

Increasing plant species richness by seeding has marginal effects on ecosystem functioning in agricultural grasslands

Martin Freitag¹, Norbert Hölzel¹, Lena Neuenkamp², Fons van der Plas³, Peter Manning⁴, Anna Abrahão⁵, Joana Bergmann⁶, Runa Boeddinghaus⁵, Ralph Bolliger⁷, Ute Hamer⁸, Ellen Kandeler⁵, Till Kleinebecker^{9,10}, Klaus-Holger Knorr¹¹, Sven Marhan⁵, Margot Neyret⁴, Daniel Prati⁷, Gaëtane Le Provost⁴, Hugo Saiz⁷, Mark van Kleunen^{12,13}, Deborah Schäfer¹⁴, Valentin H. Klaus^{15,16}

¹ Biodiversity and Ecosystem Research Group, Institute of Landscape Ecology, University of Münster, Münster, Germany

² Department of Ecology & Multidisciplinary Institute for Environment Studies “Ramon Margalef,” University of Alicante, Alicante, Spain

³ Plant Ecology and Nature Conservation Group, Wageningen University & Research, Wageningen, The Netherlands

⁴ Senckenberg Biodiversity and Climate Research Centre (SBIK-F), Senckenberg Gesellschaft für Naturforschung, Frankfurt, Germany

⁵ Institute of Soil Science and Land Evaluation, University of Hohenheim, Stuttgart, Germany

⁶ Sustainable Grassland Systems, Leibniz Centre for Agricultural Landscape Research (ZALF), Paulinenau 14641, Germany

⁷ Institute of Plant Sciences, University of Bern, Bern, Switzerland

⁸ Soil Ecology and Land Use, Institute of Landscape Ecology, University of Münster, Münster, Germany

21 ⁹ Institute of Landscape Ecology and Resource Management, Justus-Liebig-University Gießen, Gießen,
22 Germany

23 ¹⁰ Centre for International Development and Environmental Research (ZEU), Justus Liebig University
24 Giessen, Senckenbergstrasse 3, 35390 Giessen, Germany

25 ¹¹ Ecohydrology and Biogeochemistry, Institute of Landscape Ecology, University of Münster, Münster,
26 Germany

27 ¹² Ecology, Department of Biology, University of Konstanz, Universitätsstrasse 10, D-78457 Konstanz,
28 Germany

29 ¹³ Zhejiang Provincial Key Laboratory of Plant Evolutionary Ecology and Conservation, Taizhou University,
30 Taizhou 318000, China

31 ¹⁴ Botanical Garden of the University of Bern, University of Bern, Bern, Switzerland

32 ¹⁵ Institute of Agricultural Sciences, ETH Zürich, Zürich, Switzerland

33 ¹⁶ Forage Production and Grassland Systems, Agroscope, Zürich, Switzerland

34 **Keywords:** biodiversity-ecosystem functioning experiments, complementarity, plant functional traits,
35 land-use intensity gradient, restoration, seed addition, selection effect, species pool effect

36 **Article type:** Research article

37 **Word count** (all text, references, tables and figure legends): 8409 words

38 **Corresponding author:** Norbert Hölzel, Institute of Landscape Ecology, University of Münster,
39 Heisenbergstr. 2, D-48149 Münster, Germany; tel: +49 251 83-33994, fax: +49 251 83-38338, email:
40 nhoelzel@wwu.de

41 **Statement of authorship**

42 NH, UH, TK, VHK and DP designed the experiment and VHK, MF, NH, FvdP and PM conceived ideas for this
43 study. AA, RB, MF, NH, UH, MvK, EK, TK, KHK, VHK, PM, SM, SN, DP and GLP contributed data. MF analysed
44 the data and MF prepared a first draft of the manuscript with substantial input from NH. All authors
45 contributed to revising the manuscript.

46 **Data availability statement**

47 All data and code used for this study will be made available via the Biodiversity Exploratories Information
48 System BEXIS (<https://bexis.uni-jena.de/PublicData>).

49

Abstract

1. Experimental evidence shows that grassland plant diversity enhances ecosystem functioning. Yet, the transfer of results from strongly controlled biodiversity experiments to naturally assembled 'real world' ecosystems remains challenging, inhibiting clear implication for ecological restoration programs.
2. We address this issue by experimentally sowing locally absent plant species in 73 agricultural grasslands along a land-use intensity gradient, to test how ecosystem functions related to productivity and nutrient cycling respond to species enrichment.
3. To create a diversity gradient within sites a fully factorial combination of sowing and sward disturbance was chosen. Compared with the control, this resulted in an average increase of species-richness of 8.8 species in the combined treatment while sowing and sward disturbance alone led only to an average increase by 2.8 and 0.9 species, respectively. In our analysis disturbance was used as a second control treatment when assessing subsequent effects on ecosystem functioning.
4. We found that only one of 12 ecosystem functions responded to changes in species richness. In fact, ecosystem functioning was rather driven by environmental conditions and land-use intensity. This suggests that the functionally-relevant niche space was saturated in naturally assembled grasslands, and that competitive, high-functioning species were already present.
5. *Synthesis* Despite the fact that plant diversity was successfully increased, changes in ecosystem functioning were generally minor. While nature conservation and cultural ecosystem services certainly benefit from plant species enrichment, our study indicates that such an enrichment may deliver only weak increases in ecosystem functioning in both moderately intensive and traditionally managed agricultural grasslands.

INTRODUCTION

Much of the theory and knowledge on BEF relationships comes from controlled experiments, in which the functioning of randomly assembled communities of varying initial or “sown” species richness is assessed (e.g. Roscher et al., 2004; Tilman et al., 1996). Such diversity manipulations have generally shown that initially diverse assemblages promote ecosystem functioning, with positive but saturating BEF relationships being consistently observed (Cardinale et al., 2012; Hooper et al., 2012; Isbell et al., 2015). However, the transferability of findings from controlled experiments to agriculturally managed grassland ecosystems, which is nowadays increasingly used to advocate biodiversity restoration, remains challenging for two main reasons (Klaus et al., 2020; Manning et al., 2019).

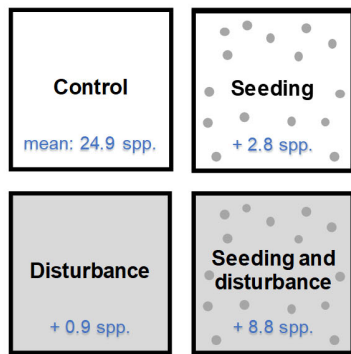
First, while BEF experiments isolate biodiversity effects by controlling for other drivers, they often neglect the role of factors which covary with biodiversity loss, including human-induced environmental change as a driver of both species-loss and functioning (De Laender et al., 2016; Srivastava & Vellend, 2005). BEF experiments are also criticized for anticipating random species loss (Lepš, 2004; Wardle, 2016), while in real-world-communities species loss or gain is non-random and species with certain functional traits are more likely to go extinct or increase in abundance (Pywell et al., 2003; Saar et al., 2012). Both simulations and experiments underpin this criticism, showing that productivity is largely unaffected by species loss if the least competitive and least productive species are most extinction prone (Gross & Cardinale, 2005; Smith & Knapp, 2003). Thus, not only biodiversity *per se* (i.e. how many species or functional types occur within a community), but also functional community composition (i.e. which species are present and/or dominant within a community) as well as environmental conditions and land use drive ecosystem functioning (Allan et al., 2015; Diaz et al., 2007; Leibold et al., 2017; van der Plas, 2019). In addition, the effect of plant diversity on ecosystem functioning can be overruled by external inputs such as intensive fertilization (Nyfeler et al., 2009), which are likely to weaken BEF relationships at high management intensity.

A second complicating factor in BEF relationships in naturally assembled communities is the distinction between initial and realized biodiversity (Hagan et al., 2021). In BEF theory, complementarity and selection effects are considered as the main mechanisms underlying the positive relationship between initially sown species-richness and ecosystem functioning of classical BEF experiments (Loreau & Hector, 2001). Complementarity effects occur when differences or interactions between co-existing species lead to positive effects on ecosystem functioning, e.g. through resource partitioning, facilitation or reduced frequency dependence (Loreau & Hector, 2001). They are considered to increase in strength the more the coexisting species differ in their traits (Barry et al., 2019). Selection effects occur when one, or several, competitive and high functioning species become dominant in a local community (Cardinale et al., 2012). Selection effects are greatest when one or few functionally important species of the initial species pool are able to outcompete all others, with diverse sown species pools more likely to contain these species (Loreau & Hector, 2001). Importantly, both theory and experimental evidence show that complementarity and selection effects (and hence overall effects of biodiversity on ecosystem functioning) are more strongly related to initial biodiversity, than realized biodiversity. For example, realized species richness can be low where selection effects are strong and only few dominant species persist from an initially diverse species pool (Cardinale et al., 2011). While it is easy to distinguish between initial and realized biodiversity in experimentally sown plant communities, this is not the case for purely observational studies, where one can only directly quantify realized biodiversity (Hagan et al., 2021).

