for site B, Table 1, no. 5), accepting the additional topographic heterogeneity across spatial scales, which we address in the next section ('Ecoscape'). We plan to do slope corrections and will express ecosystem variables per area of horizontal projection, to be consistent with geographical maps. However, we did encounter a problem when measuring root biomass and productivity with soil cylinders and in-growth cores; here, we decided to change from vertical positioning to positioning perpendicular to the slope.

#### ECOSCAPE

To cope with large-scale spatial heterogeneity of experimental sites, it is generally recommended to block experimental plots (Dutilleul 1993). Blocking is also used to avoid accidental clumping of similar treatments that may occur in completely randomized designs (Hurlbert 1984). However, when there are many different treatment combinations, blocks would have to

be very large to contain each treatment combination at least once, rendering it difficult to delineate blocks. In such cases, it can be preferable to use a completely randomized design (Table 1, no. 6). Even without blocking the potential of accidental spatial clustering of replicates of the same treatment combinations can be avoided by constraining complete plot randomization with a minimum distance rule that must be maintained between replicates. Furthermore, blocks can still be defined a posteriori to account for spatial heterogeneity at the between-plot scale. Alternatively, environmental covariates can be measured and used to adjust for potential environmental heterogeneity across a study site, an approach that we refer to here as 'ecoscape' and which we explain in detail in the following two paragraphs.

The reason to adjust for environmental heterogeneity in BEF experiments is that it may interfere with biodiversity effects on tree growth. For example, in the Sardinilla forest, BEF experiment in Panama even comparatively low environmental heterogeneity had strong effects on productivity and plant mortality (Healy, Gotelli & Potvin 2008). If environmental heterogeneity has a low-dimensional spatial pattern, it can usually be accounted for by using spatial coordinates and func-





**Fig. 3.** Establishment of BEF-China in Jiangxi Province, China. (a) Main experiment at site A after planting. The bamboo sticks mark planting positions of trees. (b) Planting trees on a 35° slope on site A. (c) One of the two nurseries of BEF-China in Dexing. The picture shows one of the five greenhouses used in 2009 to grow about 500,000 seedlings for site B.

tions of them as covariates (see e.g. Le Roux *et al.* 2013 for a grassland BEF experiment). However, if this is not the case, the environmental heterogeneity should be measured directly to characterize the ecoscape of the site at which an experiment is established. It is then possible to use the environmental measures defining the ecoscape as covariables in the statistical analysis of biodiversity effects. When using an ecoscape approach, it is important to recognize that biodiversity itself can and will influence the ecoscape as the experiment progresses. However, quantifying the effect of how biodiversity affects the ecoscape can only be performed if the initial ecoscape has been measured.

While it is necessary to control for environmental covariates when testing biodiversity effects, it is also important to be able

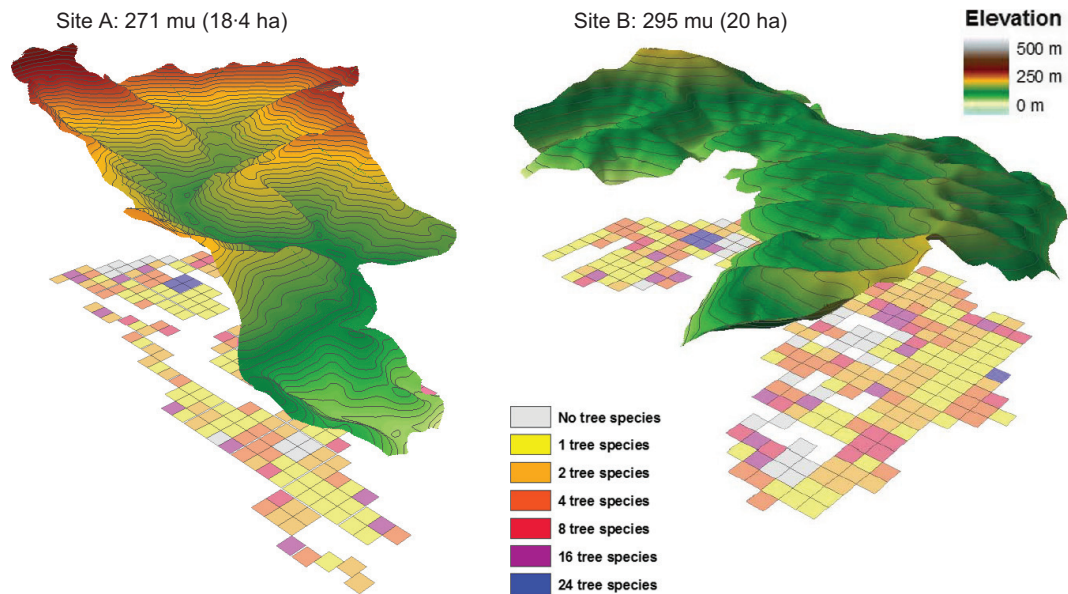
to generalize BEF relationships across different sites varying in topography, geology, soils and climate (Table 1, no. 8). For the greatest possible generalization, each experiment should be replicated at different sites in a region (Hector *et al.* 2002) and with different species pools (Wacker *et al.* 2009).

In the very heterogeneous slopes of our BEF-China experiment, and given the many diversity levels, blocking was not possible. We therefore used a completely randomized design with the constraint that replicates of the same treatment combination were located at least 100 m apart from each other. We are now characterizing the ecoscape of the experimental sites (Table 1, no. 7) by taking initial measures of environmental variables, such as soil erosion, evaporation and plant water availability in a grid across the study plots, and also measuring density and other features of herbaceous vegetation within plots (Both *et al.* 2012). We will then be able to relate these environmental variables to spatial covariates such as coordinates, altitude, slope and aspect (Yang *et al.* 2013). Later, it will allow us to analyse how diversity treatments modify the ecoscape during the course of time. Finally, to increase generality, we established plots at two sites A and B (Fig. 4) with two separate species pools that overlap by only eight of 40 total species in our BEF-China experiment.

#### SPECIES POOL

A crucial aspect in BEF experiments is the species pool, the total set of species from which individual species are chosen to create experimental communities. It should represent the full tree community that could naturally occur at a site to enable generalization with respect to the natural ecosystem. However, many biodiversity experiments start with already reduced local species richness at the highest diversity level. For example, six and 28 species used in a tropical forest BEF experiment in Panama (Potvin & Dutilleul 2009; C. Potvin pers. comm.) or 16 dipterocarp species used in a similar experiment in Borneo (Hector *et al.* 2011) represent a small proportion of species and functional diversity typically occurring in adjacent natural forests. In other cases, the opposite occurs: the five species used in a boreal forest BEF experiment in Finland (Vehviläinen & Koricheva 2006), and the 16 species used in a temperate forest BEF experiment in Germany (Scherer-Lorenzen *et al.* 2007) probably represent a higher species and functional diversity than typically occurs in adjacent natural forests.

