Running title: Experimental approach testing biodiversity effects independent of year

Title: A new experimental approach to test why biodiversity effects strengthen as ecosystems age

Anja Vogel[1,2,3,*], Anne Ebeling[3], Gerd Gleixner[4], Christiane Roscher[5,1], Stefan Scheu[6], Marcel Ciobanu[7], Eva Koller-France[8,9], Markus Lange[4], Alfred Lochner[1,2], Sebastian T. Meyer[10], Yvonne Oelmann[8], Wolfgang Wilcke[9], Bernhard Schmid[11,12,§], Nico Eisenhauer[1,2,§]

Affiliations:

[1]German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Deutscher Platz 5e, 04103 Leipzig, Germany

[2]Institute of Biology, Leipzig University, Deutscher Platz 5e, 04103 Leipzig, Germany

[3]Institute of Ecology and Evolution, Friedrich Schiller University Jena, Dornburger Straße 159, 07743 Jena, Germany

[4]Max Planck Institute for Biogeochemistry, Department of Biogeochemical Processes, Postfach 100164, 07701 Jena, Germany

[5]UFZ, Helmholtz Centre for Environmental Research, Physiological Diversity, Permoserstraße 15, 04318 Leipzig, Germany

[6]J.F. Blumenbach Institute of Zoology and Anthropology, University of Goettingen, Untere Karspüle 2, 37073 Goettingen, Germany

[7]Institute of Biological Research, Branch of the National Institute of Research and Development for Biological Sciences, Str. Republicii 48, 400015 Cluj-Napoca, Romania

[8]Geoecology, University of Tübingen, Rümelinstr. 19-23, 72070 Tübingen, Germany

[9]Institute of Geography and Geoecology, Karlsruhe Institute of Technology, Reinhard-Baumeister-Platze 1, 76131 Karlsruhe, Germany

[10]Department of Ecology and Ecosystem Management, Technische Universität München, Hans-Carl-von-Carlowitz-Platz 2, 85350 Freising-Weihenstephan, Germany

[11]Department of Geography, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland

[12]Institute of Ecology, College of Urban and Environmental Sciences, Peking University, 100871 Beijing, China

* Corresponding author

§ Authors jointly supervised this work

Keywords: biodiversity, ecosystem function, experiment, grassland, community age, soil history, plant history, nematodes

Abstract

Previous experimental studies found strengthening relationships between biodiversity and ecosystem functioning (BEF) over time. Simultaneous temporal changes of abiotic and biotic conditions, such as in the composition of soil communities, soil carbon and nutrient concentrations, plant community

assembly or selection processes, are currently discussed as potential drivers for strengthening BEF relationships. Despite the popularity of these explanations, experimental tests of underlying mechanisms of strengthening BEF relationships over time are scarce, and confounding influences of calendar year cannot be ruled out unless ecosystems of different age are compared in the same calendar years. To address this critical gap of knowledge, we reestablished the plant communities of a long-term biodiversity experiment that had started in 2002 (the Jena Experiment) with new seeds and old or new soil again in 2016. Comparing these treatments with the original communities set up in 2002, we tested whether old communities had stronger plant diversity effects on plant productivity than young ones and if this depended on soil- or plant-related processes. Our first results show that in old communities, the effect of plant diversity on productivity was indeed stronger than in young communities and that this could not be explained by the age of the soil only. However, we found significant effects of soil on the composition of soil organisms, which might be relevant for other ecosystem functions and may have stronger effects over time. Our new experimental approach enables us to test which mechanisms cause strengthening BEFrelationships for many different ecosystem functions independent of the study year.

Introduction

There is growing consensus that biodiversity increases many provisioning and regulating ecosystem services (Cardinale et al. 2012, Duffy et al. 2017, Hooper et al. 2012, Naeem et al. 2012, Eisenhauer et al. 2019), of which the best-studied function is aboveground plant biomass production (Hector et al. 1999, Marquard et al. 2009, Tilman et al. 1997). Moreover, experimental studies manipulating plant diversity found that the magnitude of biodiversity effects on ecosystem functions (BEF) increases over time, i.e. as ecosystems “age” (Guerrero-Ramírez et al. 2017, Huang et al. 2018, Meyer et al. 2016, Reich et al. 2012). This result has important implications because it suggests that experimental results

obtained from newly-assembled, young communities probably underestimate the consequences of biodiversity loss. Studying the mechanisms explaining why biodiversity effects strengthen over time as communities age is crucial in order to understand how important local species loss is for ecosystem functions in real-world ecosystems. Either increases in function at high plant diversity, decreases of function at low plant diversity, or a combination of both have been shown to cause strengthening BEF relationships over time (Guerrero-Ramírez et al. 2017, Meyer et al. 2016). These dissimilar trajectories of plant community performance illustrate context-dependent underlying mechanisms for increasing biodiversity effects (Guerrero-Ramírez et al. 2017). However, in all these previous studies, community age was confounded with calendar year, meaning that uncontrolled environmental conditions of particular years within the study periods may have contributed to biodiversity effects on particular ecosystem functions. Here, we describe a new experimental approach overcoming these shortcomings to get a deeper understanding of the mechanisms driving strengthening biodiversity effects as communities age.

Previous BEF experiments revealed changes in both plant- and soil-related processes over time, leading to hypotheses about the mechanistic links of both groups of processes in driving ecosystem functioning. The plant community level itself underlies temporal changes with potential consequences for the soil system (Reich et al. 2012). Overall, it has been shown that the contribution of multiple plant species to diversity effects on productivity increases over time (Allan et al. 2011, Cardinale et al. 2011, Cardinale et al. 2007, Fargione et al. 2007, Huang et al. 2018, Reich et al. 2012), which is commonly referred to as complementarity effects. The notion is that diverse communities use the resource pool more efficiently than species-poor communities do (Loreau and Hector 2001, Mueller et al. 2013, Roscher et al. 2008; see Barry et al. 2019 for review). By contrast, the dominance of specific productive species in mixtures also has been shown to contribute to plant diversity effects, but such selection effects seem to play a minor role in strengthening biodiversity effects as communities age (Huang et al. 2018, Reich et al. 2012).

Higher complementarity effects are associated with higher fluctuations of plant species over time in diverse communities (Allan et al. 2011). This means that the identity of those plant species, which contribute most to ecosystem function, changes over time (Allan et al. 2011, Isbell et al. 2011). Shifts in plant species composition can be associated with changes in the representation of functional traits in communities which are important for community productivity (Diaz and Cabido 2001, HilleRisLambers et al. 2004) or with their maintenance due to compensatory dynamics (Bai 2004, Yachi and Loreau 1999). The diversity and dominance of community traits indeed can explain a significant proportion of productivity changes over time and across diversity gradients (Reich et al. 2012, Roscher et al. 2012, Roscher et al. 2013). However, it is not only the abundance and composition of plant species, which determine the functional diversity of plant communities and finally ecosystem functioning. Analysis in a long-term biodiversity experiment revealed that besides phenotypic variation in trait expression along the plant diversity gradient (Gubsch et al. 2011, Lipowsky et al. 2015, Roscher et al. 2011a), selection for specific plant phenotypes also increases biodiversity effects on productivity (Lipowsky et al. 2011, van Moorsel et al. 2018, Zuppinger-Dingley et al. 2014). Such evolutionary effects point to the increase of complementarity among plant species due to niche differentiation (Tilman and Snell-Rood 2014).

Beside temporal dynamics in the plant communities themselves, changes in the soil nutrient content and soil biota over time give rise to further potential mechanisms of strengthening biodiversity effects. Long-term experiments show that the plant diversity effect on soil microbial biomass (Eisenhauer et al. 2010, Habekost et al. 2008, Lange et al. 2015, Lange et al. 2014, Strecker et al. 2016, Thakur et al. 2015) as well as the abundance and diversity of soil micro-, meso- and macrofauna (Eisenhauer et al. 2011a, Eisenhauer et al. 2011b) developed only after a time lag of four years in the Jena Experiment. As these different soil organisms play key roles in many ecosystem functions related to decomposition, nutrient cycling, and plant growth, observed shifts in soil communities were suggested to have feedback effects on plant community composition and biomass production (Eisenhauer 2012, Eisenhauer et al. 2012). Negative plant-soil feedback effects at low plant diversity (Latz et al. 2012,

119 Latz et al. 2016, Maron et al. 2011, Schnitzer et al. 2011) or facilitation at high plant diversity via biotic
120 interactions with mutualistic partners (Latz et al. 2012, Wagg et al. 2015, Wright et al. 2017) are
121 alternative mechanisms for positive and increasing BEF relationships (Eisenhauer et al. 2012).
122 Furthermore, there is evidence of increased soil-C and -N storage by increased root biomass and
123 increased microbial activity in diverse compared with species-poor plant communities (Fornara and
124 Tilman 2008, Lange et al. 2015, Mellado-Vázquez et al. 2016, Steinbeiss et al. 2008). High resource
125 availability is one potential mechanism leading to increased function at high plant diversity, as
126 indicated by the significance of soil-C content in a meta-analysis of grassland and forest biodiversity
127 experiments (Guerrero-Ramírez et al. 2017) and by increased N content in a manipulative study
128 (Fridley 2002).

129 It is, however, difficult, to separate effects of community age from effects caused by particular
130 calendar years, especially for ecosystem functions that have been measured with few temporal
131 replicates. General temporal trends, such as globally increasing N deposition (Galloway et al. 2004),
132 increasing greenhouse gas emissions (IPCC 2014), warming climate or specific weather conditions
133 during the observation period of biodiversity experiments, also might have contributed to the
134 observed effects. Elevated atmospheric CO_2 concentrations and N enrichment increase productivity
135 at least in experiments of short duration (Fridley 2002, He et al. 2002, Isbell et al. 2013, Reich et al.
136 2001). Warming can increase aboveground productivity at high plant diversity, as has been shown in
137 short-term experiments and also after several years (Cowles et al. 2016; but see De Boeck et al. 2008).
138 Moreover, extreme climate events, such as droughts and floods, have been shown to influence plant
139 productivity (Isbell et al. 2015) and BEF relationships (Wright et al. 2015). Consequently, gradual and
140 abrupt changes in environmental conditions can contribute to strengthening effects of plant diversity
141 on productivity.
142 To account for potential confounding effects of calendar years within the study period, we
143 reestablished plant communities of a long-term biodiversity experiment (the Jena Experiment) in a
144 split-plot experiment with new seeds, and old or new soil in 2016 and compared measurements in

these young plant communities with measurements taken in the same calendar year in the old communities set up in 2002 (Roscher et al. 2004). Together, our setup comprises three treatments varying in the age of the plant and soil communities and forms the so-called ΔBEF Experiment (DELTa-BEF; short for DEterminants of Long-Term Biodiversity Effects on Ecosystem Functioning Experiment). The old communities of the Jena Experiment form one treatment and have a history of plot-specific changes in biotic and abiotic soil properties as well as in plant community composition over time as outlined above. The second treatment comprises young communities without plot-specific soil and plant history, because soil was replaced by arable soil from an adjacent crop field similar to the conditions at the beginning of the Jena Experiment in 2002 (Roscher et al. 2004), and new plant seeds were derived from commercial suppliers. In a third treatment, we kept the old soil with its plot-specific history but reseeded new plant communities. With this experimental setup, we thus could study the roles of soil age and plant-community age independent of confounding effects of particular calendar years by comparing the different treatments within calendar years and by comparing the newly established plot (second treatment) with the original Jena Experiment of the same age but established and measured 14 years earlier. We hypothesized that young communities without a shared history of soil and plant communities would show the weakest BEF relationship similar to results of short-term BEF experiments (Figure 1A; Reich et al. 2012). Moreover, we hypothesized that old communities with a shared soil as well as plant history would show the steepest BEF relationship, while the treatment with old plot-specific soil history but young plant communities would have an intermediate relationship. Here, we describe the establishment of the experiment and present first treatment effects on the BEF relationship by focusing on plant biomass production and selected soil properties as ecosystem functions.

Methods

Study site

We established the ΔBEF Experiment on the main plots of the Jena Experiment. This large biodiversity experiment was established in 2002 in an alluvial plain of the Saale river in Jena (latitude 50.95, longitude 11.62, altitude 130 m a.s.l., MAP 610 mm, MAT 9.9°C (Hoffmann et al. 2014)) and is described in detail in Roscher et al. (2004). The soil is an Eutric Fluvisol (IUSS Working Group WRB 2015), which had been managed as a highly fertilized arable field for growing vegetables and cereals for at least four decades prior to the establishment of the Jena Experiment. Sixty grassland plant species typical for semi-natural mesophilic grassland communities of this region served as species pool to establish the experimental plant communities. Seeds were ordered by commercial seed suppliers. Plant communities varying in species richness (1, 2, 4, 8, 16, and 60 species) and functional group richness (1, 2, 3, 4 functional groups: legumes, grasses, small herbs, and tall herbs) were randomly assembled from the respective functional groups and sown in May 2002 in four blocks (according to a gradient in edaphic across the field site)(Roscher et al. 2004). The seed density of 1000 viable seeds per m^2 was equally divided among the species in each mixture and adjusted to germination rates from standard laboratory tests. Since 2002, the colonization of non-target plant species in the experimental plots has been precluded by hand weeding two to three times per year. Weeding has been shown not to compromise observed plant diversity effects on ecosystem functions (Weisser et al. 2017). Further management includes mowing and removal of the mown biomass twice per year (at peak biomass in late spring and late summer) according to the traditional management of such grasslands in the study region.

The ΔBEF Experiment

In May 2016, i.e. 14 years after the establishment of the Jena Experiment, we set up three subplots on all 80 plots of the main plant diversity experiment (two monoculture plots had been given up in the meantime due to very low cover (<10%); (Weisser et al. 2017)). These subplots have a size of 1.5 x 3 m, and the location of each subplot was a random choice out of five potential locations at the edge of each long-term plot (Figure 1 D, E). This design maintained the original plots and allowed the

continuation of long-term time series measurements. The three subplots forming the ΔBEF Experiment vary in soil history (new soil without soil history, -SH vs. old soil of the single experimental plots since 2002 with soil history, +SH; Table 1) and in plant community history (newly established plant communities seeded in 2016 without plant history, -PH vs. old plant communities sown in 2002 with plant history, +PH) . We completely excavated the soil and plant layer of the treatment without soil history and without plant history (-SH -PH; Figure 1B) to a depth of 30 cm (which represents the main rooting zone in the Jena Experiment (Bessler et al. 2009, Ravenek et al. 2014) and replaced it by soil of an adjacent arable field of the same soil layer. Similar to the conditions of the field site of the Jena Experiment, this new soil also had been highly fertilized and used for growing of vegetables and cereals for decades. Accordingly, the arable field soil had properties that were in the middle of the range of the soil properties determined for the plots of the Jena Experiment in 2002 (pH 7.3; C_{org} 20.5g kg^{-1}; N_{tot} 2.3 g kg^{-1} compared with pH 7.1-8.4; C_{org} 5-33g kg^{-1}; N_{tot} 1-2.7g kg^{-1} in 2002; Roscher et al. 2004).

For the second treatment with soil history but without plant history (+SH -PH), we removed the plant sod by using a digger, while keeping the soil of the respective plots (Figure 1C). To remove large roots in the soil in this treatment, we mixed and homogenized the soil to a depth of 30 cm with the digger. A lateral mixture of the soil types with their surroundings was prevented by the installation of plastic sheets (PE-HD, 1 mm thick) as soil barriers of the top 0-30 cm of the soil in the -SH -PH and +SH -PH treatments. The plastic sheets were flexible, which facilitated the installation, and were insusceptible to erosion. We recompacted the soil of both treatments using a vibrating plate before the seeding of plant communities.