Seed addition experiments are ideal to study BEF relationships while acknowledging the effect of other covariates on both biodiversity and ecosystem functioning and allow to distinguish between the effects of realized and initial biodiversity in realistic settings (Bannar-Martin et al., 2018; Manning et al., 2019). Furthermore, seed addition experiments are of high applied value, as a key factor limiting successful grassland biodiversity restoration is the low biodiversity of local species pools (Myers & Harms, 2009), which may prevent high-functioning species, from colonising a site (Hagan et al., 2021; Leibold et al.,

2017). A recent synthesis, however, contradicts the assumed generally positive effects of seed addition on aboveground biomass production (Ladouceur et al., 2020). The results of this meta-analysis indicate a context-dependency of BEF relationships on environmental conditions such as the degree of ecosystem degradation (i.e. mature grassland vs. ex-arable field), but the reasons for this context-dependency remain speculative, due to a lack of detailed information about site differences and species pools.

(a) Experimental design



x 73 sites

across gradients of
species richness, land-use
intensity and abiotic conditions

(b) Hypothetical among- and within-site relationships

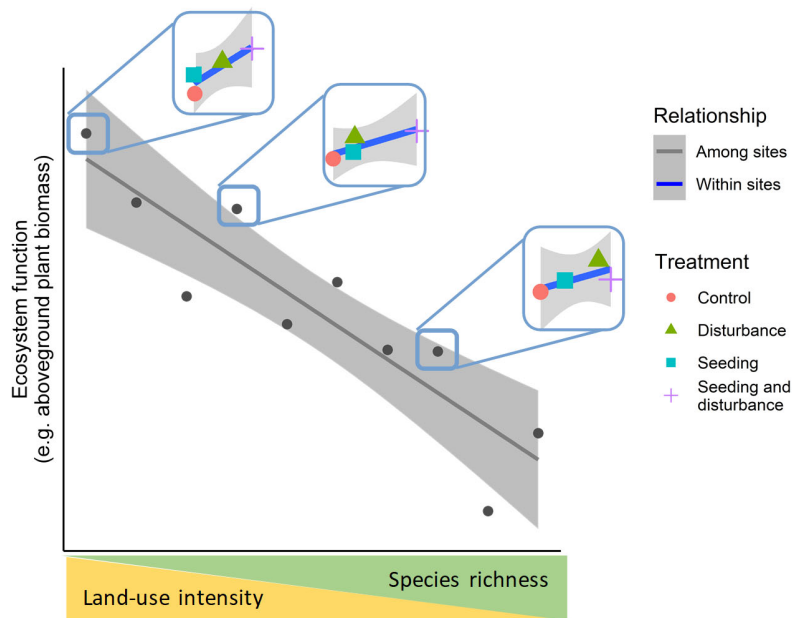


Figure 1: Conceptual graph visualizing (a) the experimental design with two-factorial disturbance and seed addition treatments in 73 grasslands (blue font reveals control plant richness and average increases in richness per treatment after four years) and (b) how negative among-site correlations of species richness with environmental drivers could mask potentially positive within site relationships between species richness and ecosystem function. We performed a seed addition experiment to create a gradient in species richness within grassland sites, where variation in species richness at least partially decoupled from variation in land-use intensity, which increases some ecosystem functions but reduces biodiversity

(Allan et al., 2015). In such a design, the analysis of *within site* variation in species richness and ecosystem function could reveal positive BEF relationships within sites, although these relationships might be negative among sites. In addition, *within site* effects of species richness on ecosystem functioning are likely to change along the land-use intensity gradient due to the replacement of richness effects by external inputs such as fertilisation at high level (Nyfeler et al., 2009).

To address the BEF knowledge gaps identified above, we performed a seed addition experiment in 73 agricultural grasslands. This allowed us to disentangle the effects of seeding-enhanced species richness on ecosystem functioning from the correlation between observed species richness and grassland ecosystem functioning across a wide range of site conditions and land use. The experiment had a two-factor design in which we crossed high-diversity seeding of native species and topsoil disturbance treatments (to facilitate colonization; Freitag et al., 2021) and measured twelve ecosystem functions related to aboveground primary productivity and nutrient cycling after 4 years (see Table 1 for the list of ecosystem functions). The seeding treatments established a species richness gradient within grassland sites (Figure. 1): After four years runtime of the experiment species richness increased on average by 2.8 species per 4 m² (10%) on the seeding-only subplots and by 8.8 species per 4 m² (35%) on the seeding-and-disturbance subplots compared to the control (Freitag et al., 2021). Moreover, there was an overall 10% increase in SPIE (effective number of equally abundant species; Chase et al., 2018) in the seeding-and-disturbance subplots, indicating that the newly sown species became essential constituents of the plant communities. Establishment success of the sown plants varied with aboveground plant biomass on site and functional traits of the sown species (Freitag et al., 2021), thus allowing us to test the effects of species richness on ecosystem functioning under a realistic community assembly scenario. First, we investigated the responses of ecosystem functions to the seeding treatments (irrespective of the changes in realized species richness), in order to address the effect of sown *initial* diversity on ecosystem functioning (Hagan et al., 2021). Second, we related ecosystem functions to *realized* species richness and

functional composition *within* each grassland site, to test how experimentally induced changes in species richness and composition translate into changes in ecosystem functions at the site level. We expected that the gain of species and functions would be highest in initially species-poor grasslands. Third, we used path analyses to assess the relative importance of species richness, functional composition, and environmental drivers (abiotic conditions and land use) for ecosystem functioning *among* grassland sites, i.e., at the landscape scale. We anticipated positive relations between species richness and ecosystem functioning *within* sites (Hypothesis I), which should be stronger for sown initial richness than for realized richness (Hypothesis II), while we expected that positive BEF relations *among* sites would be overruled by variation in land-use intensity and abiotic site conditions (Hypothesis III).

MATERIALS AND METHODS

Study design

We performed a plant-diversity-restoration experiment in 73 grasslands as part of the large-scale and long-term Biodiversity Exploratories project (Fischer et al., 2010). We installed the experiment in grasslands of three German regions (Biosphere Reserve Schwäbische Alb, 48.4°N, 9.4°E; National Park Hainich-Dün and surroundings, 51.1°N, 10.4°E; Biosphere Reserve Schorfheide-Chorin, 53.0°N, 14.0°E) that span gradients in soil characteristics, elevation and climate that are representative of large parts of Central Europe. Within each region, we selected 23-25 agricultural grasslands along gradients of land-use intensity (LUI). Farmers were interviewed annually between 2014 and 2018 to establish the amount of fertilizer used, mowing frequency and grazing intensity per grassland site (Vogt et al., 2019). We averaged fertilization ($\text{kg N ha}^{-1} \text{ yr}^{-1}$), mowing (no. of cuts yr^{-1}) and grazing (livestock units * grazing days ha^{-1}) intensities across years and standardised to their global means before calculating a LUI index according to Blüthgen et al. (2012) using the LUI tool (Ostrowski et al., 2020) in BExIS

(<http://doi.org/10.17616/R32P9Q>). A low LUI corresponds to unfertilized sheep pastures or annually-cut meadows, while a high LUI indicates highly fertilized and frequently disturbed meadows and pastures.

Within each of the 73 grasslands, we established a 2×2 factorial experiment in which high-diversity seeding and topsoil disturbance were the treatments, in autumn 2014. On each of the four $7 \text{ m} \times 7 \text{ m}$ subplots (control, seeding-only, seeding-and-disturbance, disturbance-only), we measured multiple ecosystem functions ~4 years after the experiment was established (see Appendix Figure S1 for the experimental design). We applied the disturbance by tilling the topsoil down to 10 cm with a rotary harrow or rotary cultivator in October 2014. Fragments of the former sward were left on the disturbance subplots so that plants could potentially regrow. For the seeding treatment, we compiled seed mixtures of commercially available regional seeds from common and less-common native species of the species pools of each region, including grasses (12 - 15 spp., depending on region), legumes (0 - 7 spp.) and forbs (35 - 44 spp.; Klaus et al., 2017). These mixtures were applied to the seeding subplots twice, in October 2014 and March 2015. Seeding density per species depended inversely on seed mass. In total, we sowed 5.37 g m^{-2} and 66 species in Schwäbische Alb, 4.11 g m^{-2} , 52 species in Hainich-Dün and 3.47 g m^{-2} and 47 species in Schorfheide (see Klaus et al., 2017 for further details).

We used soil pH and a topographic wetness index (TWI) to characterize the abiotic site properties of the grasslands, as these have been shown to influence local plant diversity (Le Provost et al. 2021) and ecosystem functioning (Allan et al. 2015). For pH, in May 2017, we took a composite sample of 14 soil cores of 10 cm depth in the direct vicinity of the experiment, dried the soil and measured pH in a 0.01 M CaCl_2 solution (1:2.5 soil:solution ratio). The TWI was calculated as $\log(a/\tan \beta)$ where a is the specific catchment area (cumulative upslope area which drains through a Digital Elevation Model (DEM, <http://www.bkg.bund.de>) cell, divided by per unit contour length) and $\tan \beta$ is the slope gradient in radians calculated over a local region surrounding the cell of interest (Gessler et al. 1995). TWI therefore

combines both the upslope contributing area (determining the amount of water received from upslope areas) and slope (determining the loss of water to downslope areas). We calculated TWI from raster DEM data with a cell size of 25 m using ArcGIS tools (flow direction and accumulation tools of the hydrology toolset and raster calculator) and averaged the TWI values of a 4×4 window centred on the plot, i.e. 16 DEM cells (100 m \times 100 m in total).