Beside reflecting natural species assemblages, large species pools also allow for planting higher species diversity and functional diversity per plot. Assuming that species differ in their traits because of limiting-similarity constraints (MacArthur & Levins 1967; Pacala & Tilman 1994; Weiher & Keddy 1999), larger species pools should cover larger portions of the multivariate trait space. Inconveniently, however, larger species pools also increase the number of required plots. Moreover, to assess species-specific contributions to the functioning of mixtures, all species should also be studied in monocultures (Loreau & Hector 2001). The number of monoculture plots can easily reach or exceed the number of mixture plots when a large species pool is used. Thus, one of the greatest challenges



**Fig. 4.** Arrangement of plots at both sites of BEF-China at Xingangshan (Jiangxi Province), illustrating different tree species richness levels. The chessboard pattern shows the 1-mu plots. Plots shown in grey colour are treatments without trees. Plots in yellow colour show monocultures but also include plots planted with commercially important trees, that is, *Cunninghamia lanceolata* and *Pinus massoniana*.

for large forest BEF experiments is raising sufficient numbers of individuals of all desired species. This is a tremendous logistic endeavour in countries with high tree species richness where often only a small proportion of them are used in forestry and where the majority of other species cannot be acquired from commercial nurseries (Yang *et al.* 2013).

In BEF-China, we included 40 native broad-leaved tree species and 18 shrub species in the entire species pool, which is high compared with other BEF experiments (Table 1, no. 9; Appendix Table A1). In addition, we added two commonly used commercial coniferous plantation species for monoculture comparisons (*Cunninghamia lanceolata* and *Pinus massoniana*). We were able to grow more than 1 million seedlings of our target species in two specifically contracted nurseries (Fig. 3c). At the time of planting, all seedlings had approximately the same age between 1 and 2 years and had a minimum height of 20 cm to allow for placing metal labels on them and insure their maintenance. To avoid the problem of increasing overlap of species compositions with increasing species richness (variance-reduction effect of Huston 1997; see following section), we used random extinction scenarios in which three (overlapping) pools of 16 tree species at each site were divided into nonoverlapping communities at lower diversities (see following section; Table 2, Table A1). In addition, we included an extra-'high' richness level of 24 tree species, combining species from the different pools (Table 1, no. 10). While we did not consider functional diversity as an additional design variable, we insured that our tree species represented a large range of families. This increases the chance that high species richness levels will also reflect high functional diversity levels but has the disadvantage that our design cannot distinguish between the two. Nevertheless, we will be able to use functional richness measures instead of species richness in forthcoming

analyses as we are assessing traits for all species in our experiment. Furthermore, our design, like other BEF experimental designs, always allows us to contrast communities in which particular functional groups of species such as deciduous or evergreen trees are present with communities in which they are absent.

The 18 shrub species were used to create shrub-richness levels of 2, 4 or 8 shrub species of a total pool of 10 shrub species per site. These were factorially crossed with tree species richness in the 4-mu super-plots (mentioned in the section 'Plot size' above) for one of the three random extinction scenarios at each site (yielding two random scenarios with shrub diversity treatment in total) and for the 24-tree species communities. The shrubs were planted in the interspaces of the trees, with the same density (400 individuals per mu), and thus, had the same distance as trees from each other (1.29 m) and a distance of 0.91 m to the nearest tree neighbour. In the end, our plots with the highest diversity of 24 tree plus eight shrub species per 1 mu attained a richness level comparable with the mean richness of natural forests near the experimental site, where we observed a mean richness of 23.4 tree and 18.4 shrub species per  $30 \times 30$  m CSP (Bruehlheide *et al.* 2011).

#### RANDOM EXTINCTION SCENARIO

As the fundamental question of BEF experiments is whether random species loss reduces the degree of ecosystem functioning, a central design question has always been how such species loss is best achieved (Lamont 1995; Allison 1999; Mikola, Salonen & Setälä 2002; Scherer-Lorenzen *et al.* 2005a; Balvanera *et al.* 2006; Schmid *et al.* 2008; Table 1, no. 11).

The first issue is how to create a gradient of diversity, which can be performed by sowing or planting defined num-



bers of species into plots or by deleting species from existing diverse communities. In most BEF experiments, species loss has been simulated by using predefined levels of diversity, for example 1, 2, 4, 8 and 16 species. The alternative method for simulating species loss, experimental species removal from established communities (Díaz *et al.* 2003) may appear tempting, as it would allow for working with mature trees from the start. Unfortunately, however, this approach has several severe shortcomings. The major shortcoming is a confounding of resulting biodiversity with disturbance and community density. To compensate, trees would also need to be removed from high-diversity treatments to reach the same low community density as in low-diversity plots. This in turn would lead to unnaturally low densities. Additionally, many ecosystem processes, including those below-ground, may continue to be determined by the original tree diversity over long time periods, so that removal would control standing diversity, but not the associated ecosystem processes. As roots cannot be reliably extracted, one very obvious example concerns roots biomass, decomposition and root-related processes.

The second and perhaps more important issue for determining how species loss alters ecosystem functioning is how to establish experimental replicates. Early BEF experiments used only one random extinction series with community replicates of this series (e.g. Naeem *et al.* 1994; Niklaus *et al.* 2001). However, exact replicates of a single extinction series offer no ability to distinguish plant diversity effects per se from effects of the particular community and its component species. To distinguish those effects, it is necessary to have multiple random extinction series that provide different specific species compositions at each plant diversity level (Givnish 1994). Therefore, in BEF-China, we used three partially overlapping sets of 16 species each (six total sets), of the total pool of 40 tree species (Table A1). Each species set was then used to create a separate random extinction scenario, for a total of six random extinction scenarios across the two sites.

While replication at the extinction scenario level is required to separate diversity effects from effects of particular communities, replication of particular species compositions and monocultures is necessary to separate these effects from between-plot variation (Table 1, no. 14). For example, to statistically test transgressive overyielding of mixtures compared with monocultures, replicated monocultures and mixtures must be available (Schmid *et al.* 2008). In BEF-China, to replicate species compositions without doubling the required labour, we established a full replicate of one of the three random extinction scenarios at each site. The replicated random extinction scenario was thus established once on a 4-mu superplot (allowing for the shrub-richness treatment to be applied among the four plots of this super pot) and once on a separate 1-mu plot (Table 2).

The next issue in establishing a gradient of species loss is how to best draw random communities out of the total species pool for each extinction scenario. The total number of different  $k$ -species mixtures that can be drawn from a pool of  $n$  species is given by the binomial coefficient:

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

For a pool of 16 species, the number of different 1-, 2-, 4-, 8- and 16-species compositions are 16, 120, 1820, 12,870 and 1, respectively. Realizing all of them on 1-mu plots would require a total area of about 10 km<sup>2</sup>. Typically, this is beyond the scope of any single BEF experiment. Fortunately, testing for diversity effects does not require all species combinations, but a reasonable number of representative compositions per diversity level.