As done in 2002, we sowed the same plot-specific plant species mixtures on 23 May 2016 by hand in the two treatments with young plant communities (-SH -PH and +SH -PH) with a density of 1000 germinable seeds per m^2, equally divided among the target species. Therefore, we ordered seed material from commercial suppliers in 2016, conducted germination tests under standardized

conditions, and adjusted the sown seed numbers for each species according to their respective germination rates closely replicating the protocols used during the establishment of the original Jena Experiment (Roscher et al. 2004). The seed material had no plot-specific coexistence and interaction history (-PH). Seeds were mixed with a handful of site-specific sieved and dried soil to guarantee equal seed distributions across the subplot.

Our third treatment (+SH +PH) represents an undisturbed subplot of each experimental plot of the Jena Experiment (Figure 1D). The respective old communities have a shared soil and plant history since 2002. No new seeds were sown into the plots of this third treatment.

Measurements

Before sowing of the ΔBEF Experiment, we determined the seed bank to estimate the number of viable seeds of the target communities per area. For each experimental block, we took three composite samples each consisting of four soil cores (diameter: 5.7 cm, depth: 5 cm) at different positions of the homogenized arable field soil prior to refilling the subplots (to assess the viable seed bank of (-SH -PH). In the (+SH +PH) treatment, we collected a composite sample of four soil cores to 5 cm depth (diameter: 5.7 cm) along a transect in each plot. In the (+SH -PH) treatment, we took a composite sample consisting of five soil cores (diameter: 1.8 cm, depth: 30 cm). This was done because the soil of this treatment was homogenized down to this depth to remove large roots, which might have redistributed seeds across soil layers. Furthermore, we removed for this treatment the upper ~2 cm of the soil cores from our samples, because this soil layer was also removed together with the grass sod during the setup of this treatment. To assess the number of viable seeds per m^2, we accounted for this difference in soil sampling by multiplying the sampled area (5 times a core with a diameter of 1.8 cm) by factor 6 (6 times 5 cm depth increments = 30 cm sampling depth). Overall, the sampling volume for determining the soil seed bank was low, but could not be increased to avoid too much disturbance on the subplots.Soil samples were transported to the laboratory and stored at 4°C until

processing. In order to reduce the soil volume and establish better conditions for seed germination, we applied the bulk reduction method (ter Heerdt et al. 1996). Each soil sample was washed through a cascade of two sieves (2 mm mesh size to eliminate roots and rhizomes, and 0.2 mm mesh size to remove fine-soil without seeds). Afterwards, samples were spread out in thin layers (< 5mm) in trays filled with a heat-sterilized sand-soil mixture (1:1). The trays were placed in an open greenhouse with a roof, which automatically closed during rain, and cultivated at ambient temperatures to promote germination by natural daily temperature fluctuations. They were covered with a fine gauze to prevent contamination by wind-borne seeds. In addition, control trays containing only the sterilized sand/soil mixture were installed to control for contamination originating from the substrate. Trays were watered regularly and checked for emerging seedlings. Seedlings were determined to species level and then removed. Although most seedlings emerged in the first weeks, the experiment was continued until next spring to break the dormancy of some species with chilling. The community-level seed bank was summarized from all emerging seedlings of species belonging to the target community of the respective plot.

After sowing of the plant communities in the field, we counted the number of emerging seedlings in a randomly placed quadrat of 50 x 50 cm. We did the first count between 22 and 24 June 2016, and a second count between 4 and 5 April 2017, i.e. one and 11 months after the sowing, respectively. The second count mainly served to check for seedling emergence of species with seeds that require a chilling period for dormancy breaking and do not appear at all or at very low numbers in the first growing season when sowing in spring (Heisse et al. 2007).

We counted individuals of each species in a randomly placed quadrat of 50 x 50 cm from 27 September to 11 October 2016 to determine plant density. In contrast to -SH -PH and +SH -PH, it was challenging to visually separate single individuals in the +SH +PH treatment. The numbers, therefore, represent an estimate of individuals in this treatment.

Aboveground biomass was harvested from 29 May to 6 June 2017 (12 months after sowing of the new plant communities) and 29 to 31 August 2017 (15 months after sowing the new plant communities) before mowing the entire experimental plots. The vegetation was clipped 3 cm above soil surface within two randomly placed frames (20 x 50 cm) per subplot and sorted by sown species, while non-target weeds, dead and not-identifiable plant material were removed. After drying at 70°C for 48 h, samples were weighed and dry mass of target species was summed to community-level values for each subplot. Annual biomass was calculated as the sum of the two individual harvests. We measured leaf area index in every subplot in the same way as for the soil barrier experiment before each biomass harvest (LAI-2200C, LI-COR Biosciences, Lincoln, Nebraska, USA).

We used a decimal scale (based on Londo 1976) to estimate plant species-specific cover on the whole subplot area prior to every biomass harvest. Cover data were supplemented by similar data collected in 2003 for the original Jena Experiment. This represents an initial treatment without soil history and without plant history comparable to the new -SH -PH treatment but differing in calendar year (2003 vs. 2017; see van Moorsel et al. (2018) for further justification of this approach). We used cover data to calculate realized target species richness, species evenness, and the dissimilarity between the treatments based on differences in species abundances (see below).

Two soil sampling campaigns were conducted to determine abiotic and biotic soil properties. Soil organic C concentration was determined in all treatments of the experiment in spring of 2016, so immediately after the soil treatments had been established but prior to sowing. Soil organic C in the old +SH +PH treatment was determined using a split-tube sampler (4.8 cm diameter), by taking three soil cores per plot to a depth of 30 cm. Soil cores were segmented into 5 cm-depth sections and pooled per depth sections and plot (Lange et al. 2015, Steinbeiss et al. 2008). Soil of the young treatments (-SH -PH, +SH -PH) was sampled using a soil corer (2.0 cm diameter), by taking ten soil cores per plot to a depth of 5 cm. Soil cores were pooled per plot, and the soil was then dried, sieved, and milled. Subsequently, total C was determined by combustion with an elemental analyser at 1,150°C

294 (Elementaranalysator vario Max CN, Elementar Analysensysteme GmbH, Hanau, Germany). Inorganic
295 C concentration was measured after oxidative removal of organic C for 16 h at 450°C in a muffle
296 furnace. Finally, organic C concentration was calculated as the difference between total and inorganic
297 C (Lange et al. 2015, Steinbeiss et al. 2008) for the top 5 cm soil layer.
298 Soil sampling for the other soil properties took place in June 2017 (13 months after setting up of the
299 treatments). We pooled six soil cores (4 cm in diameter, 5 cm depth) per treatment, homogenized
300 them in the field, and stored them at 4°C until further processing. We sieved one subsample of the
301 soil at <2mm and stored it eight months at –20°C for microbial analysis (see below). The other
302 subsample was used to determine concentrations of soil mineral nitrogen concentrations (NH_4^+-N and
303 NO_3^--N) and gravimetric soil moisture content. We removed visible roots and stones in this subsample
304 and extracted field-fresh soil samples with 1 M KCl solution within 24 h of sampling and filtrated the
305 samples after 30 min of shaking. The concentrations of NH_4^+-N and NO_3^--N were determined
306 colorimetrically by continuous flow analysis (SanPlus, Skalar, NL) using KCl as a blank. The
307 concentration of soil mineral nitrogen was calculated as the sum of the NH_4^+-N and NO_3^--N
308 concentrations.

309 Soil microbial-biomass carbon of approximately 5 g soil (fresh weight) was measured using an O_2-
310 microcompensation apparatus (Scheu 1992). We measured the microbial respiratory response after
311 incubation of ~24 h at hourly intervals for at least 8 h at 20°C. Substrate-induced respiration was
312 calculated from the respiratory response to D-glucose for 10 h at 20°C (Eisenhauer et al. 2013). We
313 added Glucose according to preliminary studies to saturate the catabolic enzymes of microorganisms
314 (4 mg g^{-1} dry weight dissolved in 400 µl deionized water; Strecker et al. 2016). The mean of the lowest
315 three readings within the first 10 h (between the initial peak caused by disturbing the soil and the peak
316 caused by microbial growth) was taken as maximum initial respiratory response (MIRR; µl O_2 g^{-1} soil
317 dry weight h^{-1}) and microbial biomass (µg C g^{-1} soil dry weight) was calculated as 38 × MIRR (Beck et
318 al. 1997). The values represent the biomass of all active bacteria and fungi in the soil.

Another subsample of the soil was analyzed for the soil nematode community as an indicator group of soil food web structure. Soil nematodes were extracted from 25 g of fresh soil using a modified Baermann method (Ruess 1995) and fixed with 4% formaldehyde solution at 65°C. After counting, nematodes were determined to genus level (adults) or family level (juveniles) and assigned to herbivorous, bacterivorous, fungivorous, omnivorous, and predatory trophic groups (Bongers and Bongers 1998, Yeates et al. 1993).

Data analysis

We used cover estimates to calculate Simpson evenness *D* for each subplot by $D_k = \sum_{i=1}^{n} p_{ik}^2$, where p_{ik} is the relative abundance of species *i* in a subplot *k* and *n* is the number of sown species in the subplot.

To analyze differences in species compositions of the plant and the soil nematode communities, we calculated Bray-Curtis dissimilarities (BC) for all possible treatment combinations per plot. *BC* varies between zero (= identical composition regarding species presences and abundances in both treatments) and one (= no species in common) and is calculated as $BC_{ik} = \frac{\sum_{i=1}^{n} |p_{ij} - p_{ik}|}{\sum_{i=1}^{n} |p_{ij} + p_{ik}|}$, where p_{ik} is the relative abundance of species *i* in the two subplots *j* and *k*. We used species-specific cover estimates as abundance measures for the plant communities. In case of the soil nematode communities, we used nematode counts per soil weight as abundance measures and calculated dissimilarities including the whole nematode community and single trophic groups, respectively. Note that the interpretation of *BC* varies between plant and nematode communities in our experiment. In case of the plant communities, only abundance shifts and extinctions can cause deviations of *BC* from 0, while for the nematode communities it is also species turnover or species gains and losses that can contribute to differences among communities. We therefore additionally calculated the Jaccard index

based on presence-absence data according to $J = \frac{S_{jk}}{S_j + S_k - S_{jk}}$, with S_j as the number of taxa in subplot j, S_k as the number of taxa in subplot k, and S_{jk} as the number of shared taxa in both subplots. This index addresses the turnover of communities between two treatment pairs (Anderson et al. 2011). We computed *BC* and *J* with the function *vegdist* in the package *vegan* (Oksanen et al. 2017). We performed mixed models to analyze whether dissimilarities varied among treatment pairs. Additionally, we used nonmetric multidimensional scaling (NMDS) for the ordination of the subplots based on Bray-Curtis distances of the nematode communities using the function *metaMDS* in the package *vegan*

Data of aboveground biomass and cover estimates in 2003 supplemented our three ΔBEF treatments by the initial treatment (-SH -PH in 2003). This enabled us to separate the treatment effect into three contrasts for our statistical analysis. One contrast (year) separates the effect of the different calendar years (2003 vs. 2017 respectively), one contrasts the histories of plant communities (-PH vs. +PH), and one contrasts the soil histories (-SH vs. +SH, see Table 1). We fitted a series of mixed effects models for the annual aboveground biomass in 2017. Random intercept terms accounted for the nested structure of the ΔBEF Experiment, i.e. they consisted of block, plot nested in block, and treatment nested in plot and block. As fixed effects, we fitted species richness (as a log-linear term), followed by treatment and the interaction of both (overall model A). We separated the term treatment into contrasts for year, soil history, and plant history to test for their particular effects in additional models (model B-D). In model B, we ordered the treatment contrasts in the sequence year – soil history – plant history (Table S1). This model, therefore, is testing for the *pure effect of plant history*, since we accounted for the variance due to different calendar years and soil types first. The sequence of the treatment contrast in model C was year – plant history – soil history, and it therefore tests for the *pure soil history effect*. In model D, we tested for the *pure effect of different calendar years* by fitting the treatment contrasts as plant history – soil history – year. To achieve the assumption of normality of residuals, we log-transformed biomass values. We analyzed the response variables of soil and plant

community properties in the same way, with the only exception that we skipped the contrast for different calendar years, when we had only data of 2017 but not of 2003 available. We performed all analyses using R 3.3.3 (R development core team 2011).

The Soil Barrier Experiment

We were not able to install plastic sheets in the (+SH +PH) treatment without disturbing the communities and interfering with existing installations. To quantify this potentially confounding effect (e.g., higher water runoff in subplots with than without sheet installation), we set up an additional experiment to test for the effects of soil barriers on different ecosystem properties. We used eight abandoned plots of earlier experiments at the field site, which had not been weeded since 2010, and hence had well-established plant communities. In each of these plots, we established two subplots of the same size as in the ΔBEF Experiment and installed plastic sheets by using a soil trencher in one treatment. The other treatment was left without a soil barrier and served as control. We installed three transects of 20 cm width in parallel to the long edges of the subplots to determine potential edge effects. We measured soil moisture, leaf area index and aboveground biomass in parallel to the measurements in the ΔBEF Experiment to test for the effects of soil barriers and edge effects. A portable soil moisture probe (ML3 ThetaProbe, Delta-T Devices, Cambridge, UK) was used to determine the volumetric soil water content indirectly by time-domain-reflectometry in 6 cm depth at five replicate points along each transect. In two replicate frames of 20 x 50 cm per transect, aboveground biomass was clipped at 3 cm height. We measured leaf area index exactly as described for the ΔBEF Experiment in only one transect in the middle of the subplot. Single transects of only 20 cm width would have been too small for this measure. We performed linear mixed models to test for the effect of treatment (with or without soil barrier), transect and the interaction of both. None of the measured variables showed any significant effect of the soil barriers (soil moisture: $F_{1,7} = 2.79$, $P = 0.139$; leaf area index: $F_{1,7} = 0.369$, $P = 0.563$; aboveground biomass: $F_{1,7} = 0.38$, $P = 0.559$) in any

transect (soil moisture: $F_{2,220} = 0.02$, $P = 0.975$; aboveground biomass: $F_{1,16} = 1.18$, $P = 0.293$). Therefore, we assume that the results of the ΔBEF Experiment are not biased by soil barrier effects.

Results

Establishment of the treatments

Treatment effects on plant communities

The arable soil from the neighbouring agricultural field used for the establishment of the -SH -PH treatment did not contain any viable seeds from the experimental grassland species pool. The soil of the +SH -PH treatment with plot-specific history contained a viable seed bank comprising some of the target species of the respective plots. The removal of the plant communities and soil homogenization to 30 cm depth, however, reduced the number of viable seeds in the seed bank of this treatment compared with the undisturbed +SH +PH treatment (PH in Table 2, Figure 2A). The number of viable seeds in the seed bank increased with plant species richness in both treatments (SR significant, but PH x SR not significant in Table 2, Figure 2A). The number of target species seedlings, which emerged after sowing of the new experimental communities, was independent of the soil history (SH in model C not significant in Table 2, Figure 2B) and higher in the newly sown compared to the old plant communities (PH in model C in Table 2). This was not surprising considering that there was no sowing into the old communities. Furthermore, plant species richness increased the number of emerging seedlings, but this effect was stronger in old compared to the newly sown plant communities (PH x SR in model C in Table 2, Figure 2B), also when we compared the plant communities independently of the soil type (PH x SR in model B in Table 2, Figure 2B). The soil itself did not change the diversity effect in the newly sown plant communities (SH x SR in model C not significant, Figure 2B). Mature plants had a higher density in the old compared to the newly sown plant communities (PH in model C in Table 2, Figure 2C) and this result was independent of the soil type (PH in model B in Table 2, Figure 2C). The density

of plant individuals increased with plant species richness and this effect was stronger in the old compared to the new communities (PH x SR in models B and C in Table 2). The soil itself did neither alter the density nor the plant species richness effect on the density of mature plant individuals (SH and SH x SR in model C not significant, Table 2, Figure 2C).