Plant diversity and functional composition

We surveyed plant diversity and functional composition in May 2018 and 2019, 4 years after the set-up of the experiment. We recorded all vascular plant species on 2 m \times 2 m quadrats and visually estimated the percentage cover of each species on each subplot. To characterize the functional plant community composition along resource-acquisition (specific leaf area, Reich, 2014) and mycorrhizal collaboration gradients (root diameter, Bergmann et al. 2020), we used species-specific means of specific-leaf area (SLA) from TRY (Kattge et al., 2019) and mean log-transformed measurements of fine root diameter from Lachaise et al., (2021). For both traits t , we calculated community-weighted means (CWMs) for each community c with S species as $CWM_{ct} = \sum_{i=1}^S p_{ci} \times t_i$, where p_{ci} is the relative abundance of species i in community c and t_i is the mean trait value of species i , using the R package *FD* (Laliberté & Legendre 2010) in R v4.0.5 (R Core Team 2021).

Ecosystem functions

Between 2018 and 2019, we measured twelve above- and below-ground ecosystem function indicators (hereafter ‘ecosystem functions’) related to productivity and nutrient cycling on all four subplots within sites (Table 1). See Appendix 2 for more details on ecosystem-function-indicator measurements.

Table 1: The twelve ecosystem function indicators (henceforth ecosystem functions) measured ~4 years after the set-up of the experiment. Most ecosystem functions were measured in all 73 grasslands (N=292 subplots), but sample sizes vary due to constraints in fieldwork logistics and lost samples. See Appendix 2 for more details on ecosystem-function-indicator measurements. BEF evidence was compiled from grassland studies and related experiments in similar, sown communities.

Ecosystem function	N	Measurement	Experimental vs. observational evidence
Aboveground plant biomass	292	Aboveground green biomass in g per m ²	Complementarity and selection effects increase productivity in species-rich communities (Cardinale et al. 2012); some (Bullock et al. 2007) but generally weaker evidence in observational studies (van der Plas, 2019)
Aboveground plant carbon-to-nitrogen (C:N) ratio	292	Measure of fodder quality, derived via near infrared reflectance spectroscopy	Neutral, no diversity effect on fodder quality (Schaub et al., 2020) but a positive effect of multi-species oversteering on plant biomass N (Savage et al., 2021); limited evidence for observational studies
Aboveground plant $\delta^{13}\text{C}$	284	Measure of water-use efficiency. Stable carbon isotope ratio of aboveground plant ^{13}C of community biomass	Weak evidence for increasing water-use efficiency (i.e. decreasing $\delta^{13}\text{C}$) with species richness (Bachmann et al., 2015; Guderle et al., 2018; Verheyen et al., 2008); also found in observational studies (Klaus et al., 2016)
Aboveground plant $\delta^{15}\text{N}$	284	Inverse measure of N-uptake efficiency. Stable nitrogen isotope ratio of aboveground plant ^{15}N of community biomass	Resource-partitioning in species-rich communities increases nitrogen-use efficiency (less losses of nitrogen, i.e. decreases in $\delta^{15}\text{N}$) (Gubsch et al., 2011; Mueller et al., 2013); also found in observational studies (Kleinebecker et al., 2014)
Soil N leaching risk	292	Measure of leaching risk of inorganic N from soil. Sum of NO_3^- and NH_4^+ concentrations at 20 cm depth below main rooting zone, aggregated over the growing season using ion-resin bags	Greater N uptake and nitrogen-use complementarity in diverse communities reduce subsoil nitrate concentrations (Mueller et al., 2013; Scherer-Lorenzen et al., 2003; Tilman et al. 1996); limited evidence for observational studies (e.g. Klaus et al., 2018)
Soil microbial C	188	Measure of soil microbial biomass, related to carbon cycling of terrestrial ecosystems and used as an indicator for soil quality. Measured as difference in K_2SO_4 -extractable soil organic carbon concentration of chloroform-fumigated and non-fumigated samples	Increased microbial biomass with species richness (Lama et al. 2020), due to increased belowground productivity or plant-soil feedbacks (Thakur et al. 2015); limited evidence for observational studies (Grigulis et al., 2013)
Soil microbial C:N ratio	188	Stoichiometry of soil microbial biomass, related to C and N dynamics and to the ratio of fungi/bacteria in soils. Measured as soil microbial C-to-N ratio following chloroform-fumigation-extraction	Related to fast-slow resource acquisition, negatively related to N mineralization rates (Lama et al., 2020). A positive correlation between plant species richness and microbial C:N is assumed but was not proven (Lama et al., 2020)
Root biomass	288	Belowground root biomass per m ² down to 10 cm depth, related to nutrient uptake, drought resistance, erosion control, as well as C and N cycling.	Increased belowground productivity in species-rich communities (Jochum et al., 2020; Ravenek et al., 2014); observational studies

Aboveground plant N uptake	175	Related to N cycling. Calculated from N stored in aboveground biomass per m ² relative to soil K ₂ SO ₄ -extractable N	Enhanced resource-partitioning in diverse communities leads to greater exploitation of available nitrogen (Mueller et al., 2013; Tilman et al., 1996); observational studies
Aboveground plant P uptake	274	Related to phosphorus (P) cycling. Calculated from P stored in aboveground biomass per m ² relative to soil NaHCO ₃ -extractable P (P-Olsen)	Enhanced resource-partitioning in diverse communities leads to greater exploitation of available P in experimental and observational settings (Oelmann et al., 2021)
Decomposition of high-quality litter	255	Related to C cycling of easily degradable organic material. Mass loss of standardized litter bags (green tea)	Increased decomposition with species richness via plant-soil feedbacks (Cardinale et al., 2012); weaker evidence for observational studies (van der Plas, 2019)
Decomposition of low-quality litter	255	Related to C cycling of poorly degradable organic material. Mass loss of standardized litter bags (rooibos tea)	Same as above, but related to the break-down of more recalcitrant organic material (Keuskamp et al., 2013)

Analyses

To analyse within site effects of initially sown species (Hypothesis I) and within (Hypothesis II) and among site (Hypothesis III) realized species richness on ecosystem functioning we performed the following analyses:

Increased species pool effects of initial seeding on ecosystem functions

First, to assess the effect of an increased local species pool due to seeding (as opposed to the changes in realized species richness), we modelled ecosystem functions with the dummy-coded seeding and disturbance treatments and their interaction as predictors. We added a varying intercept for grassland sites and log-transformed aboveground plant biomass, root biomass, soil microbial C concentration, N and P uptake and soil N leaching risk to meet distributional assumptions. We assumed Student-T errors for all functions, which is robust in modelling outliers and allows estimating residual correlations between functions, because responses measured on the same experimental unit might not be independent. During model fitting, we simultaneously predicted missing values of ecosystem functions via posterior prediction because we otherwise had to remove observations with missing data to estimate residual correlations among ecosystem functions (Bürkner, 2017; see Table 1 for sample sizes).

Species richness and functional composition effects within grasslands sites

Second, we examined how realized species richness and associated functional composition is linked to ecosystem functions ***within sites***. We centred the log-transformed species richness and CWMs of SLA and root diameter to the grassland mean (i.e., $x_{jk} = x_{jk} - \underline{x}_k$ for subplot j in grassland k). The grassland-mean-centred values are independent of the global-mean centred counterparts as well as other grassland-level variables and allow to separate the within-site effects on ecosystem functions from among-site effects (Enders & Tofighi, 2007). We then standardised all response and predictor variables to zero global mean and unit standard deviation (SD) to obtain comparable effect sizes (i.e. $x = (x - \underline{x})/S D_x$).

We modelled all twelve ecosystem functions simultaneously using multilevel multivariate regression with the grassland-mean-centred species richness (log-transformed), SLA and root diameter as predictors. We included varying intercepts for grassland sites and added a dummy-coded disturbance covariate (allowed to vary between grasslands) to account for the disturbance treatment. As in the previous analysis, we log-transformed the above-mentioned functions to meet distributional assumptions, assumed Student-T errors and estimated the residual correlations between functions.

Because we expected diversity effects on ecosystem functioning to be strongest at low (initial) species richness (Cardinale et al., 2012), we included an interaction term between grassland-mean-centred richness and control-treatment richness (the same control richness for all subplots within a site). We compared models with and without the interaction based on the expected log pointwise predictive density ($elpd$; derived from leave-one-out cross-validation, Vehtari et al., 2017) to judge whether the interaction improved our predictions of ecosystem function. We also tested whether an interaction with LUI improved model performance.

271 *Variation in ecosystem functions among site*

272 Third, we used a path analysis to test how diversity, functional composition and the environment relate
273 to differences in ecosystem functioning **among** sites (see path diagram in Figure 4). To separate the
274 among-site effects of species richness (log) and functional composition, we replaced the measurement of
275 variable x on subplot j within site k with their respective grassland means: $x_{jk} = \underline{x}_k$ (Enders & Tofighi
276 2007). We further centred all response and predictor variables to zero global mean and unit SD to obtain
277 comparable effect sizes (i.e., $x = (x - \underline{x})/S D_x$).

278 We related site-means of species richness, CWMs of SLA and root diameter, land-use intensity, TWI and
279 soil pH to all 12 ecosystem functions. To account for mediating effects of diversity and functional
280 composition, we simultaneously modelled species richness (log) and CWMs with land use, TWI and soil
281 pH as predictors (see path diagram in Figure 4). As before, we included a variable intercept for grassland
282 site and added a dummy-coded disturbance covariate (allowed to vary among sites) to account for the
283 disturbance treatment. As with within-site analyses, we modelled all responses simultaneously using
284 multivariate regression and estimated the residual correlations between responses, because responses
285 measured on the same experimental unit might not be independent. We simultaneously predicted
286 missing values of ecosystem functions during model fitting. We assumed Student-T errors for all functions
287 and log-transformed aboveground plant biomass, root biomass, soil microbial C concentration and soil N
288 leaching risk to meet distributional assumptions.