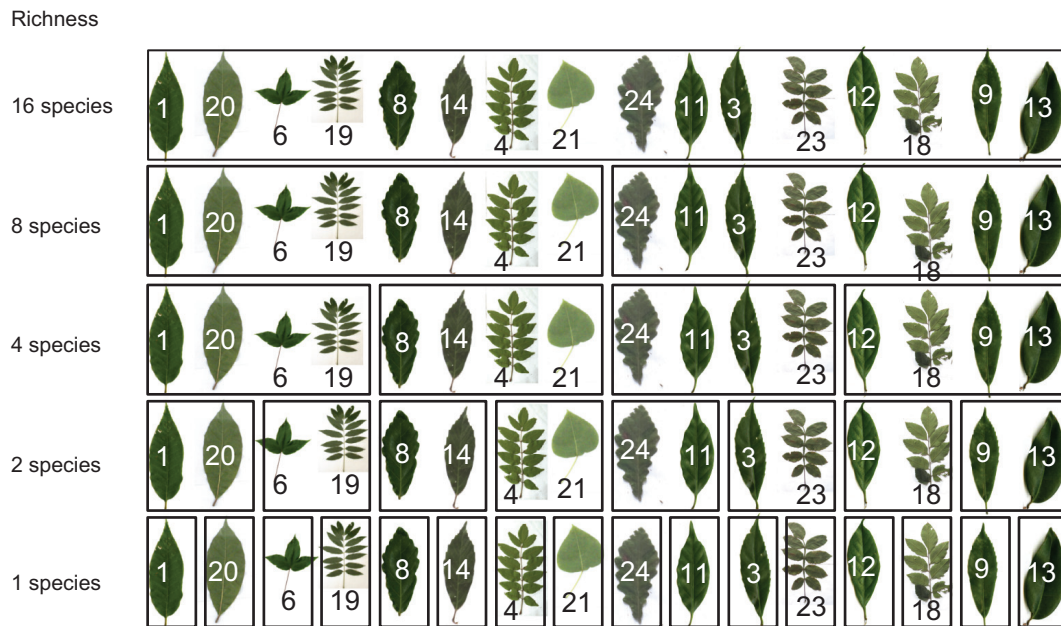
Choosing such subsets can be performed in a nested way, by successively leaving out species from more diverse communities to less diverse ones, for example, along an extinction scenario of 16–8–4–2–1 species per plot (Table 1, no. 12). In BEF experimental designs where such purposeful nesting is not performed (fully randomly-drawn designs), the number of diversity levels and plots per levels in which species occur, differs very widely between species, causing nonorthogonalities between the presence or absence of particular species and diversity levels (see e.g. Roscher *et al.* 2004) and preventing analyses of species-specific responses to biodiversity. Furthermore, fully randomly drawn non-nested designs have increasingly overlapping species compositions at higher diversity levels, leading to the so-called variance-reduction effect (Huston 1997; Bell *et al.* 2009).

In BEF-China, we have ensured that all species are equally represented at each diversity level (Table 2) by using a ‘broken-stick’ design. In this ‘broken-stick’ design (Fig. 5a), the starting species composition was randomly partitioned into nonoverlapping fractions (Bell *et al.* 2005, 2009; Salles *et al.* 2009). At each of the two sites of BEF-China, the three partly overlapping starting compositions (species sets, Table 2, Table A1) of 16 species each were randomly broken down into nonoverlapping eight-species compositions (Table 1, no. 13). To include every community of lower diversity as a subset of a community of higher diversity, we continued the partitioning for the eight-species compositions and so on, thus obtaining unique random extinction scenarios down to the monoculture of every species (Fig. 5a).

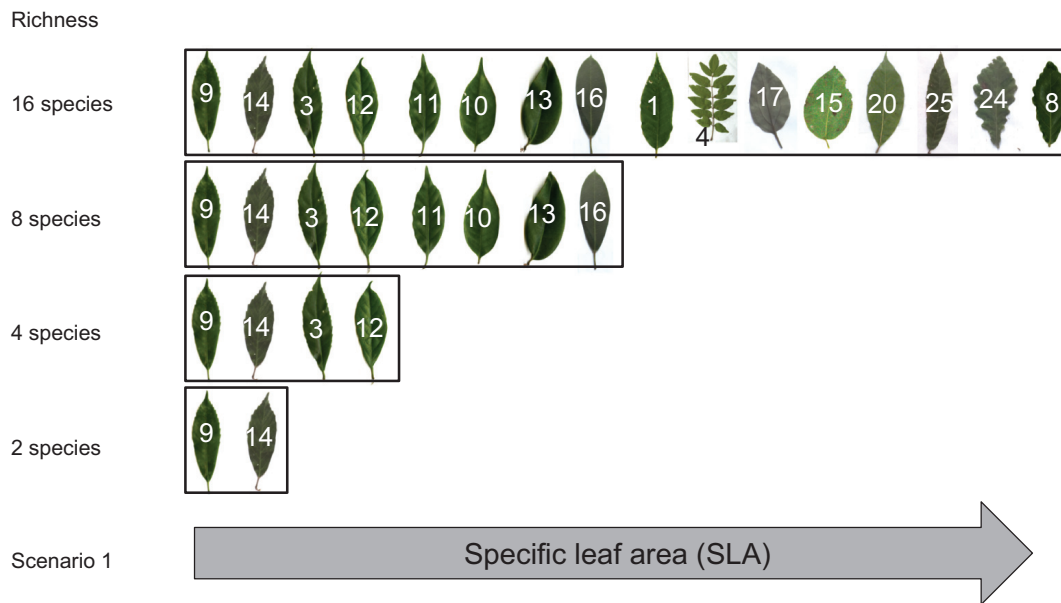
#### NONRANDOM OR TRAIT-ORIENTED EXTINCTION SCENARIOS

Most BEF experiments so far have used random extinction scenarios because species differences in extinction proneness predicted under different future environmental scenarios are largely unknown (Schmid & Hector 2004; but see Schläpfer, Pfisterer & Schmid 2005). However, random extinction scenarios may underestimate (Zavaleta & Hulvey 2004) or overestimate (Schläpfer, Pfisterer & Schmid 2005) effects of real-world biodiversity loss on ecosystem functioning because extinction in the real world might be biased towards species with particular features (e.g. Grime 2002; Lepš 2004; Schmid & Hector 2004; Solan *et al.* 2004). For example, Grime (2002) pointed out that stress-tolerant species, characterized by slow growth and long life span, may be more prone to go extinct

**(a) Random extinction scenarios**



**(b) Non-random extinction scenario**



**Fig. 5.** (a) Random partitions design in BEF-China, following a broken-stick procedure. Species are represented by numbers. For species names see Appendix Table A1. The pool of 16 species is randomly partitioned into two mixtures of eight species each. With each of the resulting mixtures, the partitioning process is continued. (b) Trait-oriented (=nonrandom) extinction scenarios in BEF-China, exemplified for specific leaf area (SLA; Scenario 1). The species richness levels are constructed from a pool of 24 species, concentrating the species with extreme trait values (the most common species and species with the lowest SLA in the two different nonrandom extinction scenarios) from the 16 species-mixtures to the two species-low-richness mixtures. For species codes see Appendix Table A1.

with increasing human disturbance. Therefore, in BEF-China, we added two nonrandom (in other words ‘trait-oriented’, ‘directed’, ‘biased’, or ‘informed’) extinction scenarios (Table 2, no. 11), one based on local rarity and one on specific leaf area (SLA). Eliminating species sequentially from the rar-

est to the second-most common can be considered a simulation of what might happen during habitat fragmentation and reduction, two on-going processes in Chinese subtropical forests (Wang, Kent & Fang 2007). We expect that at least some ecosystem functions may be differently affected by the prefer-



ential loss of rare tree species than by random species loss. For example, Schuldt *et al.* (2012) showed that herbivory on tree saplings in the GNNR increased with increasing local commonness of tree species.