Leaf area index was higher in the -SH -PH treatment without any specific soil or plant history compared to +SH -PH and +SH +PH treatments (SH in model B and C in Table 2, Figure 2D). This effect of soil history was independent of community species richness (SH x SR interaction in model C not significant, Table 2). However, old communities had a much higher leaf area index at high plant species richness compared to young communities on old soil (PH x SR in model B not significant, Table 2).

Plant species numbers in the newly sown treatments established well after sowing and showed no significant difference between the two soil treatments (SH in model C not significant, Table 2, Figure 2E). Old communities (+SH +PH) had a lower realized plant species richness compared to young communities (PH in model C in Table 2, Figure 2E), although in 2003 they established with a similar realized species richness compared to the new -SH -PH treatment (Y in model D not significant, Table 2). The lower realized number of plant species in old than in new communities was more pronounced at high sown species richness (significant PH x SR in models B and C in Table 2, Figure 2E) and independent of the soil history (SH x SR in model C not significant, Table 2). The calendar year did not affect the plant species richness effect on the realized number of species (Y x SR of model D in Table 2 not significant).

The evenness of the communities overall increased with plant species richness, but the strength of this relationship varied among treatments (T x SR of model A in Table 2, Figure 2F). Old plant communities had a lower evenness at high species richness than young communities (significant PH x SR in model C, Table 2, Figure 2F). Pure soil or plant history did not change the plant diversity effect on community evenness (PH x SR in model B and SH x SR in model C not significant, Table 2).

Furthermore, the evenness of newly sown communities was independent of the calendar year of establishment (Y and Y x SR in model D in Table 2, Figure 2F).

Although we sowed the subplots originally with the same plant species composition and proportions, the realized composition of the plant communities differed among the treatments (Treatment pair: $F_{5,300} = 29.25$, $p < 0.001$), especially in diversity treatments of more than two species (Figure 3, SR: $F_{1,57} = 66.83$, $p < 0.001$; Treatment pair x SR: $F_{5,300} = 4.04$, $p = 0.002$). The difference can be explained by the age of the plant communities (Bray-Curtis-distance of treatment pairs with same vs. different age of plant communities: $F_{1,304} = 137.36$, $p < 0.001$; two-way interaction with SR: $F_{1,304} = 15.66$, $p < 0.001$). This is indicated by the higher Bray-Curtis-dissimilarity in treatment pairs of different plant history (between +SH +PH and +SH -PH, green in Figure 3A; between +SH +PH and -SH -PH, grey in Figure 3A; between +SH +PH and 2003, grey in Figure 3B) than in treatment pairs with similar plant history (between -SH -PH and +SH -PH, brown in Figure 3A; -SH -PH and 2003, orange in Figure 3B; +SH -PH and 2003, blue in Figure 3B). Contrary to this, the history of the soil caused minor compositional shifts of plant communities (treatment pairs with same vs. different soil history: $F_{1,304} = 0.48$, $p = 0.489$; two-way interaction with SR: $F_{1,304} = 0.001$, $p = 0.975$). Accordingly, Bray-Curtis dissimilarity in treatment pairs of different soil history (between -SH -PH and +SH -PH, brown in Figure 3A) was low. Furthermore, compositional shifts due to the calendar year (Bray-Curtis dissimilarity between -SH -PH and 2003, orange in Figure 3B) were lower compared to the compositional shifts due to the age of the plant communities (Bray-Curtis dissimilarity between +SH -PH and +SH +PH as well as between +SH +PH and 2003, grey in Figure 3A and B).

Treatment effects on soil properties

Soil abiotic and biotic properties were measured from the same soil samples to test for treatment effects and to get baseline data on resource- and biotic interaction-based hypotheses underlying BEF relationships. Soil moisture differed significantly among treatments and along the plant diversity

gradient (T in model A in Table 3, Figure 4A). Soil moisture was not affected by soil history but by plant history as shown by the differences between the +SH -PH and +SH +PH treatments (PH in model B in Table 3). Additionally, the increase in soil moisture across the plant diversity gradient was much higher in old (+SH +PH treatment) than in new communities (PH x SR in model B and C, Table 3, Figure 4A). Plant-available N (nitrate and ammonium) concentrations in the soil were much higher in the +SH +PH treatment than in the newly seeded treatments without plant history (PH in model C in Table 3, Figure 4B). Beyond this, we did not find any significant differences in soil mineral N concentrations between the +SH -PH and -SH -PH treatments (SH in model C not significant, Table 3), indicating the absence of a pure soil history effect (Table 3). We did not find a significant plant species richness effect on soil mineral N concentration (SR and interactions with SR not significant in all models, Table 3, Figure 4B). Soil organic carbon content differed in all treatments and across the plant diversity gradient (T in model A in Table 3, Figure 4C). As expected, we found a lower soil organic carbon content in the new compared to the old communities (PH in model C, Table 3, Figure 4C). The organic carbon content of the +SH -PH treatment with old soil was even lower than that of the -SH -PH treatment with new soil (SH in model C, Table 3, Figure 4C). The positive effect of plant species richness on the soil organic carbon content of the old communities was lower in the new communities (PH x SR in model C, Table 3, Figure 4C) and was not affected by the soil history (SH x SR in model C not significant, Table 3).

Soil microbial biomass and nematode abundance increased with plant species richness and were affected by the treatments (Table 3, Figure 5A, B). Soil microbial biomass was highest in the +SH +PH treatment with shared soil and plant history (PH in model C significant,Table 3). Also independent of the soil history, plant history significantly increased soil microbial biomass (PH in model B in Table 3). Interestingly, soil microbial biomass was lower in +SH -PH compared to the -SH -PH treatment, resulting in a negative effect of pure soil history (SH in model C in Table 3, Figure 5A). The treatments only marginally changed the plant diversity effect (Table 3).

Similar to microbial biomass, nematode abundance was highest in the +SH +PH treatment, indicating a positive plant history effect (PH in model C in Table 3, Figure 5B). The lowest nematode abundance was found in the -SH -PH treatment (SH in model B in Table 3). Plant species richness had a stronger positive effect on nematode communities in old than in new soil (SH and SH x SR in model B in Table 3). Non-metric multidimensional scaling of the nematode communities revealed differences of nematode communities among treatments (Figure 5C). Differences in the composition of the nematode communities, measured as Bray-Curtis dissimilarities among all treatment pairs, were not significantly affected by treatment or plant species richness (Figure 5D, SR: $F_{1,39} = 0.09$, $p > 0.2$; Treatment pair: $F_{2,68} = 1.33$, $p > 0.2$). However, the treatment effect on the nematode community composition depended on the trophic groups of the nematodes (Figure 5D, trophic group x treatment pair interaction: $F_{6,336} = 7.53$, $p < 0.001$). The dissimilarity between the soil history treatments was most pronounced within the community of plant-feeding nematodes, whereas dissimilarities due to plant history did not change between the other trophic groups. Ignoring the abundances of the nematode taxa by calculating the Jaccard index, we found similar results (Figure S1, Treatment pair: $F_{2,68} = 4.83$, $p = 0.01$). The composition of plant-feeding nematodes was more dissimilar between the two soil treatments (-SH -PH and +SH -PH treatment) compared with the treatment pairs including different plant history (Figure S1, Treatment pair: $F_{2,68} = 3.40$, $p = 0.04$). The dissimilarity of the total nematode community depended on the treatment pair (Figure S1, Treatment pair: $F_{2,68} = 6.35$, $p = 0.003$) and was highest between the -SH -PH and +SH -PH treatments. Our results indicate that the composition of nematode taxa changes more strongly with soil history than with plant history, especially within the trophic group of plant-feeding nematodes.

Treatment effects on the plant diversity–productivity relationship

Aboveground biomass production differed significantly among treatments (Figure 6A, model A in Table 4). The treatments with soil history (+SH -PH and +SH +PH treatments) had lower aboveground

biomass compared to the -SH -PH treatment without soil history (SH in model B in Table 4, Figure 6A). Comparing the young plant communities only (-SH -PH and +SH -PH treatment) revealed that soil history clearly decreased aboveground biomass productivity (SH in model C in Table 4, Figure 6A). Community biomass production of the -SH -PH treatment in 2017 was as high as the community biomass production in 2003 (Y in model D in Table 4 not significant, Figure 6A). In general, aboveground biomass production increased with plant species richness in all treatments (SR in all models in Table 4, Figure 6A). However, the strength of species richness effect differed among treatments (T x SR in model A in Table 4, Figure 6A). As hypothesized, we found the steepest slope of the plant species richness-productivity relationship in the old communities with shared plant and soil history (+SH +PH; PH in model D in Table 4, Figure 6B). Separating the treatment effect revealed that pure soil history did not significantly alter the species richness effect on aboveground biomass production (SH x SR in model C not significant, Table 1, Figure 5B), and that the calendar year in young communities also did not change the plant diversity effect on aboveground productivity (Y x SR in model D not significant, Table 4, Figure 6B). This was also true when we standardized the aboveground biomass by mean biomass of the respective treatment to compare the slopes on the same scale (Table 4, Figure 6B).

Discussion

Here, we present the hypotheses, setup, and first results of a new experiment testing the mechanistic underpinnings of strengthening plant diversity effects over time in a long-term biodiversity experiment. The new ΔBEF Experiment is a cross-sectional approach and complements the long-term grassland biodiversity platform of the Jena Experiment by the opportunity to separate community age effects from effects of the respective calendar year with its specific environmental conditions. Moreover, effects of plant community-specific soil history can be tested, and long-term plots of the

538 Jena Experiment can serve as a reference for conditions of undisturbed communities with soil and
539 plant history.

540

541 *Establishment of the treatments*

542 Our analysis of the seed bank revealed that the plant communities in the -SH -PH treatment can be
543 considered as naïve, because no seeds of the experimental species pool were found in samples of the
544 arable soil. This was expected, since the management of arable fields aims to exclude grassland species
545 by conventional crop rotation, pesticide application, and fertilization for agricultural production.
546 Moreover, the plant communities of the +SH -PH treatment can also be considered as nearly naïve.
547 Although we found viable seeds in the seed bank of this treatment (on average 304 viable seeds per
548 m^2), the number was only 25% as high as in the +SH +PH treatment (on average 1284 viable seeds per
549 m^2). Furthermore, the density of emerging seedlings observed in the field in the +SH -PH treatment
550 did not differ from the completely naïve communities of the -SH -PH treatment. The treatments
551 without plant history (-SH -PH and +SH -PH) were also more similar to each other than to the old
552 communities of the +SH +PH treatment; they had a lower density of mature individuals, higher realized
553 plant species, higher evenness, and differed little in their plant species composition. Our comparisons
554 with the plant communities in 2003 (Figures 2E, 2F, 3B) confirmed that these differences between the
555 new and old plant communities reflect the compositional shifts over time in the old communities as
556 observed previously (Allan et al. 2011, Roscher et al. 2011b, 2013). Moreover, the positive effects of
557 plant species richness on the number of emerging seedlings, plant density, leaf area index, and on
558 compositional shifts were more pronounced in old than in new plant communities. Notably, in this
559 initial assessment of the ΔBEF Experiment, we did not find any significant differences between new
560 plant communities established on old or new (arable) soil. Future studies will explore if soil history
561 effects emerge over time.

562 Differences between the new (-SH -PH and +SH -PH treatments) and the old communities (+SH +PH
563 treatment) were not only present at the level of plant-community properties but also in the soil. Soil
564 moisture of the upper soil layer was significantly lower in the new than in the old communities (Figure
565 4A). Similarly, Fischer et al. (2019) reported lower topsoil water content in the first years of the Jena
566 Experiment than in later years. They explained variations in the water content of the topsoil by the
567 increasing leaf area index of the plant canopy over time. However, in the present study, leaf area index
568 was higher in the -SH -PH than in the +SH -PH treatment (Figure 2D), despite the two treatments having
569 a similar plant density (Figure 2C), and the soil moisture content was similarly low (Figure 4A). We
570 therefore suggest that higher water evaporation from the soil due to lower plant density in both newly
571 established treatments is more likely responsible for the decreased soil moisture content than are
572 differences in leaf area index. Leaf area index can still be high despite low plant densities, when the
573 vegetation is closed at higher canopy layers. In this case, evaporation could be high, despite the
574 shading effects of the canopy. However, we cannot exclude the possibility that physical factors
575 operated as well. The establishment of the treatments -SH -PH and +SH -PH involved soil disturbance
576 and homogenization similar to tillage, which is known to affect many physical soil properties (Holland
577 2004), such as increased destruction of soil aggregates (Six et al. 2000), low soil structure, and thus a
578 lower retention of soil moisture. Reduced soil moisture as well as preceding leaching might also
579 explain the much lower concentration of plant-available N in those treatments compared to +SH +PH.

580 The described absence of early soil-history effects was unexpected, particularly for soil mineral N
581 concentrations, since the arable soil of the -SH -PH treatment was derived from a conventional
582 agricultural field with fertilizer application. In the early years of the Jena Experiment, mineral N
583 concentrations increased shortly after land-use change from arable field to experimental grasslands
584 and then decreased again over time due to discontinued fertilization, increasing aboveground N
585 storage in plant biomass (Oelmann et al. 2011, Oelmann et al. 2007), and continuous removal of plant
586 biomass by mowing (Leimer et al. 2014). However, there was a short lag phase before the initial
587 increase in mineral N concentrations. It is possible that the current phase of the -SH -PH treatment

588 corresponds to this lag phase where soil mineral N remained unchanged before this increase
589 (Oelmann et al. 2011). Furthermore, an overall decrease of average nitrate concentrations with
590 increasing plant diversity has been observed in the Jena Experiment ((Lange et al. 2019) this issue),
591 which was not observed in the current study in any treatment. We found equally low concentrations
592 of plant-available N in the +SH -PH treatment as in -SH -PH, and higher concentrations of plant-
593 available N in the +SH +PH treatment. This may point towards a strong effect of the plant community
594 on soil mineral N, with younger, still establishing communities having a higher demand for N during
595 initial growth, whereas more mature, established communities may require less N and therefore
596 deplete the pool of available N in the soil to a lesser degree. This also might be reflected in the higher
597 biomass in the -SH - PH treatment, where pre-seeding N in the soil was presumably higher than in the
598 +SH -PH treatment. This would be a subject to be investigated further in future studies.

599 Further, the mixing of the soil during the setup of the experiment might contribute to the low mineral
600 N concentrations in the +SH -PH treatment compared to the +SH +PH treatment. Higher SOM
601 accumulation in old soil provides additional sources for N mineralization and may, therefore, impact
602 on mineral N concentrations in the soil (Rosenkranz et al. 2012). The increase in SOM over time in
603 undisturbed soils was shown to be strongest in the upper 5–10 cm of the soil but much less
604 pronounced at deeper soil layers (Gerzabek et al. 2005, Steinbeiss et al. 2008). Mixing might have
605 caused a homogenization of SOM in the top 30 cm of the soil and therefore reduced the SOM content
606 in upper soil layers, where we measured soil mineral N concentrations. This was further underlined by
607 the low soil organic carbon content in the +SH -PH treatment, but high soil organic carbon content in
608 the +SH +PH treatment.