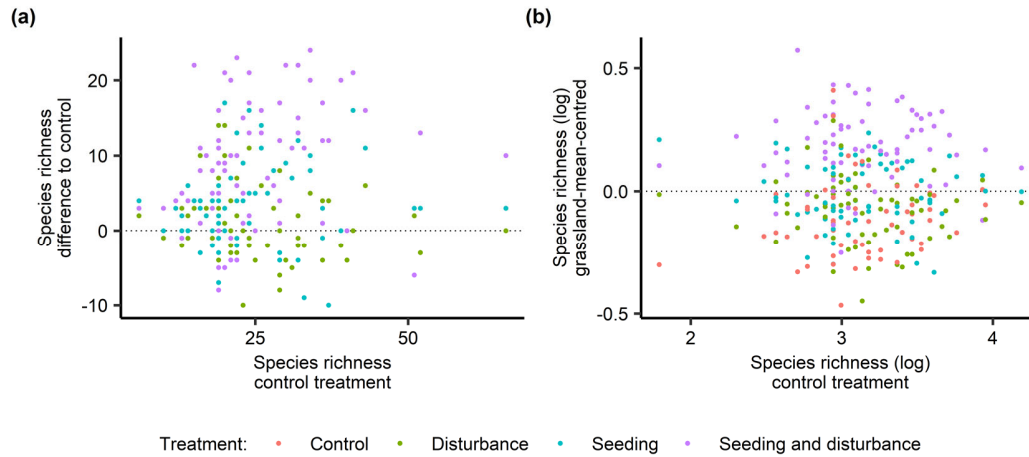
289 We fitted the multilevel multivariate models in a Bayesian framework using the Stan probabilistic
290 language (Stan Development Team, 2020) accessed via the package *brms* v2.15.0 (Bürkner, 2017) in R
291 v4.0.5 (R Core Team 2021) and used the package *loo* for leave-one-out cross-validation (Vehtari et al.,
292 2020). For all models, we specified weakly regularizing normal priors ($\mu = 0$, $\sigma = 2$ for fixed parameters; μ
293 $= 0$, $\sigma = 0.5$ for intercepts; Gelman et al., 2017). We fixed the Student-T freedom parameter to $\nu = 4$,

because the sample size ($N \leq 292$) is too low to reliably estimate ν . We ran four parallel chains and 4,000 iterations (2,000 discarded as burn-in) and ensured convergence with R-hat values ≤ 1.01 for all parameters. Model fit was assessed with posterior-predictive checks for all responses using the package *bayesplot* v1.7.2 (Gabry & Mahr, 2019; Appendix Figures S4 and S7).

RESULTS

The experimental treatments, i.e., seed addition and topsoil disturbance, increased plant richness on average by 8.8 species and created a significant gradient of *within-site* species richness (Figure 1). We observed the highest richness gains in already species-rich grasslands (Figure 2a) with low aboveground biomass productivity, where species with a low SLA established best (see also Freitag et al. 2021). Grassland-mean-centred species richness (log) was independent of CWM traits within sites (Appendix Figure S2). Similarly, species richness (log) due to the experimental treatments was unrelated to control species richness (Figure. 2b), which is highly relevant as BEF-relations are usually calculated on the basis of species richness (log).

309



310

311 **Figure 2:** Variation in plant species richness four years into the experiment (N=292). (a) Gains in species
 312 richness by combined seeding and topsoil disturbance tended to be highest in already species-rich
 313 grasslands (control plots not shown; Freitag et al., 2021). (b) Considering log-transformed species richness,
 314 however, the within-site gradient in species richness (grassland-mean-centred) was independent of
 315 control-treatment species richness.

316

317 ***Species-pool effects of seeding treatments on ecosystem functions***

318 *Within sites*, the twelve ecosystem functions were hardly affected by the seeding treatments,
 319 representing the effects of an enlarged local species pool. Only aboveground plant N uptake showed a
 320 negative response to the seeding ($\beta=-0.14$, 90% CrI: -0.28 to 0.00, Appendix Table S2) and a positive to
 321 seeding*disturbance interaction effects ($\beta=0.24$, 90% CrI: 0.03 to 0.43). This indicates a reduction of
 322 aboveground plant N uptake if seeding is applied without topsoil disturbance. Similarly, but less clear,
 323 aboveground plant $\delta^{15}\text{N}$ weakly increased with the seeding ($\beta=0.05$, 90% CrI: -0.02 to 0.12) and decreased
 324 with the seeding*disturbance interaction effects ($\beta=-0.09$, 90% CrI: -0.19 to 0.01)

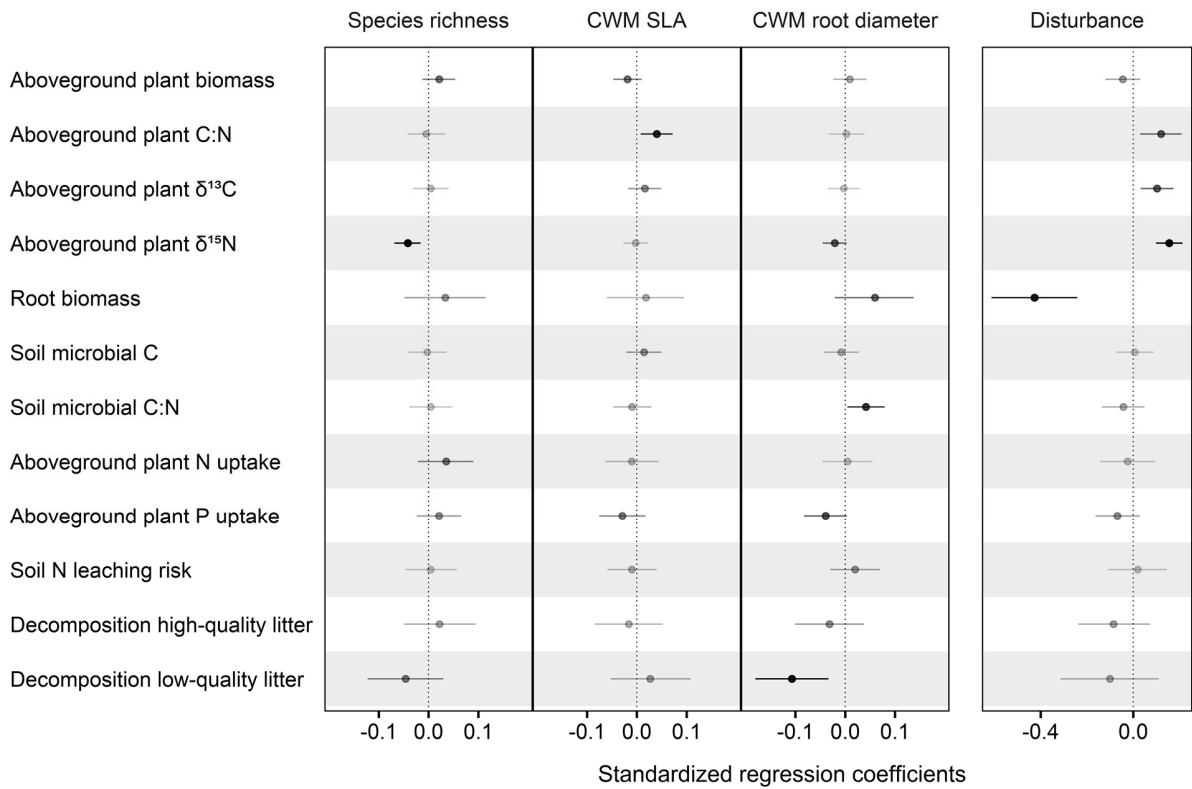


Figure 3: Standardized *within-site* effects of plant species richness and CWMs of specific leaf area (SLA) and root diameter on ecosystem functions. To highlight clear effects, effect colours range from black to light grey based on the inverse ratio between the median and the 90% credible interval width. Predictors were site-mean-centred to separate within-site effects from variation in land-use intensity and abiotic conditions among grasslands. Ecosystem functions were modelled simultaneously with residual correlations estimated using multilevel multivariate regression. A varying intercept for grassland and the dummy-coded disturbance effect (allowed to vary among sites) were included due to the experimental design features. Points indicate medians along with 90% CrI of posterior samples.

Within-site effects of realized species richness and functional composition on ecosystem functions

The effects of experimentally created *within-site* variation in species richness (realised richness) or functional traits on ecosystem functions (Figure 3) were similarly weak as the effects of an enlarged species pool. The two exceptions were aboveground plant $\delta^{15}\text{N}$, which decreased with site-mean-centred

species richness ($\beta=-0.04$, 90% CrI: -0.02 to -0.07, Appendix Table S1) and aboveground plant biomass (log-transformed), which marginally increased with higher species richness ($\beta=0.02$, 90% CrI: -0.01 to 0.05). This implies an average increase in aboveground biomass of only 2.7%, given an average 34% increase in species richness in the combined sowing and disturbance treatment. The inclusion of an interaction with control-treatment species richness or land-use intensity did not improve model predictions for any ecosystem function (loo-derived *elpd* of interaction model lower or within standard errors of *elpd* difference). This indicates that, in contrast to our hypothesis, plant species richness effects, or the lack of them, were independent of initial species richness or land-use intensity.