Specific leaf area was chosen as the key trait for the second nonrandom extinction scenario because it reflects the leaf economics spectrum (LES) from long-lived and nutrient conserving leaves (low SLA) to short-lived ones with high-nutrient contents and high net photosynthetic capacity (high SLA; Reich *et al.* 1995; Wright *et al.* 2004). Moreover, Osnas *et al.* (2013) demonstrated that many mass-based leaf traits, such as photosynthetic capacity ( $A_{\max}$ ), dark respiration rate ( $R_{\text{dark}}$ ), nitrogen and phosphorus content, covary with the LES mainly because of their proportionality to leaf area. Community-weighted mean SLA typically decreases during succession in these subtropical forests as fast-growing species are replaced by more slowly growing species (Kröber *et al.* 2012). We expect that with proceeding succession or increasing temperature or dryness through climate change, species with high SLA (i.e. deciduous ones) should go extinct first. Information on rarity and SLA was obtained from the previously mentioned 27 CSPs in the nearby GNNR, first based on expert knowledge (Fang Teng, pers. comm.) and later corroborated by extensive trait measurements (Kröber *et al.* 2012).

The trait-oriented extinction scenarios were derived from starting compositions of the 20 most common species or the 20 species with lowest SLA, assembled from the total pool of 24 species at each of the two sites. Rare species or species with large SLA, respectively, were sequentially eliminated to yield decreasing diversity levels from 16- to 8-, 4- and 2-species mixtures (Fig. 5b), the monocultures already being available from the random extinction series (Table 2). Three different species compositions were constructed at each diversity level from a set of species that shared the same degree of commonness, or magnitude of SLA, respectively. In total, there were 24 different 1-mu plots each in the rarity and SLA scenario plots in each of the two sites (Table 2). With this sample size (48 plots in each site), the experiment has a statistical power of 0.9 (Fig. 2) to detect an existing biodiversity effect comparable with that found in our CSPs. In addition, we will be able to compare the strength of diversity effects between random and trait-oriented scenarios.

#### GENETIC DIVERSITY WITHIN-TREE SPECIES AND COMMUNITIES

Genetic variation within-tree species can largely affect variation in tree performance and tree response to different environments (Whitham *et al.* 2003). However, most BEF experiments have focused on manipulating species richness and ignoring population-genetic variation within species (Balvanera *et al.* 2006) implicitly assuming it is equally distributed among plots. With some additional initial effort, it would often be possible to consider population-genetic composition and diversity at least to some degree (Table 1, no. 15).

In our BEF-China experiment, we included trees of known maternal seed families at one of our two sites (site B). We can

thus test whether the population-genetic composition of tree populations affects the response to different species diversity levels or environmental variables (e.g. our ecoscape), including variables affected by species diversity itself, such as changed light or moisture levels. If this is the case, suggested, for example, by different responses of different seed families, it implies that species diversity may affect the evolution of the constituent species (Lipowsky, Schmid & Roscher 2011).

We also address whether the population-genetic diversity of the tree populations in the experimental communities affects tree performance, for example, because lower genetic diversity leads to higher herbivory or pathogen load (Henery 2011) and change multitrophic interactions (Moreira & Mooney 2013). To this end, we manipulated the genetic diversity of 24 experimental plots at site B by establishing them for each species with trees from either single or several seed families. The resulting factorial combination of species diversity (1 or 4 species) and genetic diversity (1 or 4 seed families per species) will allow us to examine, for example whether effects of plant species diversity are reduced for genetically depauperate communities.

#### MEASUREMENTS AND DATA

The design of forest BEF experiments implies that data can be obtained at several levels of interest, mainly the plot, individual tree and within-tree levels. Examples are measurements of soil lipids at the plot level (Wu *et al.* 2012), growth rates of individual trees (X.F. Li *et al.* in preparation) or leaf toughness (Schuldt *et al.* 2012) of individual leaves belonging to those trees. Adequate combination and aggregation of such data – with their long tail typical for extensive data set compilations (Heidorn 2008), that is, with few data sets with very large amounts of similarly structured data (e.g. from climate loggers) and many smaller data sets, with single or few measures of plots, species per plot or trees per plot – is required for addressing interdisciplinary questions at the level of the ecosystem.

As measurements are taken by different researchers or research teams based in different scientific laboratories across China and Europe, it is crucial to manage naming conventions simply, reliable and online. Therefore, building upon Ecological Metadata Language (Fegraus *et al.* 2005; Michener & Jones 2012), we developed a web application tailored for the needs of BEF-China that manages naming conventions within primary data (Nadrowski *et al.* 2013). Our application is open source, developed in concert with an R package to access data (BEF-China Research Group 2013) and can also be adopted by other collaborative research platforms.

#### STATISTICAL ANALYSES OF BEF EXPERIMENTS

The discussed design aspects of BEF experiments also have consequences for their statistical analyses (Schmid *et al.* 2002, 2009; Bell *et al.* 2009; Hector, von Felten & Schmid 2010). Basically, data from forest BEF experiments can be analysed at three levels: that of the plot, the population and the individual plant. Statistically speaking, the population level corre-

sponds to the combination 'plot identity × species identity' and the individual level to the combination 'plot identity × species identity × plant identity'. However, because in monocultures, there is only one species per plot, care has to be taken when analysing species responses at these levels, and we therefore recommend separate contrasts between species within monocultures and within mixtures, allowing them to be tested at the corresponding error levels (plot for monocultures and population for mixtures). Analyses at the population and individual level may include all variation from the plot level by using a single random term for plot identity; this can be convenient when focusing on within-plot effects.

The above procedure is possible because the random term for plot identity in a BEF experiment contains all information that can be used for the plot-level analysis, that is, it also includes variation among plots due to species diversity or species composition. To address the effects of diversity or composition, we therefore aim to use contrasts within the plot term. The first contrast in case of replicated species compositions is the one between the random term species composition or 'community' and the remainder, that is, residual plot variation (equivalent to plot identity within community). The random term community can then be further split into fixed terms of interest and remainder, that is, residual community variation. Typically, such fixed-term contrasts are monocultures vs. mixtures, (log-) linear species richness, species richness as multilevel factor; these three may also be incorporated into the analysis as sequential contrasts, one nested within the other.