609 Soil biota showed significant responses to the soil-history treatment, but these were additionally
610 affected by plant removal. We expected a much lower microbial biomass in the arable soil of -SH -PH,
611 since microbial biomass was shown to increase only four years after converting an agricultural field to
612 an experimental grassland in the Jena Experiment, and this increase was largely restricted to high-

613 diversity plant communities (Eisenhauer et al. 2010, Eisenhauer et al. 2011b, Strecker et al. 2016).
614 However, microbial biomass in the +SH -PH treatment was not higher but even lower than in the arable
615 soil of the -SH -PH treatment. As we lack information on soil microbial community composition, we
616 can only speculate about the reasons for this finding. One potential explanation might be that more
617 autochthonous and less resilient microbial communities in the experimental grasslands were heavily
618 disturbed by the topsoil mixing, while microbial communities in the agricultural soils, which may be
619 better adapted to disturbance showed higher resilience due to the dominance of fast-growing
620 microbial taxa (Eisenhauer et al. 2010, Eisenhauer et al. 2017). Future research has to explore soil
621 microbial community changes in the different soil and plant history treatments along the plant
622 diversity gradient, and potential feedback effects of these microbial communities on plant growth.
623 Strong effects of soil history from the Jena Experiment on plant rhizosphere microbiomes have been
624 found in a recent glasshouse experiment (Schmid et al. 2019).

625 Soil nematode abundance was reduced in both newly established treatments in comparison to the old
626 communities of the +SH +PH treatment, but also the two new soil treatments differed marginally: the
627 nematode abundance in the new communities (-SH -PH) was lower than in the old soil (+SH -PH).
628 Exploring the functional composition of soil nematode communities provided further insights by
629 showing that the low levels in the -SH -PH treatment were due to a drop in plant-feeding nematode
630 abundances.

631 Our findings are in line with other studies reporting that plant removal causes a decrease in
632 abundances of microorganisms, micro- and mesofauna (Convey and Wynn-Williams 2002, Hirsch et
633 al. 2009, Mikola et al. 2014, Stevenson et al. 2014, Wardle et al. 1999). Plant removal and soil
634 homogenization can induce reductions in soil organic matter content or in releases of recently fixed
635 carbon by plant roots and therefore diminish important food resources of soil organisms. Mikola et al.
636 (2014) explained the decrease of microbes, nematodes, and collembolans after plant removal in arctic
637 soils with a decrease in carbon releases by plant roots rather than decreases in soil organic matter.

Even replanting could not reverse this plant-removal effect in the short-term. In our study system, soil microbial biomass increased in response to an increase in root biomass (Eisenhauer et al. 2010), while the amount of root exudates was a poor predictor of microbial biomass ((Mellado-Vázquez et al. 2016) but see (Eisenhauer et al. 2017)). In the ΔBEF Experiment, root biomass in the new communities (-SH -PH and +SH -PH treatments) was indeed lower than in the old communities (+SH +PH treatment) in the same study year as we analyzed soil microbial biomass and nematode abundance (D. Francioli, data unpublished). Thus, reduced root biomass and root-derived plant inputs to the soil might explain the lower abundances of soil biota, in addition to soil disturbance and homogenization.

However, while both, the soil treatment and the plant treatment affected the abundance of soil biota, the composition of soil nematodes responded only to the soil treatment (Figure 5C, D). Plant removal in the above-mentioned studies also caused stronger effects on abundance than on the composition of micro- and mesofauna (Mikola et al. 2014) and microbial communities (Hirsch et al. 2009). The effect of the soil history was most pronounced when we analyzed the composition of the plant-feeding nematodes only because +SH -PH and +SH +PH treatments had a more similar community composition of plant-feeding nematodes than the arable soil treatment (-SH -PH; Figure 5D and Figure S1). Overall, the soil nematode-community analysis indicates that the soil treatments of arable versus plot-specific soil differ concerning the composition of the soil biota, while the plant treatment (plant removal, soil disturbance and reseeding) changed the abundance of soil biota. While the ΔBEF Experiment allows to study soil history effects by comparing the -SH -PH and the +SH -PH treatments, it is not possible to separate effects of plant history alone by the comparison of +SH -PH and +SH +PH, because we manipulated biotic and abiotic soil properties in addition to plant community age. Plants and their inputs drive soil communities (Leff et al. 2018), which then feedback to the plants and therefore effects of plant age or plant community age can never be separated sharply from soil feedback effects. However, such soil feedback effects, which are tightly linked to the plants that they would not operate without them, can be regarded still as plant history effects, while changes in soil properties due to the physical soil disturbance could not. The treatment of +SH +PH therefore contrasts the other two

664 treatments of the ΔBEF Experiment by having a shared history of the soil and plant communities and
665 hence, serves as control.

666

667 *Treatment effects on the plant diversity-productivity relationship*

668 In line with previous work (Guerrero-Ramírez et al. 2017, Marquard et al. 2009, Reich et al. 2012) and
669 our expectations, we found a positive relationship between plant species richness and aboveground
670 biomass production in all treatments. Confirming our hypothesis, we found the weakest slope of this
671 BEF relationship in the young communities (-SH -PH and +SH -PH treatments) and the steepest slope
672 in the old communities of the +SH +PH treatment with shared soil and plant history. Moreover,
673 according to our hypothesis, the weakest slope of the BEF relationship in the -SH -PH treatment
674 without plot-specific soil history and without plant history did not significantly differ from the slope in
675 2003, i.e. one year after the establishment of the Jena Experiment. Although the conditions in 2003
676 were similar to our study year in 2017 concerning the soil history (both, the 2003 communities and -
677 SH -PH treatment was established on arable soil) and the age and composition of the plant
678 communities, climatic conditions differed considerably. In summer 2003, an extreme heat wave
679 affected Central Europe (Beniston 2004, Zaitchik et al. 2006) and resulted in almost 2°C higher spring
680 and summer mean temperatures at our field site compared to the reference period of 1961-1990. The
681 year 2017 was less extreme with only warmer spring temperatures of less than 1°C higher than the
682 average of the reference period. Experimental studies revealed that warming can change the
683 magnitude and slope of plant diversity-productivity relationships (Cowles et al. 2016, De Boeck et al.
684 2008). However, this was not the case in the present study, when we compared communities of the
685 same age but under different climate conditions. The similar magnitudes and slopes of the plant
686 diversity-productivity relationship in the -SH -PH treatment and in 2003 suggest that community age
687 effects play a larger role in shaping the BEF-relationship than climatic or other environmental
688 conditions of the respective years. Therefore, we conclude that the treatments of the ΔBEF

Experiment are successful in mimicking long-term trends in a cross-sectional study, at least for plant productivity-related processes. Consequently, our ΔBEF Experiment enables us to study ecological mechanisms and processes of BEF relationships independently of the environmental conditions and to transfer those insights to explain long-term trends in BEF relationships. Future studies will investigate a range of different ecosystem functions in response to the treatments of the ΔBEF Experiment, because the change in the strength of diversity effects on functioning and the underlying processes have been shown to differ among ecosystem functions (Meyer et al. 2016) and have important implications for the multifunctionality of ecosystems (Giling et al. 2019, Manning et al. 2018).

Our second aim in this experiment was to test for the contribution of soil history or plant history to the strengthening BEF relationship. As outlined above, pure soil history effects can be separated by comparing the treatments -SH -PH and +SH -PH. Plot-specific soil history decreased overall aboveground biomass production, but in contrast to our expectations, it did not strengthen the plant diversity–productivity relationship. Grassland-soil history in our experiment changed the abundance of soil organisms and the composition of herbivorous nematodes, while soil abiotic conditions did not differ significantly. We speculate that effects of soil organisms may need some time to materialize, e.g. via shifts in soil nutrient status, plant community composition, and antagonistic and mutualistic interactions (Eisenhauer 2012, Kulmatiski et al. 2012). Notably, in the ΔBEF Experiment, we did not fully cross soil history and plant history treatments due to logistic constraints (there is no -SH +PH treatment). This means that we can study the role of soil history effects in BEF relationships in a long-term and large-scale field experiment, while the +SH +PH treatment can serve as an undisturbed control. However, results of this long-term control and previous smaller-scale experiments (van Moorsel et al. 2018, Zuppinger-Dingley et al. 2014) indicate that plant history is an important determinant of BEF relationships, which should be studied in future work.

Conclusions

The ΔBEF Experiment allows us to study mechanisms of BEF relationships in grassland plant communities by (1) isolating community age effects from climate effects, and (2) exploring the role of soil history effects. Notably, this experimental setup mimics the starting conditions of the Jena Experiment and thus enables us to compare young with old grassland communities of the same species composition but with different plant and soil histories. Our initial results suggest that BEF relationships are robust to climatic conditions and depend on soil and plant history. While the role of soil history may need more time to emerge, our findings indicate that future studies should explore the single and interactive effects of soil history and plant history. Moreover, plant history effects need to be studied by investigating compositional shifts (Allan et al. 2013, Marquard et al. 2009, Roscher et al. 2013) and changes in plant trait expression related to resource use (Tilman and Snell-Rood 2014, Zuppinger-Dingley et al. 2014) and defense against plant antagonists (Latz et al. 2012, Weisser et al. 2017, Zuppinger-Dingley et al. 2016). The community composition and effects of soil biota on plant community composition and growth need to be explored to better understand soil history effects. The ΔBEF Experiment provides the ideal infrastructure to perform such integrative research by considering the multifunctionality of these experimental grasslands and the linkages to plant-consumer interactions above and below the ground as well as consequences for element cycling.

Acknowledgements

We gratefully acknowledge the help of the coordinators of the Jena Experiment as well as all gardeners, technicians and numerous student helpers, who maintained the long-term experimental plots and helped to establish the ΔBEF Experiment. Furthermore, we thank Uta Gerighausen, Anja Zeuner and Odette Gonzalez for soil sampling and Ronja Kober-Moritz, Rebekka Mandler, Janine

738 Schmid and Erika Snjaric for soil extraction. The Jena Experiment is funded by the German Research
739 Foundation (DFG FOR 456, FOR 1451). Moreover, A.V., A.L. and N.E. acknowledge financial support
740 from the German Centre for Integrative Biodiversity Research Halle-Jena-Leipzig, funded by the
741 German Research Foundation (FZT 118) and M.L. from the Zwillenberg-Tietz foundation. Comments
742 by two anonymous reviewers helped to improve the manuscript.

743

744 References

745 Allan, E., Jenkins, T., Fergus, A. J. F., Roscher, C., Fischer, M., Petermann, J., Weisser, W. W. and
746 Schmid, B. (2013) Experimental plant communities develop phylogenetically overdispersed
747 abundance distributions during assembly. *Ecology, 94*(2), pp. 465-477.

748

749 Allan, E., Weisser, W., Weigelt, A., Roscher, C., Fischer, M. and Hillebrand, H. (2011) More diverse
750 plant communities have higher functioning over time due to turnover in complementary
751 dominant species. *Proceedings of the National Academy of Sciences of the United States of*
752 *America, 108*(41), pp. 17034-17039.

753

754 Anderson, M. J., Crist, T. O., Chase, J. M., Vellend, M., Inouye, B. D., Freestone, A. L., Sanders, N. J.,
755 Cornell, H. V., Comita, L. S., Davies, K. F., Harrison, S. P., Kraft, N. J. B., Stegen, J. C. and
756 Swenson, N. G. (2011) Navigating the multiple meanings of β diversity: a roadmap for the
757 practicing ecologist. *Ecology Letters, 14*(1), pp. 19-28.

758

759 Bai, Y. H., Xingguo; Chen, Zuozhong; Li, Linghao (2004) Eosystem stability and compensatory effects
760 in the inner mongolia grassland. *Nature, 431*, pp. 181-184.

761

762 Barry, K. E., Mommer, L., van Ruijven, J., Wirth, C., Wright, A. J., Bai, Y., Connolly, J., De Deyn, G. B.,
763 de Kroon, H., Isbell, F., Milcu, A., Roscher, C., Scherer-Lorenzen, M., Schmid, B. and Weigelt,
764 A. (2019) The Future of Complementarity: Disentangling Causes from Consequences. *Trends*
765 *Ecol Evol, 34*(2), pp. 167-180.

766

767 Beck, T., Joergensen, R. G., Kandeler, E., Makeschin, F., Nuss, E., Oberholzer, H. R. and Scheu, S.
768 (1997) An inter-laboratory comparison of ten different ways of measuring soil microbial
769 biomass C. *Soil Biology & Biochemistry, 29*(7), pp. 1023-1032.

770

771 Beniston, M. (2004) The 2003 heat wave in Europe: A shape of things to come? An analysis based on
772 Swiss climatological data and model simulations. *Geophysical Research Letters, 31*(2), pp. 4.

773

774 Bessler, H., Temperton, V. M., Roscher, C., Buchmann, N., Schmid, B., Schulze, E. D., Weisser, W. W.
775 and Engels, C. (2009) Aboveground overyielding in grassland mixtures is associated with
776 reduced biomass partitioning to belowground organs. *Ecology, 90*(6), pp. 1520-1530.
777
778 Bongers, T. and Bongers, M. (1998) Functional diversity of nematodes. *Applied Soil Ecology, 10*(3),
779 pp. 239-251.
780
781 Cardinale, B. J., Duffy, J. E., Gonzalez, A., Hooper, D. U., Perrings, C., Venail, P., Narwani, A., Mace, G.
782 M., Tilman, D., Wardle, D. A., Kinzig, A. P., Daily, G. C., Loreau, M., Grace, J. B., Larigauderie,
783 A., Srivastava, D. S. and Naeem, S. (2012) Biodiversity loss and its impact on humanity.
784 *Nature, 486*(7401), pp. 59-67.
785
786 Cardinale, B. J., Matulich, K. L., Hooper, D. U., Byrnes, J. E., Duffy, E., Gamfeldt, L., Balvanera, P.,
787 O'Connor, M. I. and Gonzalez, A. (2011) The Functional Role of Producer Diversity in
788 Ecosystems. *American Journal of Botany, 98*(3), pp. 572-592.
789
790 Cardinale, B. J., Wrigh, J. P., Cadotte, M. W., Carroll, I. T., Hector, A., Srivastava, D. S., Loreau, M. and
791 Weis, J. J. (2007) Impacts of plant diversity on biomass production increase through time
792 because of species complementarity. *Proceedings of the National Academy of Sciences of the*
793 *United States of America, 104*(46), pp. 18123-18128.
794
795 Convey, P. and Wynn-Williams, D. D. (2002) Antarctic soil nematode response to artificial climate
796 amelioration. *European Journal of Soil Biology, 38*(3), pp. 255-259.
797
798 Cowles, J. M., Wragg, P. D., Wright, A. J., Powers, J. S. and Tilman, D. (2016) Shifting grassland plant
799 community structure drives positive interactive effects of warming and diversity on
800 aboveground net primary productivity. *Global Change Biology, 22*(2), pp. 741-749.
801
802 De Boeck, H. J., Lemmens, C., Zavalloni, C., Gielen, B., Malchair, S., Carnol, M., Merckx, R., Van den
803 Berge, J., Ceulemans, R. and Nijs, I. (2008) Biomass production in experimental grasslands of
804 different species richness during three years of climate warming. *Biogeosciences, 5*(2), pp.
805 585-594.
806
807 Diaz, S. and Cabido, M. (2001) Vive la difference: plant functional diversity matters to ecosystem
808 processes. *Trends in Ecology & Evolution, 16*(11), pp. 646-655.
809
810 Duffy, J. E., Godwin, C. M. and Cardinale, B. J. (2017) Biodiversity effects in the wild are common and
811 as strong as key drivers of productivity. *Nature, 549*(7671), pp. 261-264.
812
813 Eisenhauer, N. (2012) Aboveground-belowground interactions as a source of complementarity
814 effects in biodiversity experiments. *Plant and Soil, 351*(1-2), pp. 1-22.
815
816 Eisenhauer, N., Bessler, H., Engels, C., Gleixner, G., Habekost, M., Milcu, A., Partsch, S., Sabais, A. C.
817 W., Scherber, C., Steinbeiss, S., Weigelt, A., Weisser, W. W. and Scheu, S. (2010) Plant

diversity effects on soil microorganisms support the singular hypothesis. *Ecology, 91*(2), pp. 485-496.