For all combinations of traits and functions studied, we only found three significant trait-function relationships. The first of these was that SLA increased aboveground biomass C:N ratio ($\beta=0.04$, 90% CrI: 0.01 to 0.07). The second and third was that high fine root diameter was associated with slower decomposition of low-quality litter ($\beta=-0.11$, 90% CrI: -0.18 to -0.03) and an increase in the soil microbial C:N ratio ($\beta=0.04$, 90% CrI: 0.00 to 0.08). Topsoil disturbance had a lasting effect on several ecosystem functions; it strongly reduced root biomass ($\beta=-0.43$, 90% CrI: -0.61 to -0.24), but increased aboveground plant $\delta^{13}\text{C}$ ($\beta=0.10$, 90% CrI: 0.03 to 0.17) and $\delta^{15}\text{N}$ ($\beta=0.16$, 90% CrI: 0.10 to 0.21).

Among-site effects of species richness on ecosystem functioning

The path analysis revealed strong effects of land-use intensity and topographic wetness (TWI) on plant diversity and functional composition (Figure 4). Land-use intensity was associated with decreasing species richness (Figure 4a, $\beta=-0.37$, 90% CrI: -0.53 to -0.22, Appendix Table S3) and increasing SLA ($\beta=0.30$, 90% CrI: 0.11 to 0.48). Regarding changes in functions, species richness was negatively related to aboveground plant $\delta^{15}\text{N}$, soil microbial C and soil N concentration (Figure 4e,g,k). Meanwhile, land-use intensity and TWI had positive effects on aboveground plant biomass and aboveground plant $\delta^{15}\text{N}$, further accompanied by positive indirect effects via SLA (Figure 4b,e). In line with hypothesis III, our results clearly

361 showed that environmental drivers and land use, not plant species richness, are the main drivers of
 362 ecosystem functioning among sites.

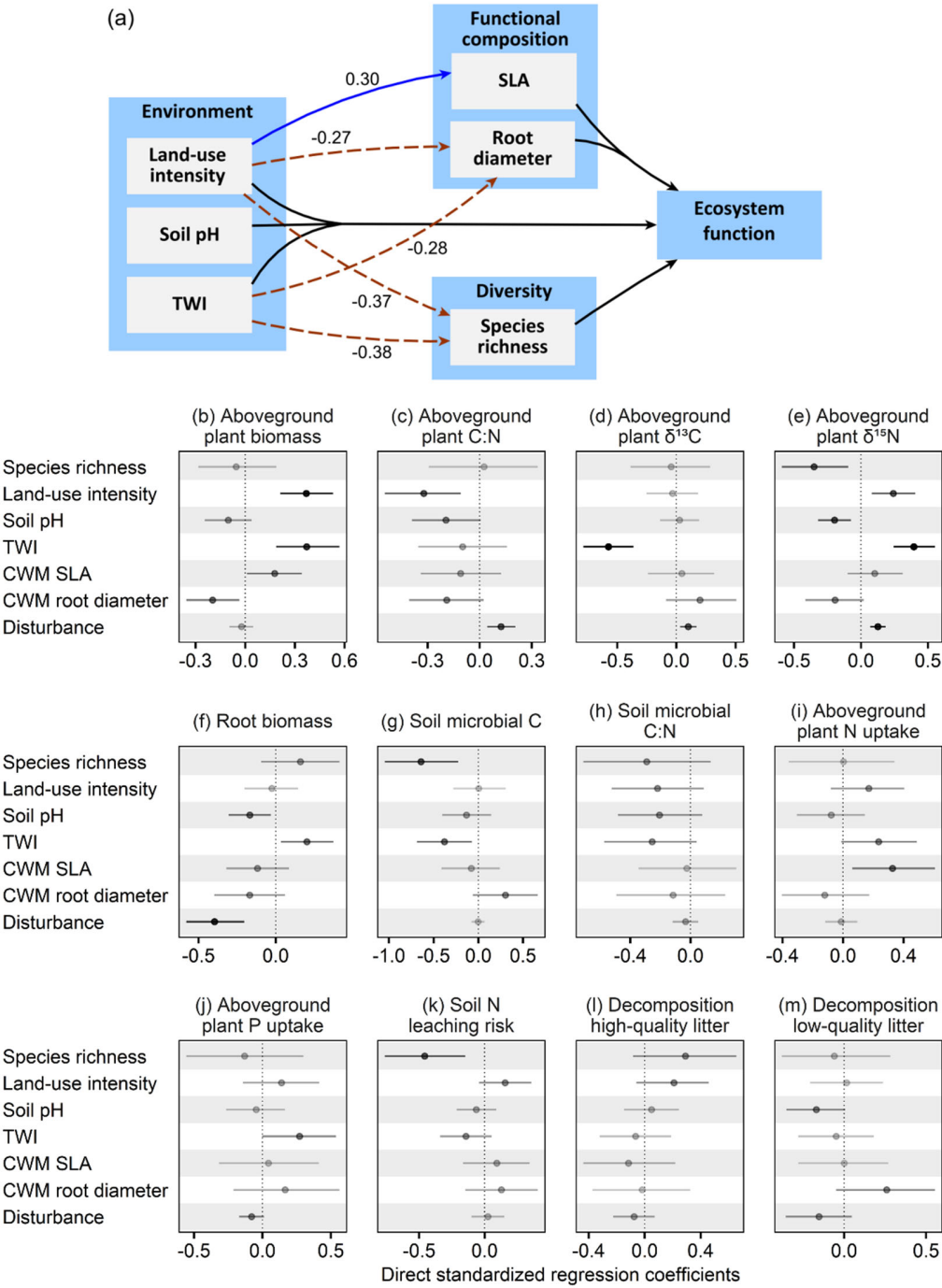


Figure 4: (a) Path analysis diagram with standardized *among-site* effects of land-use intensity and abiotic conditions on plant species richness (log) and traits CWMs. Posterior means and 90% CrI are shown, but soil pH effects not shown as the 90% CrI clearly overlapped zero (Appendix Table S3 for all effects). Residual correlation between plant species richness and specific leaf area (SLA) was -0.14 and between plant species richness and root diameter was 0.17. (b)-(m) Direct standardized among-site effects on ecosystem functions as modelled in the path analysis. Ecosystem functions were modelled simultaneously, and residual correlations estimated using multilevel multivariate regression. A varying intercept for grassland and a disturbance effect (allowed to vary among sites) were included to account for the experimental design. Shown are posterior means, (b)-(m) along with 90% CrI in panels (b)-(m). See Appendix Figure S8 for residual correlations.

DISCUSSION

To overcome the limitations of both observational and strongly controlled BEF studies, we created a *within-site* plant species richness gradient via seeding that enabled us to separate the effects of *among-site* (covarying with environmental drivers) and *within-site* effects of species richness on ecosystem functioning (Enders & Tofighi, 2007). We found only weak evidence for changes in ecosystem functioning with increased local species-pool richness (hypothesis I) and realized plant species richness (hypothesis II) within the studied agricultural grasslands (*within-site* effects). Contrary, *among sites* species richness was clearly associated with three out of twelve ecosystem-functions along confounding gradients of abiotic environmental conditions and land-use intensity.

Besides aboveground biomass production that increased only marginally, the only function affected by *within-site* variation in (realized) species richness was aboveground plant $\delta^{15}\text{N}$ (Figure 3). Aboveground plant $\delta^{15}\text{N}$ became more negative with higher species richness, indicating reduced N losses in species-rich

communities. This finding is in line with results of BEF experiments (Gubsch et al., 2011) and has been previously reported for a large set of grasslands including those studied here (Kleinebecker et al., 2014). Increased resource partitioning in diverse communities may have led to a more complete use of available N resources, which minimizes the risk of N losses due to denitrification or leaching, processes that cause residual enrichment in $\delta^{15}\text{N}$ (Kahmen et al., 2006).

Variation in the functional community composition within sites was weakly related to ecosystem functioning. Yet, contrary to species richness, increases in SLA were positively related to aboveground plant $\delta^{15}\text{N}$. This suggests a more open N cycle in ‘fast’ resource-acquisitive communities characterized by high SLA, possibly because fast resource-uptake goes along with lower leaf and root lifespan and faster N mineralization (Reich, 2014), which increases N turnover and the risk of N losses. Similar patterns have been shown for the relation between plant traits, soil carbon and soil microbial properties (Manning et al., 2015; de Vries et al., 2012).

Although weakly, aboveground plant $\delta^{15}\text{N}$ also tended to decrease with an increased local species pool (hypothesis I), especially when seeding was combined with topsoil disturbance. Yet, against our expectation, increasing the species pool by seeding did not reveal further BEF relationships except for an unclear effect of seeding on aboveground plant N uptake. Beyond this effect, we could not find further evidence for more efficient N use in other N-related functions, such as the soil N leaching risk or microbial C:N.

When focussing on variation in diversity and ecosystem functions *among sites*, we again found a negative correlation of species richness with aboveground plant $\delta^{15}\text{N}$, but also negative associations between species richness and soil microbial C and the soil N leaching risk. Yet, we are cautious to interpret these among-site effects of species richness as causal relationships. The negative effect of species richness on N leaching risk, for example, may also reflect that species richness is a better indicator of local soil

conditions and long-term nutrient inputs on site (resulting in competitive exclusion, Harpole et al., 2016) than our measurements of recent land-use intensity (Kleinebecker et al., 2018).