Other fixed-term contrasts may be crossed with these, for example, the presence of within communities of particular functional groups or species (e.g. Yang *et al.* 2013). Additional levels come into the statistical analyses when repeated measures are made on plots, populations or plant individuals. These include measures of the same dependent variable at different time points or in different vertical positions as well as measures of different variables. Sometimes repeated measures can be aggregated into single-value measures of change (e.g. slope over time), variation (e.g. stability as the inverse of the coefficient of variation) or multifunctionality (Zavaleta *et al.* 2010; Isbell *et al.* 2011).

## Outlook

With our short manual to forest BEF experiments, we hope to contribute to the continued development of this new type of research in ecology and environmental studies. Forest BEF experiments provide exciting and timely research options requiring interdisciplinary collaboration and laborious set-ups. Achieving specific research goals and synergy with previous experiments requires manifold design decisions and calls for investing appropriate time and expertise when designing new experiments. Because of the complexity of issues related to the multifaceted aspects of biodiversity, the large effort in setting up and managing an experiment and the involvement of multiple research teams from different fields or expertise, careful planning and transparent decisions should be the guiding principle.

We do not suggest that all future forest BEF experiments should follow or build on the approach we used in BEF-China. Rather, it will be important that a diversity of approaches is followed to answer the diversity of questions facing us today in the context of global change and biodiversity loss. With our short manual, we aim to indicate which issues typically come up in BEF experiments and which reasoning can be involved in choosing between different design, measurement and analysis options. Ideally, experiments, such as BEF-China, should be interdisciplinary platforms complementary to each other in generating information and knowledge to tackle global problems, providing scientists the opportunity to 'think big'. BEF-China aims at being such an open research platform and encourages researchers worldwide to participate.

## Acknowledgements

We are indebted to the BEF-China students for their assistance in establishing the experiment, in particular to Martin Baruffol, Yang Bo, Martin Böhnke, Matteo Brezzi, Shane Frank, Christian Geißler, Wenzel Kröber, Anne Lang, Yinlei Ma, Gao Mo, Ricarda Pohl, Yann Salmon, Miaomiao Shi, Stefan Trogisch, Yu-Ting Wu, Yuanjie Xu and Xueqin Zeng as well as to Chen Lin and Kaitian Wang for coordinating the field work. We also highly appreciate the competence of Fang Teng from the Gutianshan National Nature Reserve and of the two nurseries in Dexing and Xingangshan. BEF-China is mainly funded by the German Research Foundation (DFG FOR 891/1 and 2), with additional funds from the National Natural Science Foundation of China (NSFC 30710103907, 30930005, 31170457 and 31210103910) and the Swiss National Science Foundation (SNSF). We also benefitted from various travel grants and summer schools financed by the Sino-German Centre for Research Promotion in Beijing (GZ 524, 592, 698, 699 and 785). Finally, we appreciate the comments of four anonymous referees, which contributed considerably to improving the manuscript.

## References

- Allison, G.W. (1999) The implications of experimental design for biodiversity manipulations. *American Naturalist*, **153**, 26–45.
- Bala, G., Caldeira, K., Wickett, M., Phillips, T.J., Lobell, D.B., Delire, C. & Mirin, A. (2007) Combined climate and carbon-cycle effects of large-scale deforestation. *Proceedings of the National Academy of Sciences USA*, **104**, 6550–6555.
- Balvanera, P., Pfisterer, A.B., Buchmann, N., He, J.S., Nakashizuka, T., Raffaelli, D. & Schmid, B. (2006) Quantifying the evidence for biodiversity effects on ecosystem functioning and services. *Ecology Letters*, **9**, 1146–1156.
- Barthlott, W., Mutke, J., Rafiqpoor, M.D., Kier, G. & Krefth, H. (2005) Global centres of vascular plant diversity. *Nova Acta Leopoldina*, **92**, 61–83.
- BEF-China Research Group (2013) *BEF data portal code repository*. *BEF-China data portal*. Retrieved January 31, 2013, from <https://github.com/befdata/>
- Bell, T., Newman, J.A., Silverman, B.W., Turner, S.L. & Lilley, A.K. (2005) The contribution of species richness and composition to bacterial services. *Nature*, **436**, 1157–1160.
- Bell, T., Lilley, A.K., Hector, A., Schmid, B., King, L. & Newman, J.A. (2009) A linear model method for biodiversity–ecosystem functioning experiments. *American Naturalist*, **174**, 836–849.
- Both, S., Fang, T., Baruffol, M., Schmid, B., Bruelheide, H. & Erfmeier, A. (2012) Effects of tree sapling diversity and nutrient addition on herb-layer invasibility in young experimental communities of subtropical species. *Open Journal of Ecology*, **2**, 1–12.
- Bruelheide, H., Böhnke, M., Both, S., Fang, T., Assmann, T., Baruffol, M. *et al.* (2011) Community assembly during secondary forest succession in a Chinese subtropical forest. *Ecological Monographs*, **81**, 25–41.
- Cardinale, B.J., Srivastava, D.S., Duffy, J.E., Wright, J.P., Downing, A.L., Sankaran, M. & Jouseau, C. (2006) Effects of biodiversity on the functioning of trophic groups and ecosystems. *Nature*, **443**, 989–992.
- Cardinale, B.J., Matulich, K.L., Hooper, D.U., Byrnes, J.E., Duffy, E., Gamfeldt, L., Balvanera, P., O'Connor, M.I. & Gonzalez, A. (2011) The functional role of producer diversity in ecosystems. *American Journal of Botany*, **98**, 572–592.