Eisenhauer, N., Dobies, T., Cesarz, S., Hobbie, S. E., Meyer, R. J., Worm, K. and Reich, P. B. (2013) Plant diversity effects on soil food webs are stronger than those of elevated CO₂ and N deposition in a long-term grassland experiment. *Proceedings of the National Academy of Sciences, 110*(17), pp. 6889-6894.

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

Eisenhauer, N., Migunova, V. D., Ackermann, M., Ruess, L. and Scheu, S. (2011a) Changes in Plant Species Richness Induce Functional Shifts in Soil Nematode Communities in Experimental Grassland. *Plos One, 6*(9), pp. 9.

Eisenhauer, N., Milcu, A., Sabais, A. C. W., Bessler, H., Brenner, J., Engels, C., Klarner, B., Maraun, M., Partsch, S., Roscher, C., Schonert, F., Temperton, V. M., Thomisch, K., Weigelt, A., Weisser, W. W. and Scheu, S. (2011b) Plant Diversity Surpasses Plant Functional Groups and Plant Productivity as Driver of Soil Biota in the Long Term. *Plos One, 6*(1), pp. 11.

Eisenhauer, N., Reich, P. B. and Scheu, S. (2012) Increasing plant diversity effects on productivity with time due to delayed soil biota effects on plants. *Basic and Applied Ecology, 13*(7), pp. 571-578.

Fargione, J., Tilman, D., Dybzinski, R., HilleRisLambers, J., Clark, C., Harpole, W. S., Knops, J. M. H., Reich, P. B. and Loreau, M. (2007) From selection to complementarity: shifts in the causes of biodiversity-productivity relationships in a long-term biodiversity experiment. *Proceedings of the Royal Society B-Biological Sciences, 274*(1611), pp. 871-876.

Fischer, C., Leimer, S., Roscher, C., Ravenek, J., de Kroon, H., Kreutziger, Y., Baade, J., Beßler, H., Eisenhauer, N., Weigelt, A., Mommer, L., Lange, M., Gleixner, G., Wilcke, W., Schröder, B. and Hildebrandt, A. (2019) Plant species richness and functional groups have different effects on soil water content in a decade-long grassland experiment. *Journal of Ecology, 107*(1), pp. 127-141.

Fornara, D. A. and Tilman, D. (2008) Plant functional composition influences rates of soil carbon and nitrogen accumulation. *Journal of Ecology, 96*(2), pp. 314-322.

Fridley, J. D. (2002) Resource availability dominates and alters the relationship between species diversity and ecosystem productivity in experimental plant communities. *Oecologia, 132*(2), pp. 271-277.

Galloway, J. N., Dentener, F. J., Capone, D. G., Boyer, E. W., Howarth, R. W., Seitzinger, S. P., Asner, G. P., Cleveland, C. C., Green, P. A., Holland, E. A., Karl, D. M., Michaels, A. F., Porter, J. H.,

Townsend, A. R. and Vöosmarty, C. J. (2004) Nitrogen Cycles: Past, Present, and Future. *Biogeochemistry, 70*(2), pp. 153-226.

Gerzabek, M. H., Strebl, F., Tulipan, M. and Schwarz, S. (2005) Quantification of organic carbon pools for Austria's agricultural soils using a soil information system. *Canadian Journal of Soil Science, 85*(Special Issue), pp. 491-498.

Giling, D. P., Beaumelle, L., Phillips, H. R. P., Cesarz, S., Eisenhauer, N., Ferlian, O., Gottschall, F., Guerra, C., Hines, J., Sendek, A., Siebert, J., Thakur, M. P. and Barnes, A. D. (2019) A niche for ecosystem multifunctionality in global change research. *Global Change Biology, 25*(3), pp. 763-774.

Gubsch, M., Buchmann, N., Schmid, B., Schulze, E. D., Lipowsky, A. and Roscher, C. (2011) Differential effects of plant diversity on functional trait variation of grass species. *Annals of Botany, 107*(1), pp. 157-169.

Guerrero-Ramírez, N. R., Craven, D., Reich, P. B., Ewel, J. J., Isbell, F., Koricheva, J., Parrotta, J. A., Auge, H., Erickson, H. E., Forrester, D. I., Hector, A., Joshi, J., Montagnini, F., Palmborg, C., Piotto, D., Potvin, C., Roscher, C., van Ruijven, J., Tilman, D., Wilsey, B. and Eisenhauer, N. (2017) Diversity-dependent temporal divergence of ecosystem functioning in experimental ecosystems. *Nature Ecology & Evolution, 1*(11), pp. 1639-1642.

Habekost, M., Eisenhauer, N., Scheu, S., Steinbeiss, S., Weigelt, A. and Gleixner, G. (2008) Seasonal changes in the soil microbial community in a grassland plant diversity gradient four years after establishment. *Soil Biology and Biochemistry, 40*(10), pp. 2588-2595.

He, J. S., Bazzaz, F. A. and Schmid, B. (2002) Interactive effects of diversity, nutrients and elevated CO2 on experimental plant communities. *Oikos, 97*(3), pp. 337-348.

Hector, A., Schmid, B., Beierkuhnlein, C., Caldeira, M. C., Diemer, M., Dimitrakopoulos, P. G., Finn, J. A., Freitas, H., Giller, P. S., Good, J., Harris, R., Hogberg, P., Huss-Danell, K., Joshi, J., Jumpponen, A., Korner, C., Leadley, P. W., Loreau, M., Minns, A., Mulder, C. P. H., O'Donovan, G., Otway, S. J., Pereira, J. S., Prinz, A., Read, D. J., Scherer-Lorenzen, M., Schulze, E. D., Siamantziouras, A. S. D., Spehn, E. M., Terry, A. C., Troumbis, A. Y., Woodward, F. I., Yachi, S. and Lawton, J. H. (1999) Plant diversity and productivity experiments in European grasslands. *Science, 286*(5442), pp. 1123-1127.

Heisse, K., Roscher, C., Schumacher, J. and Schulze, E. D. (2007) Establishment of grassland species in monocultures: different strategies lead to success. *Oecologia, 152*(3), pp. 435-447.

HilleRisLambers, J., Harpole, W. S., Tilman, D., Knops, J. and Reich, P. B. (2004) Mechanisms responsible for the positive diversity-productivity relationship in Minnesota grasslands. *Ecology Letters, 7*(8), pp. 661-668.

907 Hirsch, P. R., Gilliam, L. M., Sohi, S. P., Williams, J. K., Clark, I. M. and Murray, P. J. (2009) Starving the
908 soil of plant inputs for 50 years reduces abundance but not diversity of soil bacterial
909 communities. *Soil Biology and Biochemistry, 41*(9), pp. 2021-2024.

910
911 Hoffmann, K., Bivour, W., Früh, B., Koßmann, M. and Voß, P.-H. (2014) *Klimauntersuchungen in Jena*
912 *für die Anpassung an den Klimawandel und seine erwarteten Folgen. Ein Ergebnisbericht,*
913 *Berichte des Deutschen Wetterdienstes,* Offenbach am Main: Selbstverlag des Deutschen
914 Wetterdienstes.

915
916 Holland, J. M. (2004) The environmental consequences of adopting conservation tillage in Europe:
917 reviewing the evidence. *Agriculture, Ecosystems & Environment, 103*(1), pp. 1-25.

918
919 Hooper, D. U., Adair, E. C., Cardinale, B. J., Byrnes, J. E. K., Hungate, B. A., Matulich, K. L., Gonzalez,
920 A., Duffy, J. E., Gamfeldt, L. and O'Connor, M. I. (2012) A global synthesis reveals biodiversity
921 loss as a major driver of ecosystem change. *Nature, 486*(7401), pp. 105-U129.

922
923 Huang, Y., Chen, Y., Castro-Izaguirre, N., Baruffol, M., Brezzi, M., Lang, A., Li, Y., Härdtle, W., von
924 Oheimb, G., Yang, X., Liu, X., Pei, K., Both, S., Yang, B., Eichenberg, D., Assmann, T., Bauhus,
925 J., Behrens, T., Buscot, F., Chen, X.-Y., Chesters, D., Ding, B.-Y., Durka, W., Erfmeier, A., Fang,
926 J., Fischer, M., Guo, L.-D., Guo, D., Gutknecht, J. L. M., He, J.-S., He, C.-L., Hector, A., Hönig,
927 L., Hu, R.-Y., Klein, A.-M., Kühn, P., Liang, Y., Li, S., Michalski, S., Scherer-Lorenzen, M.,
928 Schmidt, K., Scholten, T., Schuldt, A., Shi, X., Tan, M.-Z., Tang, Z., Trogisch, S., Wang, Z., Welk,
929 E., Wirth, C., Wubet, T., Xiang, W., Yu, M., Yu, X.-D., Zhang, J., Zhang, S., Zhang, N., Zhou, H.-
930 Z., Zhu, C.-D., Zhu, L., Bruelheide, H., Ma, K., Niklaus, P. A. and Schmid, B. (2018) Impacts of
931 species richness on productivity in a large-scale subtropical forest experiment. *Science,*
932 *362*(6410), pp. 80-83.

933
934 IPCC (2014) *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the*
935 *Fifth Assessment Report of the Intergovernmental Panel on Climate Change,* Geneva,
936 Switzerland: IPCC.

937
938 Isbell, F., Calcagno, V., Hector, A., Connolly, J., Harpole, W. S., Reich, P. B., Scherer-Lorenzen, M.,
939 Schmid, B., Tilman, D., van Ruijven, J., Weigelt, A., Wilsey, B. J., Zavaleta, E. S. and Loreau, M.
940 (2011) High plant diversity is needed to maintain ecosystem services. *Nature, 477*(7363), pp.
941 199-U96.

942
943 Isbell, F., Craven, D., Connolly, J., Loreau, M., Schmid, B., Beierkuhnlein, C., Bezemer, T. M., Bonin, C.,
944 Bruelheide, H., de Luca, E., Ebeling, A., Griffin, J. N., Guo, Q. F., Hautier, Y., Hector, A.,
945 Jentsch, A., Kreyling, J., Lanta, V., Manning, P., Meyer, S. T., Mori, A. S., Naeem, S., Niklaus, P.
946 A., Polley, H. W., Reich, P. B., Roscher, C., Seabloom, E. W., Smith, M. D., Thakur, M. P.,
947 Tilman, D., Tracy, B. F., van der Putten, W. H., van Ruijven, J., Weigelt, A., Weisser, W. W.,
948 Wilsey, B. and Eisenhauer, N. (2015) Biodiversity increases the resistance of ecosystem
949 productivity to climate extremes. *Nature, 526*(7574), pp. 574-U263.

950
951 Isbell, F., Reich, P. B., Tilman, D., Hobbie, S. E., Polasky, S. and Binder, S. (2013) Nutrient enrichment,
952 biodiversity loss, and consequent declines in ecosystem productivity. *Proceedings of the*
953 *National Academy of Sciences of the United States of America, 110*(29), pp. 11911-11916.

IUSS Working Group WRB (2015) World Reference Base for Soil Resources 2014, International soil classification system for naming soils and creating legends for soil maps, update 2015. in,Rome: FAO.

Kulmatiski, A., Beard Karen, H. and Heavilin, J. (2012) Plant–soil feedbacks provide an additional explanation for diversity–productivity relationships. *Proceedings of the Royal Society B: Biological Sciences, 279*(1740), pp. 3020-3026.

Lange, M., Eisenhauer, N., Sierra, C. A., Bessler, H., Engels, C., Griffiths, R. I., Mellado-Vázquez, P. G., Malik, A. A., Roy, J., Scheu, S., Steinbeiss, S., Thomson, B. C., Trumbore, S. E. and Gleixner, G. (2015) Plant diversity increases soil microbial activity and soil carbon storage. *Nature Communications, 6*, pp. 6707.

Lange, M., Habekost, M., Eisenhauer, N., Roscher, C., Bessler, H., Engels, C., Oelmann, Y., Scheu, S., Wilcke, W., Schulze, E.-D. and Gleixner, G. (2014) Biotic and Abiotic Properties Mediating Plant Diversity Effects on Soil Microbial Communities in an Experimental Grassland. *Plos One, 9*(5), pp. e96182.

Lange, M., Koller-France, E., Hildebrandt, A., Oelmann, Y., Wilcke, W. and Gleixner, G. (2019) How plant diversity impacts the linkage of water, nutrient and carbon cycles. *Advances in Ecological Research,* (61).

Latz, E., Eisenhauer, N., Rall, B. C., Allan, E., Roscher, C., Scheu, S. and Jousset, A. (2012) Plant diversity improves protection against soil-borne pathogens by fostering antagonistic bacterial communities. *Journal of Ecology, 100*(3), pp. 597-604.

Latz, E., Eisenhauer, N., Rall, B. C., Scheu, S. and Jousset, A. (2016) Unravelling Linkages between Plant Community Composition and the Pathogen-Suppressive Potential of Soils. *Scientific Reports, 6*, pp. 23584.

Leff, J. W., Bardgett, R. D., Wilkinson, A., Jackson, B. G., Pritchard, W. J., De Long, J. R., Oakley, S., Mason, K. E., Ostle, N. J., Johnson, D., Baggs, E. M. and Fierer, N. (2018) Predicting the structure of soil communities from plant community taxonomy, phylogeny, and traits. *The ISME Journal, 12*(7), pp. 1794-1805.

Leimer, S., Wirth, C., Oelmann, Y. and Wilcke, W. (2014) Biodiversity effects on nitrate concentrations in soil solution: a Bayesian model. *Biogeochemistry, 118*(1), pp. 141-157.

Lipowsky, A., Roscher, C., Schumacher, J., Michalski, S. G., Gubsch, M., Buchmann, N., Schulze, E.-D. and Schmid, B. (2015) Plasticity of functional traits of forb species in response to biodiversity. *Perspectives in Plant Ecology, Evolution and systematics, 17*(1), pp. 66-77.

Lipowsky, A., Schmid, B. and Roscher, C. (2011) Selection for monoculture and mixture genotypes in a biodiversity experiment. *Basic and Applied Ecology, 12*(4), pp. 360-371.

Londo, G. (1976) The decimal scale for releves of permanent quadrats. *Vegetatio, 33*(1), pp. 61-64.

Loreau, M. and Hector, A. (2001) Partitioning selection and complementarity in biodiversity experiments. *Nature, 412*(6842), pp. 72-76.

Manning, P., van der Plas, F., Soliveres, S., Allan, E., Maestre, F. T., Mace, G., Whittingham, M. J. and Fischer, M. (2018) Redefining ecosystem multifunctionality. *Nature Ecology & Evolution, 2*(3), pp. 427-436.