In sum, our results suggest that plant species enrichment of up to nine species per 4 m² had only minor effects on ecosystem functioning in our moderately intensive to traditionally managed agricultural grasslands. Both possible approaches of analysing within-site BEF relationships (initial species pool versus realized species richness, see Hagan et al., 2021) yielded comparable and similarly weak results in terms of altered ecosystem functioning. There are three possible and interlinked explanations for our findings:

First, BEF relationships quickly approach saturation (Cardinale et al. 2012), often strongly levelling off at around 10 species per community (Jochum et al., 2020; Tilman et al., 1996). However, even in the least diverse grasslands in our study we recorded 10 species per 4 m² on the control treatments, and we found no evidence that ecosystem functioning in these species-poor communities was more clearly affected by enhanced species richness. Although species gains perhaps led to a more complete exploitation of N, this indication of enhanced resource partitioning hardly translated into changes in other ecosystem functions except for a rather insignificant increase in aboveground plant biomass. It seems that strong biodiversity effects, based on considerable increases in complementarity via resource partitioning, might be confined to even lower diversity levels. The functional benefits of grassland species enrichment may be greatest in very species poor grassland systems, such as those sown for silage cropping (Finn et al. 2013). Moreover, complementarity effects might be slow and delayed and there is evidence from BEF experiments that biodiversity–ecosystem functioning relationships become stronger over time (Guerrero-Ramírez et al., 2017).

Second, small-statured species with low SLA had higher average establishment success than high-SLA species, and low-SLA species preferably established in low-productivity grasslands (Freitag et al., 2021). This may explain why increases in richness did not substantially raise productivity, as high productivity,

433 'fast' functioning species (i.e. species with high SLA) from the local species pool were already present.
434 Stabilizing effects of low-SLA species on ecosystems functioning, however, may become visible in the long
435 term through the buffering of extreme events such as exceptional droughts (Tilman & Downing, 1994).

436 Third, among grassland sites, the strong influence of environmental drivers has obviously masked and
437 overruled diversity effects. Here, land-use intensity and topographic wetness strongly affected both
438 species richness and plant community traits. Environmental drivers also strongly influenced ecosystem
439 functions, for example, aboveground plant production, aboveground plant C:N ratio (biomass quality) and
440 N uptake. Observational studies comparing BEF relationships among grassland sites therefore typically
441 find weak diversity effects, often because diversity and functional composition are the result of land use
442 and abiotic conditions (van der Plas, 2019). However, in our study we also hardly found positive effects of
443 raised species richness on ecosystem functioning within sites, where differences in land use and abiotic
444 conditions between treatments are negligible.

445 In summary, our results showed only very limited increases in ecosystem functioning as a response to
446 seeding-induced richness gains. This begs the question as to under which circumstances restoring plant
447 diversity can be more beneficial. Enhancing ecosystem functioning by restoration of plant diversity might
448 be confined to previously degraded or extremely species-poor systems on ex-arable land, where
449 enhanced plant species richness by seeding indeed has led to higher productivity (Bullock et al., 2007;
450 Ladouceur et al., 2020), higher biomass quality (Savage et al., 2021), accelerated litter decomposition
451 (Zirbel et al., 2019), and reduced soil inorganic N losses (Klopf et al., 2017; Mueller et al., 2013). Many
452 grassland BEF experiments were created on ex-arable soils, which are, unlike old grasslands soils,
453 degraded in view of low soil C_{org} content and reduced soil fauna (Lange et al., 2015). Increased root and
454 microbial biomass inputs as well as an improved soil structure via plant-soil feedbacks have been
455 identified as main mechanisms that may strongly contribute to complementarity effects observed in

experiments (Barry et al., 2019; Eisenhauer et al., 2017; Klopff et al., 2017). Yet, in contrast to BEF experiments on ex-arable land, the soils of the permanent agricultural grasslands in this study may be less degraded and hence benefit not as much from increased plant species richness in terms of soil organic carbon accumulation and other soil properties (Lange et al., 2015). This indicates that some BEF relationships may be of greater applicability to previously degraded, immature grasslands with a strongly depleted species pool, where dispersal limitation is a crucial driver that controls the arrival of high-performing species. In more mature grasslands with less degraded soils, expectations of considerable increases in ecosystem functioning after plant diversity restoration might not be met, even if the species pool of a grassland was significantly enriched, as in this study. Yet, other functions or ecosystem services that more directly depend on plant diversity, such as pollination, flowering aspects and grassland aesthetics, are more likely to improve with plant diversity restoration (Allan et al. 2015).

From our *among site* analysis we further conclude that the control of both plant species richness and community assembly by environmental drivers and especially land use overrides BEF relationships and limits the transferability of experimental BEF results to the maintenance and restoration of real-world grassland ecosystems. In this context, our approach of creating *within site* gradients in species richness by seeding is a key step to bridge the gap between classical BEF experiments and observational studies under realistic settings of community assembly and ecosystem management.

ACKNOWLEDGEMENTS

Laboratory analyses were carried out at the Institute of Landscape Ecology, Münster, and the Institute of Soil Science and Land Evaluation, Hohenheim. Representing numerous helpers, we thank Ulrike-Berning Mader, Judith Hinderling, Torsten Meene, Svenja Kunze and many students for invaluable help during field and lab work and Sascha Nowak for performing soil chemical analyses. We also thank the managers of the

479 three Exploratories, Kirsten Reichel-Jung, Iris Steitz, Sandra Weithmann, Florian Straub, Katrin Lorenzen,
480 Juliane Vogt, Martin Gorke and Miriam Teuscher for their work in realizing this experiment and
481 maintaining the plot and project infrastructure, Christiane Fischer and Jule Mangels for giving support
482 through the central office, Michael Owonibi and Andreas Ostrowski for managing the central database,
483 and Markus Fischer, Eduard Linsenmair, Dominik Hessenmöller, Jens Nieschulze, Francois Buscot, Ernst-
484 Detlef Schulze, Wolfgang W. Weisser, and the late Elisabeth Kalko for their role in setting up the
485 Biodiversity Exploratories project. We thank the administration of the Hainich national park, the UNESCO
486 Biosphere Reserve Swabian Alb and the UNESCO Biosphere Reserve Schorfheide-Chorin as well as all land
487 owners for the excellent collaboration. The work has been funded by the German Research Foundation
488 (DFG) Priority Program 1374 'Infrastructure-Biodiversity-Exploratories' (FI 1246/15-2; HA 4597/6-3; HO
489 3830/2-4; KL 2265/4-4, KL 1866/12-1). Anna Abrahão was partly funded by the Coordenação de
490 Aperfeiçoamento de Pessoal de 452 Nível Superior - Brasil (CAPES) (Finance Code 001, process
491 88881.172163/2018-01). Lena Neuenkamp has been supported by the Marie Skłodowska Curie Individual
492 Fellowship (MYFUN, Grant number: 835472). Field work permits were issued by the responsible state
493 environmental offices of Baden-Württemberg, Thüringen, and Brandenburg.

REFERENCES

- Allan, E., Manning, P., Alt, F., Binkenstein, J., Blaser, S., Blüthgen, N. et al. (2015) Land use intensification alters ecosystem multifunctionality via loss of biodiversity and changes to functional composition. *Ecology Letters*, 18, 834–843.
- Bachmann, D., Gockele, A., Ravenek, J.M., Roscher, C., Strecker, T., Weigelt, A., et al. (2015). No evidence of complementary water use along a plant species richness gradient in temperate experimental grasslands. *PLOS ONE*, 10, e0116367.
- Bannar-Martin, K.H., Kremer, C.T., Ernest, S.K.M., Leibold, M.A., Auge, H., Chase, J. et al. (2018) Integrating community assembly and biodiversity to better understand ecosystem function: The Community Assembly and the Functioning of Ecosystems (CAFE) approach. *Ecology Letters*, 21, 167–180.
- Barry, K.E., Mommer, L., van Ruijven, J., Wirth, C., Wright, A.J., Bai, Y. et al. (2019) The future of complementarity: Disentangling causes from consequences. *Trends in Ecology and Evolution*, 34, 167–180.
- Bergmann, J., Weigelt, A., Plas, F. van der, Laughlin, D.C., Kuyper, T.W., Guerrero-Ramirez, N. et al. (2020) The fungal collaboration gradient dominates the root economics space in plants. *Science Advances*, 6, eaba3756.
- Blüthgen, N., Dormann, C.F., Prati, D., Klaus, V.H., Kleinebecker, T., Hölzel, N. et al. (2012) A quantitative index of land-use intensity in grasslands: Integrating mowing, grazing and fertilization. *Basic and Applied Ecology*, 13, 207–220.
- Bullock, J.M., Pywell, R.F. & Walker, K.J. (2007) Long-term enhancement of agricultural production by restoration of biodiversity. *Journal of Applied Ecology*, 44, 6–12.

516 Bürkner, P.-C. (2017). Brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical*
517 *Software*, 80, 1–28.

518 Cardinale, B.J. (2011) Biodiversity improves water quality through niche partitioning. *Nature*, 472, 86–89.

519 Cardinale, B.J., Duffy, J.E., Gonzalez, A., Hooper, D.U., Perrings, C., Venail, P. et al. (2012) Biodiversity loss
520 and its impact on humanity. *Nature*, 486, 59–67.