- Chisholm, R.A., Muller-Landau, H.C., Rahman, K.A., Bebbler, D.P., Bin, Y., Bohlman, S.A. *et al.* (2013) Scale-dependent relationships between tree species richness and ecosystem function in forests. *Journal of Ecology*, **101**, 1214–1224.
- Crutsinger, G.M., Collins, M.D., Fordyce, J.A., Gompert, Z., Nice, C.C. & Sanders, N.J. (2006) Plant genotypic diversity predicts community structure and governs an ecosystem process. *Science*, **313**, 966–968.
- Díaz, S., Symstad, A.J., Chapin, F.S., Wardle, D.A. & Huenneke, L.F. (2003) Functional diversity revealed by removal experiments. *Trends in Ecology and Evolution*, **18**, 140–146.
- Duffy, J.E. (2009) Why biodiversity is important to the functioning of real-world ecosystems. *Frontiers in Ecology and the Environment*, **7**, 437–444.
- Durieux, L., Machado, L.A.T. & Laurent, H. (2003) The impact of deforestation on cloud cover over the Amazon arc of deforestation. *Remote Sensing of Environment*, **86**, 132–140.
- Dutilleul, P. (1993) Spatial heterogeneity and the design of ecological field experiments. *Ecology*, **74**, 1646–1658.
- Fegraus, E.H., Andelman, S., Jones, M.B. & Schildhauer, M. (2005) Maximizing the value of ecological data with structured metadata: an introduction to Ecological Metadata Language (EML) and principles for metadata creation. *Bulletin of the Ecological Society of America*, **86**, 158–168.
- von Felten, S., Hector, A., Buchmann, N., Niklaus, P.A., Schmid, B. & Scherer-Lorenzen, M. (2009) Belowground nitrogen partitioning in experimental grassland plant communities of varying species richness. *Ecology*, **90**, 1389–1399.
- Gamfeldt, L., Snäll, T., Bagchi, R., Jonsson, M., Gustafsson, L., Kjellander, P. *et al.* (2013) Higher levels of multiple ecosystem services are found in forests with more tree species. *Nature Communications*, **4**, 1340.
- Geißler, C., Nadrowski, K., Kühn, P., Baruffol, M., Bruelheide, H., Schmid, B. & Scholten, T. (2013) Kinetic energy of throughfall in subtropical forests of SE China – effects of tree canopy structure, functional traits, and biodiversity. *PLoS One*, **8**, e49618.
- Givnish, T.J. (1994) Does diversity beget stability? *Nature*, **371**, 113–114.
- Grime, J.P. (1997) Biodiversity and ecosystem function: the debate deepens. *Science*, **277**, 1260–1261.
- Grime, J.P. (2002) Declining plant diversity: empty niches or functional shifts? *Journal of Vegetation Science*, **13**, 457–460.
- Hantsch, L., Braun, U., Scherer-Lorenzen, M. & Bruelheide, H. (2013) Tree species richness and tree species identity effects on occurrence of foliar fungal pathogens in a tree diversity experiment. *Ecosphere*, **4**, 81.
- Healy, C., Gotelli, N.J. & Potvin, C. (2008) Partitioning the effects of biodiversity and environmental heterogeneity for productivity and mortality in a tropical tree plantation. *Journal of Ecology*, **96**, 903–913.
- Hector, A. & Bagchi, R. (2007) Biodiversity and ecosystem multifunctionality. *Nature*, **448**, 188–190.
- Hector, A., von Felten, S. & Schmid, B. (2010) Analysis of variance with unbalanced data: an update for ecology & evolution. *Journal of Animal Ecology*, **79**, 308–316.
- 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. *Ecology Letters*, **5**, 502–511.
- Hector, A., Philipson, C., Saner, P., Chamagne, J., Dzulkiifi, D., O'Brien, M. *et al.* (2011) The Sabah Biodiversity Experiment: a long-term test of the role of tree diversity in restoring tropical forest structure and functioning. *Philosophical Transactions of the Royal Society B*, **366**, 3303–3315.
- Heidorn, P.B. (2008) Shedding light on the dark data in the long tail of science. *Library Trends*, **57**, 280–299.
- Henry, M. (2011) The constraints of selecting for insect resistance in plantation trees. *Agricultural and Forest Entomology*, **13**, 111–120.
- Hooper, D.U., Chapin, F.S. III, Ewel, J.J., Hector, A., Inchausti, P., Lavorel, S. *et al.* (2005) Effects of biodiversity on ecosystem functioning: a consensus of current knowledge. *Ecological Monographs*, **75**, 3–35.
- Hurlbert, S.H. (1984) Pseudoreplication and the design of ecological field experiments. *Ecological Monographs*, **54**, 187–211.
- Huston, M.A. (1997) Hidden treatments in ecological experiments: re-evaluating the ecosystem function of biodiversity. *Oecologia*, **110**, 449–460.
- Isbell, F., Calcagno, V., Hector, A., Connolly, J., Harpole, W.S., Reich, P.B. *et al.* (2011) High plant diversity is needed to maintain ecosystem services. *Nature*, **477**, 199–202.
- Kröber, W., Böhnke, M., Welk, E., Wirth, C. & Bruelheide, H. (2012) Leaf trait-environment relationships in a subtropical broadleaved forest in South-East China. *PLoS One*, **7**, e35742.
- Lamont, B.B. (1995) Testing the effect of ecosystem composition structure on its functioning. *Oikos*, **74**, 283–295.
- Lang, A.C., Härdtle, W., Baruffol, M., Böhnke, M., Bruelheide, H., Schmid, B., von Wehrden, H. & von Oheimb, G. (2012) Mechanisms promoting tree species coexistence: experimental evidence with saplings of subtropical forest ecosystems of China. *Journal of Vegetation Science*, **23**, 837–846.