Maron, J. L., Marler, M., Klironomos, J. N. and Cleveland, C. C. (2011) Soil fungal pathogens and the relationship between plant diversity and productivity. *Ecology Letters, 14*(1), pp. 36-41.

Marquard, E., Weigelt, A., Temperton, V. M., Roscher, C., Schumacher, J., Buchmann, N., Fischer, M., Weisser, W. W. and Schmid, B. (2009) Plant species richness and functional composition drive overyielding in a six-year grassland experiment. *Ecology, 90*(12), pp. 3290-3302.

Mellado-Vázquez, P. G., Lange, M., Bachmann, D., Gockele, A., Karlowsky, S., Milcu, A., Piel, C., Roscher, C., Roy, J. and Gleixner, G. (2016) Plant diversity generates enhanced soil microbial access to recently photosynthesized carbon in the rhizosphere. *Soil Biology and Biochemistry, 94*, pp. 122-132.

Meyer, S. T., Ebeling, A., Eisenhauer, N., Hertzog, L., Hillebrand, H., Milcu, A., Pompe, S., Abbas, M., Bessler, H., Buchmann, N., De Luca, E., Engels, C., Fischer, M., Gleixner, G., Hudewenz, A., Klein, A.-M., de Kroon, H., Leimer, S., Loranger, H., Mommer, L., Oelmann, Y., Ravenek, J. M., Roscher, C., Rottstock, T., Scherber, C., Scherer-Lorenzen, M., Scheu, S., Schmid, B., Schulze, E.-D., Staudler, A., Strecker, T., Temperton, V., Tscharntke, T., Vogel, A., Voigt, W., Weigelt, A., Wilcke, W. and Weisser, W. W. (2016) Effects of biodiversity strengthen over time as ecosystem functioning declines at low and increases at high biodiversity. *Ecosphere, 7*(12), pp. e01619-n/a.

Mikola, J., Sørensen, L. I. and Kytöviita, M.-M. (2014) Plant removal disturbance and replant mitigation effects on the abundance and diversity of low-arctic soil biota. *Applied Soil Ecology, 82*, pp. 82-92.

Mueller, K. E., Tilman, D., Fornara, D. A. and Hobbie, S. E. (2013) Root depth distribution and the diversity–productivity relationship in a long-term grassland experiment. *Ecology, 94*(4), pp. 787-793.

Naeem, S., Duffy, J. E. and Zavaleta, E. (2012) The Functions of Biological Diversity in an Age of Extinction. *Science, 336*(6087), pp. 1401-1406.

Nico Eisenhauer, Holger Schielzeth, Andrew Barnes, Kathryn Barry, Aletta Bonn, Ulrich Brose, Helge Bruelheide, Nina Buchmann, François Buscot, Anne Ebeling, Olga Ferlian, Grégoire T. Freschet, Darren P. Giling, Stephan Hättenschwiler, Helmut Hillebrand, Jes Hines, Forest Isbell, Eva Koller-France, Birgitta König-Ries, Hans de Kroon, Sebastian T. Meyer, Alexandru

Milcu, Charles A. Nock, Jana S. Petermann, Christiane Roscher, Christoph Scherber, Michael Scherer-Lorenzen, Bernhard Schmid, Stefan A. Schnitzer, Teja Tscharntke, Manfred Türke, Nicole M. van Dam, Anja Vogel, Cameron Wagg, David A. Wardle, Alexandra Weigelt, Wolfgang W. Weisser, Christian Wirth and Jochum, M. (2019) A multitrophic, eco-evolutionary perspective on biodiversity–ecosystem function research. *Advances in Ecological Research,* (61).

Oelmann, Y., Buchmann, N., Gleixner, G., Habekost, M., Roscher, C., Rosenkranz, S., Schulze, E. D., Steinbeiss, S., Temperton, V. M., Weigelt, A., Weisser, W. W. and Wilcke, W. (2011) Plant diversity effects on aboveground and belowground N pools in temperate grassland ecosystems: Development in the first 5 years after establishment. *Global Biogeochemical Cycles, 25*, pp. 11.

Oelmann, Y., Wilcke, W., Temperton, V. M., Buchmann, N., Roscher, C., Schumacher, J., Schulze, E. D. and Weisser, W. W. (2007) Soil and plant nitrogen pools as related to plant diversity in an experimental grassland. *Soil Science Society of America Journal, 71*(3), pp. 720-729.

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlin, D., Minchin, P., O'Hara, R. B., Simpson, G., Solymos, P. and others (2017) vegan: Community Ecology Package. in.

Ravenek, J. M., Bessler, H., Engels, C., Scherer-Lorenzen, M., Gessler, A., Gockele, A., De Luca, E., Temperton, V. M., Ebeling, A., Roscher, C., Schmid, B., Weisser, W. W., Wirth, C., de Kroon, H., Weigelt, A. and Mommer, L. (2014) Long-term study of root biomass in a biodiversity experiment reveals shifts in diversity effects over time. *Oikos, 123*(12), pp. 1528-1536.

Reich, P. B., Knops, J., Tilman, D., Craine, J., Ellsworth, D., Tjoelker, M., Lee, T., Wedin, D., Naeem, S., Bahauddin, D., Hendrey, G., Jose, S., Wrage, K., Goth, J. and Bengston, W. (2001) Plant diversity enhances ecosystem responses to elevated CO2 and nitrogen deposition. *Nature, 410*(6839), pp. 809-812.

Reich, P. B., Tilman, D., Isbell, F., Mueller, K., Hobbie, S. E., Flynn, D. F. B. and Eisenhauer, N. (2012) Impacts of Biodiversity Loss Escalate Through Time as Redundancy Fades. *Science, 336*(6081), pp. 589-592.

Roscher, C., Schmid, B., Buchmann, N., Weigelt, A. and Schulze, E. D. (2011a) Legume species differ in the responses of their functional traits to plant diversity. *Oecologia, 165*(2), pp. 437-452.

Roscher, C., Schumacher, J., Baade, J., Wilcke, W., Gleixner, G., Weisser, W. W., Schmid, B. and 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*(2), pp. 107-121.

Roscher, C., Schumacher, J., Gubsch, M., Lipowsky, A., Weigelt, A., Buchmann, N., Schmid, B. and Schulze, E.-D. (2012) Using plant functional traits to explain diversity-productivity relationships. *Plos One, 7*(5), pp. e36760.

Roscher, C., Schumacher, J., Lipowsky, A., Gubsch, M., Weigelt, A., Pompe, S., Kolle, O., Buchmann, N., Schmid, B. and Schulze, E. D. (2013) A functional trait-based approach to understand community assembly and diversity-productivity relationships over 7 years in experimental grasslands. *Perspectives in Plant Ecology Evolution and Systematics, 15*(3), pp. 139-149.

Roscher, C., Thein, S., Schmid, B. and Scherer-Lorenzen, M. (2008) Complementary nitrogen use among potentially dominant species in a biodiversity experiment varies between two years. *Journal of Ecology, 96*(3), pp. 477-488.

Roscher, C., Weigelt, A., Proulx, R., Marquard, E., Schumacher, J., Weisser, W. W. and Schmid, B. (2011b) Identifying population- and community-level mechanisms of diversity-stability relationships in experimental grasslands. *Journal of Ecology, 99*(6), pp. 1460-1469.

Rosenkranz, S., Wilcke, W., Eisenhauer, N. and Oelmann, Y. (2012) Net ammonification as influenced by plant diversity in experimental grasslands. *Soil Biology & Biochemistry, 48*, pp. 78-87.

Ruess, L. (1995) Studies On the Nematode Fauna of an Acid Forest Soil: Spatial Distribution and Extraction. *Nematologica, 41*, pp. 229-239.

Scheu, S. (1992) Automated Measurement of the Respiratory Response of Soil Microcompartments - Active Microbial Biomass in Earthworm Faeces. *Soil Biology & Biochemistry, 24*(11), pp. 1113-1118.

Schmid, M. W., Hahl, T., van Moorsel, S. J., Wagg, C., De Deyn, G. B. and Schmid, B. (2019) Feedbacks of plant identity and diversity on the diversity and community composition of rhizosphere microbiomes from a long-term biodiversity experiment. *Molecular Ecology, 28*(4), pp. 863-878.

Schnitzer, S. A., Klironomos, J. N., HilleRisLambers, J., Kinkel, L. L., Reich, P. B., Xiao, K., Rillig, M. C., Sikes, B. A., Callaway, R. M., Mangan, S. A., van Nes, E. H. and Scheffer, M. (2011) Soil microbes drive the classic plant diversity–productivity pattern. *Ecology, 92*(2), pp. 296-303.

Six, J., Elliott, E. T. and Paustian, K. (2000) Soil macroaggregate turnover and microaggregate formation: a mechanism for C sequestration under no-tillage agriculture. *Soil Biology and Biochemistry, 32*(14), pp. 2099-2103.

Steinbeiss, S., Bessler, H., Engels, C., Temperton, V. M., Buchmann, N., Roscher, C., Kreutziger, Y., Baade, J., Habekost, M. and Gleixner, G. (2008) Plant diversity positively affects short-term soil carbon storage in experimental grasslands. *Global Change Biology, 14*(12), pp. 2937-2949.

Stevenson, B. A., Hunter, D. W. F. and Rhodes, P. L. (2014) Temporal and seasonal change in microbial community structure of an undisturbed, disturbed, and carbon-amended pasture soil. *Soil Biology and Biochemistry, 75*, pp. 175-185.

Strecker, T., Macé, O. G., Scheu, S. and Eisenhauer, N. (2016) Functional composition of plant communities determines the spatial and temporal stability of soil microbial properties in a long-term plant diversity experiment. *Oikos, 125*(12), pp. 1743-1754.

ter Heerdt, G. N. J., Verweij, G. L., Bekker, R. M. and Bakker, J. P. (1996) An Improved Method for Seed-Bank Analysis: Seedling Emergence After Removing the Soil by Sieving. *Functional Ecology, 10*(1), pp. 144-151.

Thakur, M. P., Milcu, A., Manning, P., Niklaus, P. A., Roscher, C., Power, S., Reich, P. B., Scheu, S., Tilman, D., Ai, F. X., Guo, H. Y., Ji, R., Pierce, S., Ramirez, N. G., Richter, A. N., Steinauer, K., Strecker, T., Vogel, A. and Eisenhauer, N. (2015) Plant diversity drives soil microbial biomass carbon in grasslands irrespective of global environmental change factors. *Global Change Biology, 21*(11), pp. 4076-4085.

Tilman, D., Lehman, C. L. and Thomson, K. T. (1997) Plant diversity and ecosystem productivity: Theoretical considerations. *Proceedings of the National Academy of Sciences of the United States of America, 94*(5), pp. 1857-1861.

Tilman, D. and Snell-Rood, E. (2014) Diversity breeds complementarity. *Nature, 515*(7525), pp. 44-45.

van Moorsel, S. J., Hahl, T., Wagg, C., De Deyn, G. B., Flynn, D. F. B., Zuppinger-Dingley, D. and Schmid, B. (2018) Community evolution increases plant productivity at low diversity. *Ecology Letters, 21*(1), pp. 128-137.

Wagg, C., Barendregt, C., Jansa, J. and van der Heijden, M. G. A. (2015) Complementarity in both plant and mycorrhizal fungal communities are not necessarily increased by diversity in the other. *Journal of Ecology, 103*(5), pp. 1233-1244.

Wardle, D. A., Bonner, K. I., Barker, G. M., Yeates, G. W., Nicholson, K. S., Bardgett, R. D., Watson, R. N. and Ghani, A. (1999) Plant removals in perennial grassland: Vegetation dynamics, decomposers, soil biodiversity, and ecosystem properties. *Ecological Monographs, 69*(4), pp. 535-568.

Weisser, W. W., Roscher, C., Meyer, S. T., Ebeling, A., Luo, G., Allan, E., Beßler, H., Barnard, R. L., Buchmann, N., Buscot, F., Engels, C., Fischer, C., Fischer, M., Gessler, A., Gleixner, G., Halle, S., Hildebrandt, A., Hillebrand, H., de Kroon, H., Lange, M., Leimer, S., Le Roux, X., Milcu, A., Mommer, L., Niklaus, P. A., Oelmann, Y., Proulx, R., Roy, J., Scherber, C., Scherer-Lorenzen, M., Scheu, S., Tscharntke, T., Wachendorf, M., Wagg, C., Weigelt, A., Wilcke, W., Wirth, C., Schulze, E.-D., Schmid, B. and Eisenhauer, N. (2017) Biodiversity effects on ecosystem functioning in a 15-year grassland experiment: Patterns, mechanisms, and open questions. *Basic and Applied Ecology, 23*(Supplement C), pp. 1-73.

Wright, A. J., Ebeling, A., de Kroon, H., Roscher, C., Weigelt, A., Buchmann, N., Buchmann, T., Fischer, C., Hacker, N., Hildebrandt, A., Leimer, S., Mommer, L., Oelmann, Y., Scheu, S., Steinauer, K., Strecker, T., Weisser, W., Wilcke, W. and Eisenhauer, N. (2015) Flooding disturbances

increase resource availability and productivity but reduce stability in diverse plant communities. *Nature Communications, 6*, pp. 6.

Wright, A. J., Wardle, D. A., Callaway, R. and Gaxiola, A. (2017) The Overlooked Role of Facilitation in Biodiversity Experiments. *Trends in Ecology & Evolution, 32*(5), pp. 383-390.

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

Yeates, G. W., Bongers, T., De Goede, R. G. M., Freckman, D. W. and Georgieva, S. S. (1993) Feeding Habits in Soil Nematode Families and Genera—An Outline for Soil Ecologists. *Journal of Nematology, 25*(3), pp. 315-331.

Zaitchik, B. F., Macalady, A. K., Bonneau, L. R. and Smith, R. B. (2006) Europe's 2003 heat wave: A satellite view of impacts and land-atmosphere feedbacks. *International Journal of Climatology, 26*(6), pp. 743-769.

Zuppinger-Dingley, D., Flynn, D. F. B., De Deyn, G. B., Petermann, J. S. and Schmid, B. (2016) Plant selection and soil legacy enhance long-term biodiversity effects. *Ecology, 97*(4), pp. 918-928.

Zuppinger-Dingley, D., Schmid, B., Petermann, J. S., Yadav, V., De Deyn, G. B. and Flynn, D. F. B. (2014) Selection for niche differentiation in plant communities increases biodiversity effects. *Nature, 515*(7525), pp. 108-+.

Tables

Table 1: Name and description of the experimental treatments of the ΔBEF experiment. Here, no soil or plant history (-SH, -PH) refers to a legacy of non-experimental conditions, while soil or plant history (+SH, +PH) refers to a legacy of 15 years of experimental grassland diversity treatments. The levels of the soil and plant histories were combined to the treatments of the ΔBEF experiment.