521 Chase, J.M., McGill, B.J., McGlinn, D.J., May, F., Blowes, S.A., Xiao, X. *et al.* (2018) Embracing scale-
522 dependence to achieve a deeper understanding of biodiversity and its change across communities.
523 *Ecology Letters*, 21, 1737–1751.

524 De Laender, F., Rohr, J.R., Ashauer, R., Baird, D.J., Berger, U., Eisenhauer, N. et al. (2016) Reintroducing
525 environmental change drivers in biodiversity–ecosystem functioning research. *Trends in Ecology &*
526 *Evolution*, 31, 905–915.

527 de Vries, F.T., Manning, P., Tallowin, J.R.B., Mortimer, S.R., Pilgrim, E.S., Harrison, K.A. et al. (2012) Abiotic
528 drivers and plant traits explain landscape-scale patterns in soil microbial communities. *Ecology Letters*,
529 15, 1230–1239.

530 Diaz, S., Lavorel, S., de Bello, F., Quetier, F., Grigulis, K. & Robson, T.M. (2007) Incorporating plant
531 functional diversity effects in ecosystem service assessments. *Proceedings of the National Academy of*
532 *Sciences*, 104, 20684–20689.

533 Eisenhauer, N., Lanoue, A., Strecker, T., Scheu, S., Steinauer, K., Thakur, M.P. et al. (2017) Root biomass
534 and exudates link plant diversity with soil bacterial and fungal biomass. *Scientific Reports*, 7, 44641.

535 Enders, C.K. & Tofighi, D. (2007) Centering predictor variables in cross-sectional multilevel models: A new
536 look at an old issue. *Psychological Methods*, 12, 121–138.

537 Finn, J. A., Kirwan, L., Connolly, J., Sebastià, M. T., Helgadottir, A., Baadshaug, O. H. et al. (2013) Ecosystem
 538 function enhanced by combining four functional types of plant species in intensively managed grassland
 539 mixtures: A 3-year continental-scale field experiment. *Journal of Applied Ecology*, 50, 365–375.

540 Fischer, M., Bossdorf, O., Gockel, S., Hänsel, F., Hemp, A., Hessenmöller, D. et al. (2010) Implementing
 541 large-scale and long-term functional biodiversity research: The Biodiversity Exploratories. *Basic and*
 542 *Applied. Ecology*, 11, 473–485.

543 Freitag, M., Klaus, V.H., Bolliger, R., Hamer, U., Kleinebecker, T., Prati, D. et al. (2021) Restoration of plant
 544 diversity in permanent grassland by seeding: Assessing the limiting factors along land-use gradients.
 545 *Journal of Applied Ecology*, 58, 1681–1692.

546 Gabry, J. & Mahr, T. (2019) *Bayesplot: Plotting for Bayesian models. R package version 1.7.2*. Available at:
 547 <http://mc-stan.org/bayesplot>.

548 Gelman, A., Simpson, D. & Betancourt, M. (2017) The prior can often only be understood in the context
 549 of the likelihood. *Entropy*, 19, 555.

550 Gessler, P.E., Moore, I.D., McKenzie, N.J. & Ryan, P.J. (1995) Soil-landscape modelling and spatial
 551 prediction of soil attributes. *International Journal of Geographical Information Systems*, 9, 421–432.

552 Grigulis, K., Lavorel, S., Krainer, U., Legay, N., Baxendale, C., Dumont, M., et al. (2013). Relative
 553 contributions of plant traits and soil microbial properties to mountain grassland ecosystem services.
 554 *Journal of Ecology*, 101, 47–57.

555 Gross, K. & Cardinale, B.J. (2005) The functional consequences of random vs. Ordered species extinctions.
 556 *Ecology Letters*, 8, 409–418.

557 Gubsch, M., Roscher, C., Gleixner, G., Habekost, M., Lipowsky, A., Schmid, B. et al. (2011) Foliar and soil
558 $\delta^{15}\text{N}$ values reveal increased nitrogen partitioning among species in diverse grassland communities. *Plant*
559 *Cell & Environment*, 34, 895–908.

560 Guderle, M., Bachmann, D., Milcu, A., Gockele, A., Bechmann, M., Fischer, C., et al. (2018). Dynamic niche
561 partitioning in root water uptake facilitates efficient water use in more diverse grassland plant
562 communities. *Functional Ecology*, 32, 214–227.

563 Guerrero-Ramírez, N.R., Craven, D., Reich, P.B., Ewel, J.J., Isbell, F., Koricheva, J. et al. (2017) Diversity-
564 dependent temporal divergence of ecosystem functioning in experimental ecosystems. *Nature Ecology &*
565 *Evolution*, 1, 1639.

566 Hagan, J.G, Vanschoenwinkel, B., & Gamfeldt, L. (2021) We should not necessarily expect positive
567 relationships between biodiversity and ecosystem functioning in observational field data. *Ecology Letters*,
568 24, 2537–2548.

569 Harpole, W.S., Sullivan, L.L., Lind, E.M., Firn, J., Adler, P.B., Borer, E.T. et al. (2016) Addition of multiple
570 limiting resources reduces grassland diversity. *Nature*, 537, 93–96.

571 Hooper, D.U., Adair, E.C., Cardinale, B.J., Byrnes, J.E.K., Hungate, B.A., Matulich, K.L. et al. (2012) A global
572 synthesis reveals biodiversity loss as a major driver of ecosystem change. *Nature*, 486, 105–108.

573 Isbell, F., Tilman, D., Polasky, S. & Loreau, M. (2015) The biodiversity-dependent ecosystem service debt.
574 *Ecology Letters*, 18, 119–134.

575 Jochum, M., Fischer, M., Isbell, F., Roscher, C., van der Plas, F., Boch, S. et al. (2020) The results of
576 biodiversity–ecosystem functioning experiments are realistic. *Nature Ecology & Evolution*, 4, 1485–1494.

577 Kahmen, A., Renker, C., Unsicker, S.B. & Buchmann, N. (2006) Niche complementarity for nitrogen: An
578 explanation for the biodiversity and ecosystem functioning relationship? *Ecology*, 87, 1244–1255.

579 Kattge, J., Bönisch, G., Diaz, S., Lavorel, S., Prentice, I.C., Leadley, P. et al. (2020) TRY plant trait database–
580 enhanced coverage and open access. *Global Change Biology* 26, 119–188.

581 Keuskamp, J.A., Dingemans, B.J.J., Lehtinen, T., Sarneel, J.M. & Hefting, M.M. (2013). Tea Bag Index: a
582 novel approach to collect uniform decomposition data across ecosystems. *Methods in Ecology and*
583 *Evolution*, 4, 1070–1075.

584 Klaus, V.H., Whittingham, M.J., Báldi, A., Eggers, S., Francksen, R.M., Hiron, M. et al. (2020) Do biodiversity-
585 ecosystem functioning experiments inform stakeholders how to simultaneously conserve biodiversity and
586 increase ecosystem service provisioning in grasslands? *Biological Conservation*, 245, 108552.

587 Klaus, V.H., Kleinebecker, T., Busch, V., Fischer, M., Hölzel, N., Nowak, S. et al. (2018) Land use intensity,
588 rather than plant species richness, affects the leaching risk of multiple nutrients from permanent
589 grasslands. *Global Change Biology*, 24, 2828–2840.

590 Klaus, V.H., Schäfer, D., Kleinebecker, T., Fischer, M., Prati, D. & Hölzel, N. (2017). Enriching plant diversity
591 in grasslands by large-scale experimental sward disturbance and seed addition along gradients of land-
592 use intensity. *Journal of Plant Ecology*, 10, 581–591.

593 Klaus, V.H., Hölzel, N., Prati, D., Schmitt, B., Schöning, I., Schrumpf, M., et al. (2016). Plant diversity
594 moderates drought stress in grasslands: Implications from a large real-world study on ¹³C natural
595 abundances. *Science of The Total Environment*, 566–567, 215–222.

596 Kleinebecker, T., Hölzel, N., Prati, D., Schmitt, B., Fischer, M. & Klaus, V.H. (2014) Evidence from the real
597 world: ¹⁵N natural abundances reveal enhanced nitrogen use at high plant diversity in Central European
598 grasslands. *Journal of Ecology*, 102, 456–465.

599 Kleinebecker, T., Busch, V., Hölzel, N., Hamer, U., Schäfer, D., Prati, D., Fischer, M., Hemp, A. et al. (2018)
600 And the winner is ... ! A test of simple predictors of plant species richness in agricultural grasslands.
601 *Ecological Indicators*, 87, 296–301.

602 Klopff, R.P., Baer, S.G., Bach, E.M. & Six, J. (2017) Restoration and management for plant diversity enhances
603 the rate of belowground ecosystem recovery. *Ecological Applications*, 27, 355–362.

604 Lachaise, T., Bergmann, J., Rillig, M.C. & Kleunen, M. van. (2021) Below- and aboveground traits explain
605 local abundance, and regional, continental and global occurrence frequencies of grassland plants. *Oikos*,
606 130, 110–120.

607 Ladouceur, E., Harpole, W. S., Blowes, S.A., Roscher, C., Auge, H., Seabloom, E.W. et al. (2020) Reducing
608 dispersal limitation via seed addition increases species richness but not above-ground biomass. *Ecology*
609 *Letters*, 23, 1442–1450.