- Le Roux, X., Schmid, B., Barnard, R.L., Clays-Josserand, A., Falcao-Salles, J., Guillaumaud, N. *et al.* (2013) Plant functional diversity rather than species richness drives soil functions in experimental grasslands. *PLoS One*, **8**, e61069.
- Leadley, P.W. & Körner, C. (1996) Effects of elevated CO<sub>2</sub> on plant species dominance in a highly diverse calcareous grassland. *Carbon Dioxide, Populations, and Communities* (eds C. Körner & F.A. Bazzaz), pp. 159–175. Academic Press, San Diego, CA, USA.
- Lepš, J. (2004) What do the biodiversity experiments tell us about consequences of plant species loss in the real world? *Basic and Applied Ecology*, **5**, 529–534.
- Lipowsky, A., Schmid, B. & Roscher, C. (2011) Selection for monoculture and mixture genotypes in a biodiversity experiment. *Basic and Applied Ecology*, **12**, 360–371.
- Loreau, M. & Hector, A. (2001) Partitioning selection and complementarity in biodiversity experiments. *Nature*, **412**, 72–76.
- Loreau, M., Naeem, S. & Inchausti, P. (2002) *Biodiversity and Ecosystem Functioning*. Oxford University Press, New York.
- MacArthur, R. & Levins, R. (1967) The limiting similarity, convergence and divergence of coexisting species. *American Naturalist*, **101**, 377–385.
- Marquard, E., Weigelt, A., Roscher, C., Gubsch, M., Lipowsky, A. & Schmid, B. (2009a) Positive biodiversity–productivity relationship due to increased plant density. *Journal of Ecology*, **97**, 696–704.
- Marquard, E., Weigelt, A., Temperton, V.M., Roscher, C., Schumacher, J., Buchmann, N., Fischer, M., Weisser, W.W. & Schmid, B. (2009b) Plant species richness and functional composition drive overyielding in a 6-year grassland experiment. *Ecology*, **90**, 3290–3302.
- Michener, W.K. & Jones, M.B. (2012) Ecoinformatics: supporting ecology as a data-intensive science. *Trends in Ecology and Evolution*, **27**, 85–93.
- Mikola, J., Salonen, V. & Setälä, H. (2002) Studying the effects of plant species richness on ecosystem functioning: does the choice of experimental design matter? *Oecologia*, **133**, 594–598.
- Moreira, X. & Mooney, K.A. (2013) Influence of plant genetic diversity on interactions between higher trophic levels. *Biology Letters*, **9**, 20130133.
- Nadrowski, K., Wirth, C. & Scherer-Lorenzen, M. (2010) Is forest diversity driving ecosystem function and service? *Current Opinion in Environmental Sustainability*, **2**, 1–5.
- Nadrowski, K., Ratcliffe, S., Bönisch, G., Bruelheide, H., Kattge, J., Liu, X.J. *et al.* (2013) Harmonizing, annotating, and sharing data in biodiversity-ecosystem functioning research. *Methods in Ecology and Evolution*, **4**, 201–205.
- Naeem, S., Thompson, L.J., Lawler, S.P., Lawton, J.H. & Woodfin, R.M. (1994) Declining biodiversity can alter the performance of ecosystems. *Nature*, **368**, 734–737.
- Niklaus, P.A., Leadley, P.W., Schmid, B. & Körner, C. (2001) A long-term study on biodiversity elevated CO<sub>2</sub> interactions in grassland. *Ecological Monographs*, **71**, 341–356.
- Nyland, R.D. (2002) *Silviculture: Concepts and Applications*, 2nd edn. McGraw-Hill, Boston.
- von Oheimb, G., Lang, A.C., Bruelheide, H., Forrester, D.I., Wäsche, I., Yu, M.J. & Härdtle, W. (2011) Individual-tree radial growth in a species-rich subtropical forest: the role of local neighbourhood competition. *Forest Ecology and Management*, **261**, 499–507.
- Osnas, J.L.D., Lichstein, J.W., Reich, P.B. & Pacala, S.W. (2013) Global leaf trait relationships: mass, area, and the leaf economics spectrum. *Science*, **340**, 741–744.
- Outbor, N.J. (1993) Isolation, population size and extinction: the classical and metapopulation approaches applied to vascular plants along the Dutch Rhine-system. *Oikos*, **66**, 298–308.
- Pacala, S.W. & Deutschman, D.H. (1995) Details that matter – the spatial distribution of individual trees maintains forest ecosystem function. *Oikos*, **74**, 357–365.
- Pacala, S.W. & Tilman, D. (1994) Limiting similarity in mechanistic and spatial models of plant competition in heterogeneous environments. *American Naturalist*, **143**, 222–257.
- Pasari, J.R., Levi, T., Zavaleta, E.S. & Tilman, D. (2013) Several scales of biodiversity affect ecosystem multifunctionality. *Proceedings of the National Academy of Sciences USA*, **119**, 10219–10222.
- Paul, G.S., Montagnini, F., Berlyn, G.P., Craven, D.J., Breugel, M. & Hall, J.S. (2012) Foliar herbivory and leaf traits of five native tree species in a young plantation of Central Panama. *New Forests*, **43**, 69–87.
- Petermann, J.S., Fergus, A.J., Turnbull, L.A. & Schmid, B. (2008) Janzen-Connell effects are both widespread and strong enough to maintain functional diversity in grasslands. *Ecology*, **89**, 2399–2406.
- Potvin, C. & Dutilleul, P. (2009) Neighborhood effects and size-asymmetric competition in a tree plantation varying in diversity. *Ecology*, **90**, 321–327.