Name of treatment	Soil history	Age of soil history (yr)	Soil history contrast for mixed models	Plant history	Age of plant history (yr)	Plant history contrast for mixed models	Year contrast for mixed models
-SH -PH	conventional (crop production)	1	-SH	unspecific (commercial supplier)	1	-PH	2017
+SH -PH	plotspecific (experimental grassland)	15	+SH	unspecific (commercial supplier)	1	-PH	2017
+SH +PH	plotspecific (experimental grassland)	15	+SH	plotspecific (experimental grassland)	15	+PH	2017
2003	conventional (crop production)	1	-SH	unspecific (commercial supplier)	1	-PH	2003

Table 2: Effects of experimental treatments on plant community properties. We obtained results from different linear mixed-effects model analyses: treatment was fitted as a factor with three to four levels in model A and depicted into a year (2003 vs. 2017, according to Table 1), a soil history (-SH vs. +SH, according to Table 1) and a plant history contrast (-PH vs. +PH, according to Table 1) in models B, C and D with varying order of the contrasts. In this way, model B tests for the pure plant history effect by contrasting treatment +SH -PH against treatment +SH +PH, model C tests for the pure soil history effect by contrasting treatment -SH -PH against treatment +SH –PH, and model D tests for the effect of the calendar year by contrasting treatment -SH -PH against 2003. ***: $p \leq 0.001$, **: $p \leq 0.01$, *: $p \leq 0.05$, .: $p \leq 0.1$. Plant species richness was modelled as a log-linear term.

		Seedbank (# viable seeds)					Number of seedlings					Density				LAI spring			
	Model	numDf	denDf	F	*p*		numDf	denDf	F	*p*		denDf	F	*p*		denDf	F	*p*	
(Intercept)	A, B, C, D	1	74	107.39	<0.000	***	1	148	1553.03	0.000	***	138	2956.42	0.000	***	144	206.03	0.000	***
Plant species richness (SR), log-linear	A, B, C, D	1	71	22.79	<0.000	***	1	71	9.43	0.003	**	71	22.50	0.000	***	71	28.09	0.000	***
Treatments (T)	**A**						**2**	**148**	**39.29**	**0.000**	***	138	22.02	0.000	***	**144**	**35.80**	**0.000**	***
Year (Y)	B																		
Soil history (SH)	B						1	148	16.93	0.000	***	138	14.60	0.000	***	144	70.47	0.000	***
Plant history (PH)	**B**	**1**	**74**	**53.75**	**<0.000**	***	**1**	**148**	**61.64**	**0.000**	***	**138**	**29.43**	**0.000**	***	**144**	**1.14**	**0.288**	
Y	C																		
PH	C						1	148	78.44	0.000	***	138	43.60	0.000	***	144	10.59	0.001	**
SH	**C**						**1**	**148**	**0.13**	**0.717**		**138**	**0.44**	**0.510**		**144**	**61.01**	**0.000**	***
PH	D																		
SH	D																		
Y	**D**																		
T x SR	**A**						**2**	**148**	**4.09**	**0.019**	*	**138**	**8.46**	**0.000**	***	**144**	**6.78**	**0.002**	**
Y : SR	B																		

SH x SR	B						1	148	3.52	0.063	.	138	9.38	0.003	**	144	7.58	0.007	**
PH x SR	**B**	**1**	**74**	**0.14**	**0.712**		**1**	**148**	**4.66**	**0.033**	*****	**138**	**7.54**	**0.007**	******	**144**	**5.97**	**0.016**	*****
Y : SR	C																		
PH x SR	C						1	148	7.88	0.006	**	138	15.17	0.000	***	144	12.11	0.001	***
SH x SR	**C**						**1**	**148**	**0.30**	**0.586**		**138**	**1.75**	**0.188**		**144**	**1.45**	**0.231**	
PH x SR	D																		
SH x SR	D																		
Y : SR	**D**																		

Table 2 continued

		Realized plant species richness					Simson evenness spring			
	Model	numDf	denDf	F	*p*		denDf	F	*p*	
(Intercept)	A, B, C, D	1	220	8155.10	<0.001	***	234	625.88	<0.001	***
Plant species richness (SR), log-linear	A, B, C, D	1	71	4859.88	<0.001	***	75	319.18	<0.001	***
Treatments (T)	**A**	**3**	**220**	**17.68**	**<0.001**	*******	**234**	**1.51**	**0.212**	
Year (Y)	B	**1**	220	14.26	<0.001	***	234	2.29	0.131	
Soil history (SH)	B	1	220	9.15	0.003	**	234	0.26	0.609	
Plant history (PH)	**B**	**1**	**220**	**29.64**	**<0.001**	*******	**234**	**1.99**	**0.160**	
Y	C	1	220	14.26	<0.001	***	234	2.29	0.131	
PH	C	1	220	38.78	<0.001	***	234	0.93	0.335	
SH	**C**	**1**	**220**	**0.01**	**0.941**		**234**	**1.32**	**0.252**	
PH	D	1	220	50.75	<0.001	***	234	2.00	0.159	
SH	D	1	220	0.48	0.490		234	0.19	0.659	
Y	**D**	**1**	**220**	**1.81**	**0.180**		**234**	**2.35**	**0.127**	

T x SR	**A**	3	220	16.24	<0.001	***	**234**	**2.94**	**0.034**	*
Y x SR	B	**1**	**220**	**11.53**	**0.001**	***	234	0.33	0.566	
SH x SR	B	**1**	**220**	**9.22**	**0.003**	**	234	5.82	0.017	*
PH x SR	**B**	**1**	**220**	**27.99**	**<0.001**	***	**234**	**2.68**	**0.103**	
Y x SR	C	1	220	11.53	0.001	***	234	0.33	0.566	
PH x SR	C	1	220	37.20	<0.001	***	234	6.88	0.009	**
SH x SR	**C**	**1**	**220**	**0.00**	**0.962**		**234**	**1.61**	**0.205**	
PH x SR	D	1	220	47.23	<0.001	***	234	7.10	0.008	**
SH x SR	D	1	220	0.43	0.514		234	0.87	0.351	
Y x SR	**D**	**1**	**220**	**1.08**	**0.301**		**234**	**0.85**	**0.357**	

Df = degrees of freedom.

Table 3: Effects of experimental treatments on abiotic and biotic soil properties. We obtained results from different linear mixed effects model analyses: treatment was fitted as an overall factor with three levels in model A and depicted into a soil history (-SH vs. +SH, according to Table 1) and a plant history contrast (-PH vs. +PH, according to Table 1) in models B and C with varying order of the contrasts. In this way, model B tests for the pure plant history effect by contrasting treatment +SH -PH against treatment +SH +PH, and model C tests for the pure soil history effect by contrasting treatment -SH -PH against treatment +SH –PH. ***: p ≤ 0.001, **: p ≤ 0.01, *: p ≤ 0.05, .: p ≤ 0.1. Plant species richness was modelled as a log-linear term.

		Soil moisture					Soil mineral N				Soil organic carbon				Soil microbial biomass				So
	Model	numDf	denDf	F	*p*		denDf	F	*p*		denDf	F	*p*		denDf	F	*p*		denDf
(Intercept)	A, B, C	1	148	1567.58	<0.001	***	116	8.40	0.005	**	148	410.99	<0.001	***	148	5379.88	<0.001	***	79
Plant species richness (SR), log-linear	A, B, C	1	71	37.16	<0.001	***	56	0.16	0.693		71	38.982	<0.001	***	71	17.08	<0.001	***	39
Treatments (T)	**A**	**2**	**148**	**40.80**	**<0.001**	***	**116**	**22.07**	**<0.001**	***	**148**	**175.69**	**<0.001**	***	**148**	**107.79**	**<0.001**	***	**79**
Soil history (SH)	B	1	148	31.98	<0.001	***	116	12.54	0.001	***	148	21.833	<0.001	***	148	5.93	0.016	*	79
Plant history (PH)	**B**	**1**	**148**	**49.63**	**<0.001**	***	**116**	**31.60**	**<0.001**	***	**148**	**329.55**	**<0.001**	***	**148**	**209.65**	**<0.001**	***	**79**
Plant history (PH)	C	1	148	79.72	<0.001	***	116	44.08	<0.001	***	148	179.16	<0.001	***	148	189.24	<0.001	***	79
Soil history (SH)	**C**	**1**	**148**	**1.89**	**0.171**		**116**	**0.06**	**0.805**		**148**	**172.22**	**<0.001**	***	**148**	**26.33**	**<0.001**	***	**79**
T x SR	**A**	**2**	**148**	**15.50**	**<0.001**	***	**116**	**0.23**	**0.792**		**148**	**21.195**	**<0.001**	***	**148**	**2.91**	**0.058**	.	**79**
SH x SR	B	1	148	13.48	<0.001	***	116	0.10	0.750		148	19.805	<0.001	***	148	3.54	0.062	.	79
PH x SR	**B**	**1**	**148**	**17.52**	**<0.001**	***	**116**	**0.37**	**0.547**		**148**	**22.585**	**<0.001**	***	**148**	**2.28**	**0.133**		**79**
PH x SR	C	1	148	29.81	<0.001	***	116	0.47	0.496		148	40.206	<0.001	***	148	5.06	0.026	*	79
SH x SR	**C**	**1**	**148**	**1.18**	**0.279**		**116**	**0.00**	**0.980**		**148**	**2.1841**	**0.142**		**148**	**0.76**	**0.384**		**79**

Df = degrees of freedom.

Table 4: Effects of experimental treatments on annual aboveground biomass and standardized annual aboveground biomass. Standardization was done by division of aboveground biomass by the mean biomass per treatment. We obtained results from different linear mixed effects model analyses: treatment was fitted as a factor with four levels in model A and depicted into a year (2003 vs. 2017, according to Table 1), a soil history (-SH vs. +SH, according to Table 1) and a plant history contrast (-PH vs. +PH, according to Table 1) in models B, C and D with varying order of the contrasts. In this way, model B tests for the pure plant history effect by contrasting treatment +SH -PH against treatment +SH +PH, model C tests for the pure soil history effect by contrasting treatment -SH -PH against treatment +SH –PH, and model D tests for the effect of the calendar year by contrasting treatment -SH -PH against 2003. ***: $p \leq 0.001$, **: $p \leq 0.01$, *: $p \leq 0.05$, .: $p \leq 0.1$. Plant species richness was modelled as a log-linear term.

			Annual biomass				Standardized biomass		
	Model	numDf	denDf	F	*p*		F	*p*	
(Intercept)	A, B, C, D	1	222	7750.20	<0.001	***	4.89	0.028	*
Plant species richness (SR), log-linear	A, B, C, D	1	71	50.33	<0.001	***	44.54	<0.001	***
Treatments (T)	**A**	**3**	**222**	**49.84**	**<0.001**	***	**2.08**	**0.104**	
Year (Y)	B	1	222	60.30	<0.001	***	2.80	0.096	.
Soil history (SH)	B	1	222	79.97	<0.001	***	2.41	0.122	
Plant history (PH)	**B**	**1**	**222**	**9.25**	**0.003**	**	**1.03**	**0.312**	
Y	C	1	222	60.30	<0.001	***	2.80	0.096	.
PH	C	1	222	50.48	<0.001	***	2.73	0.100	.
SH	**C**	**1**	**222**	**38.74**	**<0.001**	***	**0.70**	**0.402**	
PH	D	1	222	86.25	<0.001	***	4.48	0.035	*
SH	D	1	222	61.88	<0.001	***	1.54	0.217	
Y	**D**	**1**	**222**	**1.39**	**0.240**		**0.22**	**0.640**	
T x SR	**A**	**3**	**222**	**9.40**	**<0.001**	***	**8.91**	**<0.001**	***
Y x SR	B	1	222	9.70	0.002	**	8.44	0.004	**
SH x SR	B	1	222	10.94	0.001	**	10.82	0.001	**

PH x SR	**B**	**1**	**222**	**7.55**	**0.007**	**	**7.48**	**0.007**	**
Y x SR	C	1	222	9.70	0.002	**	8.44	0.004	**
PH x SR	C	1	222	16.27	<0.001	***	16.11	<0.001	***
SH x SR	**C**	**1**	**222**	**2.22**	**0.137**		**2.19**	**0.140**	
PH x SR	D	1	222	23.43	<0.001	***	22.58	<0.001	***
SH x SR	D	1	222	4.36	0.038	*	3.94	0.049	*
Y x SR	**D**	**1**	**222**	**0.40**	**0.528**		**0.22**	**0.637**	

Df = degrees of freedom.

Table S1: Details on the partitioning of the treatment factor into a year, a soil history, and a plant history contrast in Models B, C, and D of Tables 2-4. The same partitioning applies to the interaction term of the treatment contrast with sown plant species richness. Model D and the contrast for year was only included, if data of 2003 were available.

Mixed effects model	Fitting sequence of contrasts	Partitioning of treatment levels
Model A	Treatment	3-4 level factor: +SH +PH, +SH -PH, -SH -PH (,2003)
Model B	(1. Year)	2-level factor (2003, 2017): +SH +PH, +SH -PH, -SH -PH versus 2003
	2. Soil history	2-level factor (+SH, -SH): +SH +PH, +SH -PH versus -SH -PH
	3. Plant history	2-level factor (+PH, -PH): +SH +PH versus +SH -PH
Model C	(1. Year)	2-level factor (2003, 2017): +SH +PH, +SH -PH, -SH -PH versus 2003
	2. Plant history	2-level factor (+PH, -PH): -SH -PH, +SH -PH versus +SH +PH
	3. Soil history	2-level factor (+SH, -SH): -SH -PH versus +SH -PH
(Model D)	1. Plant history	2-level factor (+PH, -PH): 2003, +SH -PH, -SH -PH versus +SH +PH
	2. Soil history	2-level factor (+SH, -SH): 2003, -SH -PH versus +SH -PH
	3. Year	2-level factor (2003, 2017): 2003 versus -SH -PH

Figure legends

Figure 1

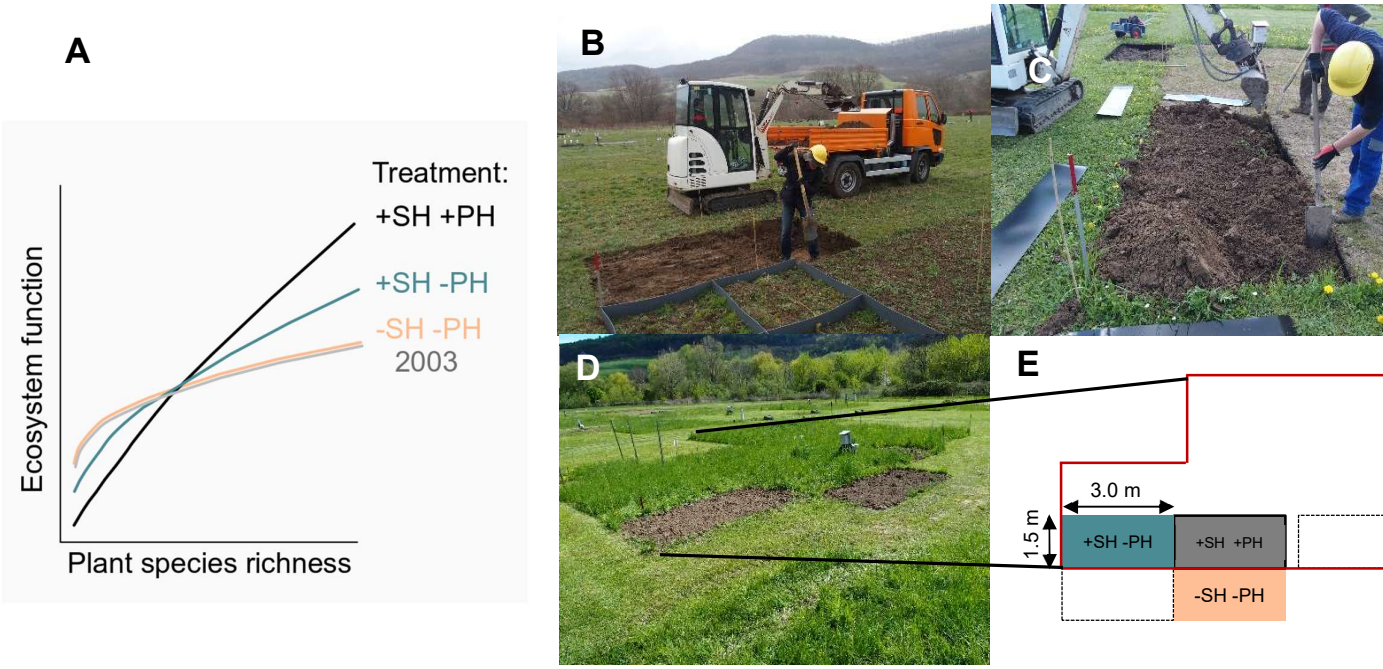

Figure 1: Hypothetical relationship of plant species richness and ecosystem functions in the different treatments of the ΔBEF-experiment (A). The soil of the -SH -PH treatment was excavated and replaced by arable soil (B), while the soil in +SH -PH was extricated from large roots (C). (D, E) shows the location of the treatments on one plot as an example. The location of treatments was assigned randomly on five potential positions (E).