610 Laliberté, E. & Legendre, P. (2010) A distance-based framework for measuring functional diversity from
611 multiple traits. *Ecology*, 91, 299–305.

612 Lama, S., Velescu, A., Leimer, S., Weigelt, A., Chen, H., Eisenhauer, N. et al. (2020) Plant diversity
613 influenced gross nitrogen mineralization, microbial ammonium consumption and gross inorganic N
614 immobilization in a grassland experiment. *Oecologia*, 193, 731–748.

615 Lange, M., Eisenhauer, N., Sierra, C. A., Bessler, H., Engels, C., Griffiths, R. I. et al. (2015) Plant diversity
616 increases soil microbial activity and soil carbon storage. *Nature Communications*, 6, 1–8.

617 Leibold, M.A., Chase, J.M. & Ernest, S.K.M. (2017) Community assembly and the functioning of
618 ecosystems: How metacommunity processes alter ecosystems attributes. *Ecology*, 98, 909–919.

619 Le Provost, G., Thiele, J., Westphal, C., Penone, C., Allan, E., Neyret, M. et al. (2021) Contrasting responses
620 of above-and belowground diversity to multiple components of land-use intensity. *Nature*
621 *Communications*, 12, 1–13.

622 Lepš, J. (2004) What do the biodiversity experiments tell us about consequences of plant species loss in
623 the real world? *Basic and Applied Ecology*, 5, 529–534.

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

626 Manning, P., de Vries, F. T., Tallowin, J. R. B., Smith, R., Mortimer, S. R., Pilgrim, E. S. et al. (2015) Simple
 627 measures of climate, soil properties and plant traits predict national-scale grassland soil carbon stocks.
 628 *Journal of Applied Ecology*, 52, 1188–1196.

629 Manning, P., Loos, J., Barnes, A.D., Batáry, P., Bianchi, F.J.J.A., Buchmann, N. et al. (2019) Chapter Ten -
 630 Transferring biodiversity-ecosystem function research to the management of ‘real-world’ ecosystems. In:
 631 *Mechanisms underlying the relationship between biodiversity and ecosystem function*, Advances in
 632 Ecological Research (eds. Eisenhauer, N., Bohan, D.A. & Dumbrell, A.J.). Academic Press, pp. 323–356.

633 Mueller, K.E., Hobbie, S.E., Tilman, D. & Reich, P.B. (2013) Effects of plant diversity, N fertilization, and
 634 elevated carbon dioxide on grassland soil N cycling in a long-term experiment. *Global Change Biology*, 19,
 635 1249–61.

636 Myers, J.A. & Harms, K.E. (2009) Seed arrival, ecological filters, and plant species richness: A meta-analysis.
 637 *Ecology Letters*, 12, 1250–60.

638 Nyfeler, D., Huguenin-Elie, O., Suter, M., Frossard, E., Connolly, J. & Lüscher, A. (2009) Strong mixture
 639 effects among four species in fertilized agricultural grassland led to persistent and consistent transgressive
 640 overyielding. *Journal of Applied Ecology*, 46, 683–691.

641 Oelmann, Y., Lange, M., Leimer, S., Roscher, C., Aburto, F., Alt, F. et al. (2021) Above-and belowground
 642 biodiversity jointly tighten the P cycle in agricultural grasslands. *Nature Communications*, 12, 1–9.

643 Oelmann, Y., Buchmann, N., Gleixner, G., Habekost, M., Roscher, C., Rosenkranz, S., et al. (2011a). Plant
 644 diversity effects on aboveground and belowground N pools in temperate grassland ecosystems:
 645 Development in the first 5 years after establishment. *Global Biogeochemical Cycles*, 25.

646 Oelmann, Y., Richter, A.K., Roscher, C., Rosenkranz, S., Temperton, V.M., Weisser, W.W., *et al.* (2011b).
 647 Does plant diversity influence phosphorus cycling in experimental grasslands? *Geoderma*, 167–168, 178–
 648 187.

649 Olsen, S.R. (1954). *Estimation of Available Phosphorus in Soils by Extraction with Sodium Bicarbonate*. U.S.
 650 Department of Agriculture, Washington D.C.

651 Ostrowski, A., Lorenzen, K., Petzold, E. & Schindler, S. (2020) *Land use intensity index (LUI) calculation tool*
 652 *of the Biodiversity Exploratories project for grassland survey data from three different regions in Germany*
 653 *since 2006, BEXIS 2 module*. Zenodo, <https://doi.org/10.5281/zenodo.3865579>.

654 Pywell, R.F., Bullock, J.M., Roy, D.B., Warman, L.I.Z., Walker, K.J. & Rothery, P. (2003) Plant traits as
 655 predictors of performance in ecological restoration. *Journal of Applied Ecology*, 40, 65–77.

656 R Core Team. (2021) *R: A language and environment for statistical computing v4.0.4*. R Foundation for
 657 Statistical Computing, Vienna, Austria.

658 Ravenek, J.M., Bessler, H., Engels, C., Scherer-Lorenzen, M., Gessler, A., Gockele, A., *et al.* (2014). Long-
 659 term study of root biomass in a biodiversity experiment reveals shifts in diversity effects over time. *Oikos*,
 660 123, 1528–1536.

661 Reich, P.B. (2014) The world-wide ‘fast–slow’ plant economics spectrum: A traits manifesto. *Journal of*
 662 *Ecology*, 102, 275–301.

663 Roscher, C., Schumacher, J., Baade, J., Wilcke, W., Gleixner, G., Weisser, W.W. *et al.* (2004) The role of
 664 biodiversity for element cycling and trophic interactions: An experimental approach in a grassland
 665 community. *Basic & Applied Ecology*, 5, 107–121.

666 Saar, L., Takkis, K., Pärtel, M. & Helm, A. (2012) Which plant traits predict species loss in calcareous
 667 grasslands with extinction debt? *Diversity & Distribution*, 18, 808–817.

668 Savage, J., Woodcock, B. A., Bullock, J. M., Nowakowski, M., Tallowin, J. R. & Pywell, R. F. (2021)
669 Management to support multiple ecosystem services from productive grasslands. *Sustainability*, 13, 6263.

670 Schaub, S., Finger, R., Leiber, F., Probst, S., Kreuzer, M., Weigelt, A. et al. (2020) Plant diversity effects on
671 forage quality, yield and revenues of semi-natural grasslands. *Nature Communications*, 11, 1-11.

672 Scherer-Lorenzen, M., Palmborg, C., Prinz, A., & Schulze, E. D. (2003) The role of plant diversity and
673 composition for nitrate leaching in grasslands. *Ecology*, 84, 1539–1552.

674 Smith, M.D. & Knapp, A.K. (2003) Dominant species maintain ecosystem function with non-random
675 species loss. *Ecology Letters*, 6, 509–517.

676 Srivastava, D.S. & Vellend, M. (2005) Biodiversity-ecosystem function research: Is it relevant to
677 conservation? *Annual Review in Ecology, Evolution & Systematics*, 36, 267–294.

678 Stan Development Team. (2020). *Stan: A C++ library for probability and sampling* v2.19.3. [http://mc-](http://mc-stan.org/)
679 [stan.org/](http://mc-stan.org/).

680 Thakur, M.P., Milcu, A., Manning, P., Niklaus, P.A., Roscher, C., Power, S., et al. (2015). Plant diversity
681 drives soil microbial biomass carbon in grasslands irrespective of global environmental change factors.
682 *Global Change Biology*, 21, 4076–85.

683 Tilman, D. & Downing, J.A. (1994) Biodiversity and stability in grasslands. *Nature*, 367, 363–365.

684 Tilman, D., Wedin, D. & Knops, J. (1996) Productivity and sustainability influenced by biodiversity in
685 grassland ecosystems. *Nature*, 379, 718–720.

686 van der Plas, F. (2019) Biodiversity and ecosystem functioning in naturally assembled communities.
687 *Biological Reviews*, 94, 1220–1245.

688 Vehtari, A., Gabry, J., Yao, Y. & Gelman, A. (2020) *Loo: Efficient leave-one-out cross-validation and WAIC*
689 *for Bayesian models. R package version 2.1.0*. Available at: <https://CRAN.R-project.org/package=loo>.

690 Vehtari, A., Gelman, A. & Gabry, J. (2017) Practical Bayesian model evaluation using leave-one-out cross-
691 validation and WAIC. *Statistics and Computing*, 27, 1413–1432.

692 Verheyen, K., Bulteel, H., Palmborg, C., Olivié, B., Nijs, I., Raes, D., *et al.* (2008). Can complementarity in
693 water use help to explain diversity–productivity relationships in experimental grassland plots? *Oecologia*,
694 156, 351–361.

695 Vogt, J., Klaus, V., Both, S., Fürstenau, C., Gockel, S., Gossner, M. *et al.* (2019) Eleven years’ data of
696 grassland management in Germany. *Biodiversity. Data Journal*, 7, e36387.

697 Wardle, D.A. (2016) Do experiments exploring plant diversity–ecosystem functioning relationships inform
698 how biodiversity loss impacts natural ecosystems? *Journal of Vegetation Science*, 27, 646–653.

699 Zirbel, C.R., Grman, E., Bassett, T. & Brudvig, L.A. (2019) Landscape context explains ecosystem
700 multifunctionality in restored grasslands better than plant diversity. *Ecology*, 100, 02634.