- Potvin, C. & Gotelli, N. (2008) Biodiversity enhances individual performance but does not affect survivorship in tropical trees. *Ecology Letters*, **11**, 217–223.
- Pretzsch, H. & Schütze, G. (2009) Transgressive overyielding in mixed compared with pure stands of Norway spruce and European beech in Central Europe: evidence on stand level and explanation on individual tree level. *European Journal of Forest Research*, **128**, 183–204.
- Quijas, S., Jackson, L.E., Maass, M., Raffaelli, D., Schmid, B. & Balvanera, P. (2012) Plant diversity and generation of ecosystem services at the landscape scale: expert knowledge assessment. *Journal of Applied Ecology*, **49**, 929–940.
- Reich, P.B., Walters, M.B., Kloeppel, B.D. & Ellsworth, D.S. (1995) Different photosynthesis-nitrogen relations in deciduous hardwood and evergreen coniferous tree species. *Oecologia*, **104**, 24–30.
- Reich, P.B., Tilman, D., Isbell, F., Mueller, K., Hobbie, S.E., Flynn, D.F.B. & Eisenhauer, N. (2012) Impacts of biodiversity loss escalate through time as redundancy fades. *Science*, **336**, 589–592.
- Roscher, C., Schumacher, J., Baade, J., Wilcke, W., Gleixner, G., Weisser, W.W., Schmid, B. & Schulze, E.-D. (2004) The role of biodiversity for element cycling and trophic interactions: an experimental approach in a grassland community. *Basic and Applied Ecology*, **5**, 107–121.
- Roscher, C., Temperton, V.M., Scherer-Lorenzen, M., Schmitz, M., Schumacher, J., Schmid, B., Buchmann, N., Weisser, W.W. & Schulze, E.D. (2005) Overyielding in experimental grassland communities – irrespective of species pool or spatial scale. *Ecology Letters*, **8**, 419–429.
- Saha, S., Kohnle, U., Brang, P., Ehring, A., Geisel, J., Leder, B. *et al.* (2012) Growth and quality of young oaks (*Quercus robur* and *Q. petraea*) grown in cluster plantings in Central Europe: a weighted meta-analysis. *Forest Ecology and Management*, **283**, 106–118.
- Salles, J.F., Poly, F., Schmid, B. & Le Roux, X. (2009) Community niche predicts the functioning of denitrifying bacterial assemblages. *Ecology*, **90**, 3324–3332.
- Scherer-Lorenzen, M. (2013) The functional role of biodiversity in the context of global change. *Forests and Global Change* (eds D.A. Coomes, D.F.R.P. Burslem & W.D. Simonson), pp. 195–237. Cambridge University Press, Cambridge.
- Scherer-Lorenzen, M., Körner, C. & Schulze, E.-D. (2005b) The functional significance of forest diversity: a synthesis. *Forest Diversity and Function: Temperate and Boreal Systems* (eds M. Scherer-Lorenzen, C. Körner & E.D. Schulze), pp. 377–389. Ecological Studies, 176. Springer, Berlin.
- Scherer-Lorenzen, M., Potvin, C., Koricheva, J., Schmid, B., Hector, A., Bornik, Z., Reynolds, G. & Schulze, E.-D. (2005a) The design of experimental tree plantations for functional biodiversity research. *Forest Diversity and Function: Temperate and Boreal Systems* (eds M. Scherer-Lorenzen, C. Körner & E.D. Schulze), pp. 347–376. Ecological Studies, 176. Springer, Berlin.
- Scherer-Lorenzen, M., Schulze, E.D., Don, A., Schumacher, A. & Weller, E. (2007) Exploring the functional significance of forest diversity: a new long-term experiment with temperate tree species (BIOTREE). *Perspectives in Plant Ecology Evolution and Systematics*, **9**, 53–70.
- Schläpfer, F., Pfisterer, A. & Schmid, B. (2005) Non-random species extinction and plant production: implications for ecosystem functioning. *Journal of Applied Ecology*, **42**, 13–24.
- Schmid, B. & Hector, A. (2004) The value of biodiversity experiments. *Basic and Applied Ecology*, **5**, 535–542.
- Schmid, B., Hector, A., Huston, M.A., Inchausti, P., Nijs, I., Leadley, P.W. & Tilman, D. (2002) The design and analysis of biodiversity experiments. *Biodiversity and Ecosystem Functioning: Synthesis and Perspectives* (eds M. Loreau, S. Naeem & P. Inchausti), pp. 61–75. Oxford University Press, Oxford.
- Schmid, B., Hector, A., Saha, P. & Loreau, M. (2008) Biodiversity effects and transgressive overyielding. *Journal of Plant Ecology*, **1**, 95–102.
- Schmid, B., Balvanera, P., Cardinale, B.J., Godbold, J., Pfisterer, A.B., Raffaelli, D., Solan, M. & Srivastava, D.S. (2009) Consequences of species loss for ecosystem functioning: meta-analyses of data from biodiversity experiments. *Biodiversity, Ecosystem Functioning, and Human Wellbeing* (eds S. Naeem, D.E. Bunker, A. Hector, M. Loreau & C. Perrings), pp. 14–29. Oxford University Press, Oxford.
- Schnitzer, S.A., Klironomos, J.N., HilleRisLambers, J., Kinkel, L.L., Reich, P.B., Xiao, K. *et al.* (2011) Soil microbes drive the classic plant diversity-productivity pattern. *Ecology*, **92**, 296–303.
- Schuldt, A., Bruelheide, H., Durka, W., Eichenberg, D., Fischer, M., Kröber, W. *et al.* (2012) Plant traits affecting herbivory on tree recruits in highly diverse subtropical forests. *Ecology Letters*, **15**, 732–739.
- Solan, M., Cardinale, B.J., Downing, A.L., Engelhardt, K.A.M., Ruesink, J.L. & Srivastava, D.S. (2004) Extinction and ecosystem function in the marine benthos. *Science*, **306**, 1177–1180.
- Spehn, E.M., Scherer-Lorenzen, M., Schmid, B., Hector, A., Caldeira, M.C., Dimitrakopoulos, P.G. *et al.* (2002) The role of legumes as a component of biodiversity in a cross-European study of grassland biomass nitrogen. *Oikos*, **98**, 205–218.
- Stoll, P. & Prati, D. (2001) Intraspecific aggregation alters competitive interactions in experimental plant communities. *Ecology*, **82**, 319–327.
- Tilman, D., Wedin, D. & Knops, J. (1996) Productivity and sustainability influenced by biodiversity in grassland ecosystems. *Nature*, **379**, 718–720.
- Vehviläinen, H. & Koricheva, J. (2006) Moose and vole browsing patterns in experimentally assembled pure and mixed forest stands. *Ecography*, **29**, 497–506.
- Verheyen, K. & Scherer-Lorenzen, M. (2012) Tree tessellation. *International Innovation. Environment*, **31**.
- Visser, S. (1995) Ectomycorrhizal fungal succession in Jack pine stands following wildfire. *New Phytologist*, **129**, 389–401.
- Wacker, L., Baudois, O., Eichenberger-Glinz, S. & Schmid, B. (2008) Environmental heterogeneity increases complementarity in experimental grassland communities. *Basic and Applied Ecology*, **9**, 467–474.
- Wacker, L., Baudois, O., Eichenberger-Glinz, S. & Schmid, B. (2009) Diversity effects in early- and mid-successional species pools along a nitrogen gradient. *Ecology*, **90**, 637–648.
- Wang, X.H., Kent, M. & Fang, X.F. (2007) Evergreen broad-leaved forest in Eastern China: its ecology and conservation and the importance of resprouting in forest restoration. *Forest Ecology and Management*, **245**, 76–87.
- Wang, K., Shi, X.Z., Yu, D.S., Shi, D.M., Chen, J.M., Xu, B.B., Liang, Y. & Li, D.C. (2005) Environmental factors affecting temporal and spatial dynamics of soil erosion in Xhinguo county South China. *Pedosphere*, **15**, 620–627.
- Watt, A.D., Hunter, M. & Stork, N.E. (eds) (1997) *Forests and Insects*. Chapman and Hall, London.
- Weigelt, A., Weisser, W.W., Buchmann, N. & Scherer-Lorenzen, M. (2009) Biodiversity and multifunctional agriculture: equal productivity in high-diversity low-input and low-diversity high-input grasslands. *Biogeosciences*, **6**, 1695–1706.
- Weiber, E. & Keddy, P. (1999) *Ecological Assembly Rules. Perspectives, Advances, Retreats*. Cambridge University Press, Cambridge.
- Weiner, J., Andersen, S.B., Wille, W.K.M., Griepentrog, H.W. & Olsen, J.M. (2010) Evolutionary Agroecology: the potential for cooperative, high density, weed-suppressing cereals. *Evolutionary Applications*, **3**, 473–479.
- Whitham, T.G., Young, W.P., Martinsen, G.D., Gehring, C.A., Schweitzer, J.A., Shuster, S.M. *et al.* (2003) Community and ecosystem genetics: a consequence of the extended phenotype. *Ecology*, **84**, 559–573.
- de Wit, C.T. (1960) On competition. *Verslagen Landbouwkundige Onderzoekingen*, **66**, 1–82.
- Worm, B., Barbier, E.B., Beaumont, N., Duffy, J.E., Folke, C., Halpern, B.S. *et al.* (2006) Impacts of biodiversity loss on ocean ecosystem services. *Science*, **314**, 787–790.
- Wright, I.J., Reich, P.B., Westoby, M., Ackerly, D.D., Baruch, Z., Bongers, F. *et al.* (2004) The worldwide leaf economics spectrum. *Nature*, **428**, 821–827.
- Wu, Y.T., Gutknecht, J., Nadrowski, K., Geißler, C., Kühn, P., Scholten, T. *et al.* (2012) Relationships between soil microorganisms, plant communities and soil characteristics in Chinese subtropical forests. *Ecosystems*, **15**, 624–636.
- Yang, X.F., Bauhus, J., Both, S., Fang, T., Härdtle, W., Kröber, W. *et al.* (2013) Establishment success in a forest biodiversity and ecosystem functioning experiment in subtropical China (BEF-China). *European Journal of Forest Research*, **132**, 593–606.
- Zavaleta, E.S. & Hulvey, K.B. (2004) Realistic species losses disproportionately reduce grassland resistance to biological invaders. *Science*, **306**, 1175–1177.
- Zavaleta, E.S., Pasari, J.R., Hulvey, K.B. & Tilman, G.D. (2010) Sustaining multiple ecosystem functions in grassland communities requires higher biodiversity. *Proceedings of the National Academy of Sciences USA*, **107**, 1443–1446.

Received 12 June 2013; accepted 7 October 2013

Handling Editor: Helene Muller-Landau

## Supporting Information

Additional Supporting Information may be found in the online version of this article.

**Table A1.** List of species used in the BEF-China experiment according to the Flora of China (<http://www.efloras.org> and <http://frps.eflora.cn>). The site column shows the experimental site (A, B) where the species was planted. In addition, for broad-leaved trees, the number after the site shows the species set that refers to the three random extinction scenarios each at the experimental site A and B