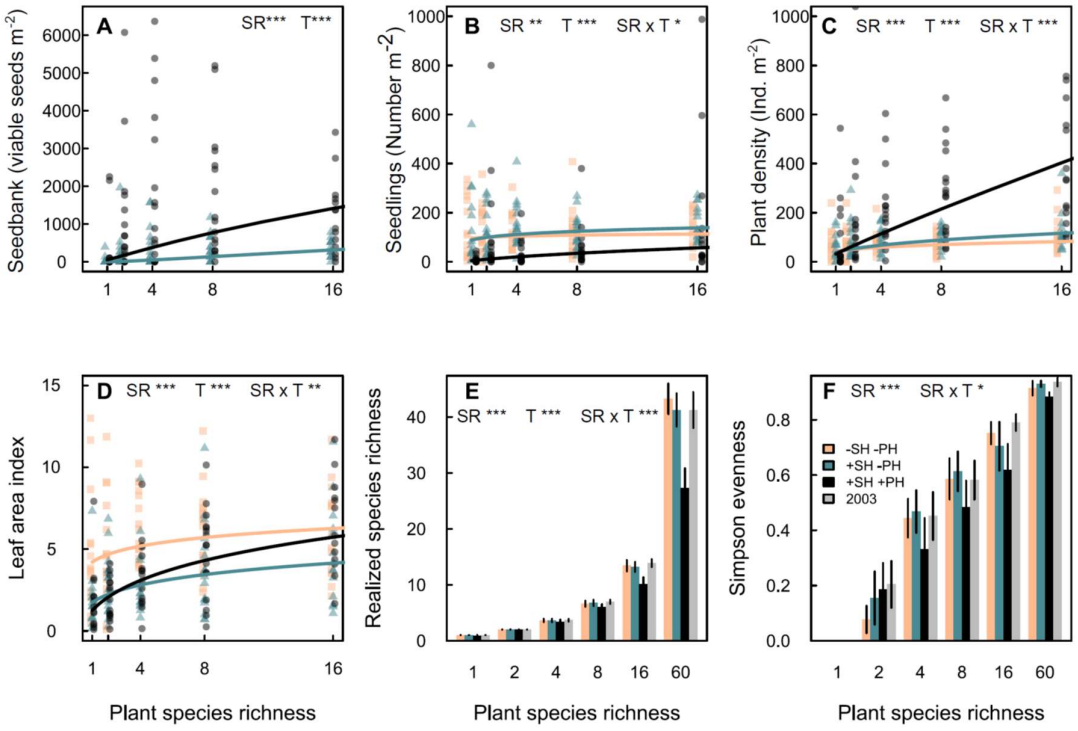

Figure 2: Plant community properties of the experimental treatments: (A) viable seeds in the seedbank belonging to the target species of each community. The treatment -SH -PH did not contain any viable seeds from the experimental species pool. (B) summed number of seedlings emerged in summer 2016 and spring 2017, (C) number of plant individuals counted in established communities in summer 2017, (D) leaf area index during peak biomass in spring 2017. (E) realized plant species richness, and (F) Simpson evenness of the plant communities based on cover estimates in spring 2017 and 2003. Lines represent the predicted means and were derived from a mixed model with a log-linear term for plant species richness (SR), treatment (T) and the interaction of both (SR x T) as fixed effects. Significant effects are indicated on top ***: $p \leq 0.001$, **: $p \leq 0.01$, *: $p \leq 0.05$, .: $p \leq 0.1$.

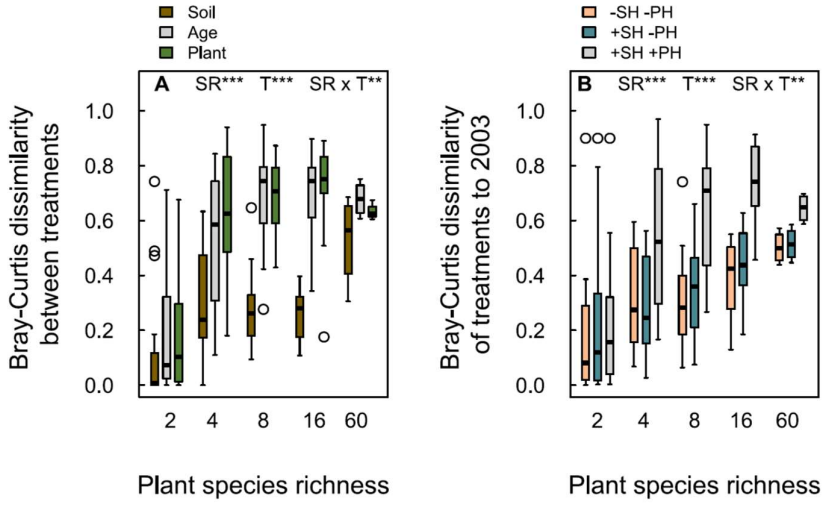

Figure 3: Pairwise Bray-Curtis dissimilarities in plant community composition (A) between two treatment pairs differing in soil history only (-SH -PH and +SH -PH, brown), plant history (+SH -PH and +SH +PH, green) as well as plant and soil history (-SH -PH and +SH +PH, grey). Bray-Curtis dissimilarities in plant community composition (B) between the single treatments and 2003 (orange: -SH -PH vs 2003; blue: +SH -PH vs 2003; grey: +SH +PH vs 2003) for the single plant species richness levels. Significant effects are indicated on top ***: p ≤ 0.001, **: p ≤ 0.01, *: p ≤ 0.05, .: p ≤ 0.1.

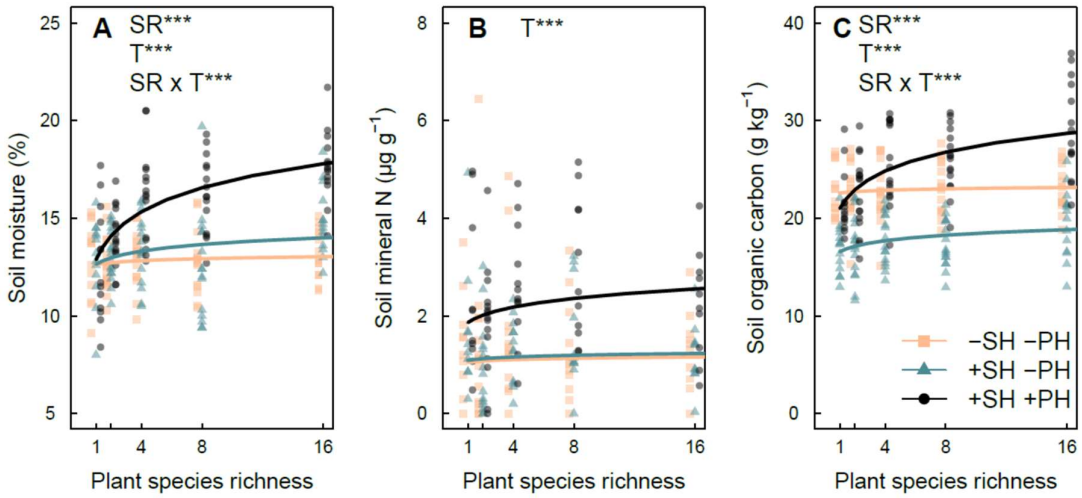

Figure 4: Soil abiotic properties of the experimental treatments: soil moisture (A), soil mineral nitrogen concentration (sum of nitrate- and ammonium-N, B), and soil organic carbon content at 0–5 cm depth. Lines represent the predicted means and were derived from a mixed model with a log-linear term for plant species richness (SR), treatment (T) and the interaction of both (T x SR) as fixed effects. Significant effects of the models are indicated on top ***: p ≤ 0.001, **: p ≤ 0.01, *: p ≤ 0.05, .: p ≤ 0.1.

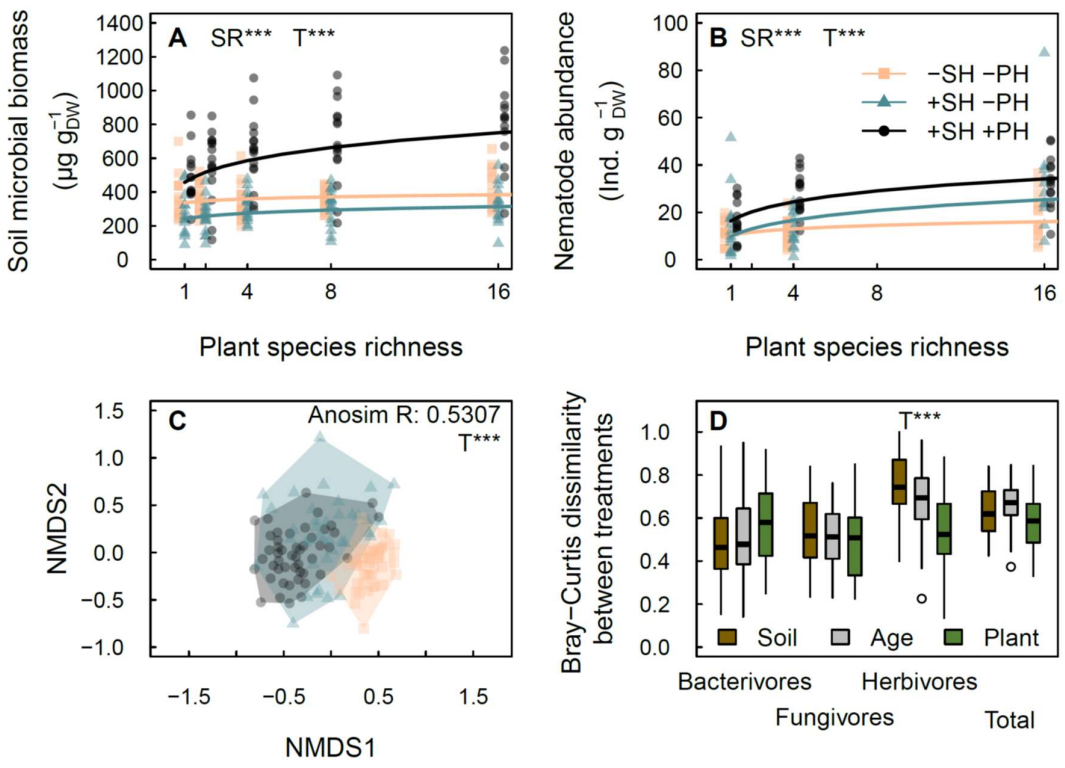

Figure 5: Soil biotic properties of the experimental treatments: microbial biomass (A) and soil nematode abundance (B). Lines represent the predicted means and were derived from a mixed model with a log-linear term for plant species richness (SR), treatment (T) and the interaction of both (T x SR) as fixed effects. (C) shows the NMDS ordination of the nematode communities of all subplots based on pairwise Bray-Curtis distances. The Bray-Curtis dissimilarities between all possible treatment pairs (D) for the single trophic groups separately and for the nematode community as a whole.

Dissimilarities were based on soil (between -SH -PH and +SH -PH, brown), age (between -SH -PH and +SH +PH, grey), and plant history (between +SH -PH and +SH +PH, green). Significant effects of the models are indicated on top ***: $p \leq 0.001$, **: $p \leq 0.01$, *: $p \leq 0.05$, .: $p \leq 0.1$.

Figure 5

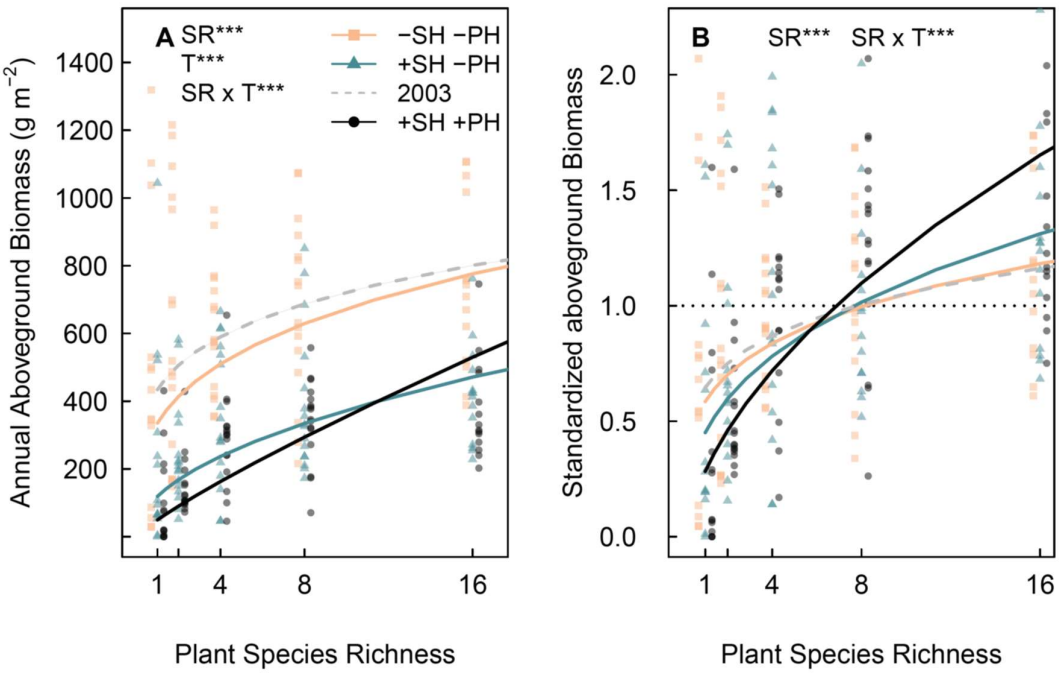

Figure 6: Annual aboveground biomass production (A) and standardized aboveground biomass (B) of the experimental treatments across the plant species richness gradient. Lines represent the predicted means and were derived from a mixed model with a log-linear term for plant species richness (SR), treatment (T) and the interaction of both (T x SR) as fixed effects. Significant effects are indicated on bottom ***: $p \leq 0.001$, **: $p \leq 0.01$, *: $p \leq 0.05$, .: $p \leq 0.1$.

Figure S1: Jaccard dissimilarities of nematode communities between all possible treatment pairs for the single trophic groups separately and for the nematode community as a whole. Dissimilarities were based on soil (between -SH -PH and +SH -PH, brown), age (between -SH -PH and +SH +PH, grey), and plant history (between +SH -PH and +SH +PH, green). Significant effects of the models are indicated on top ***: p ≤ 0.001, **: p ≤ 0.01, *: p ≤ 0.05, .: p ≤ 0.1.