Major shifts in species' relative abundance in grassland mixtures alongside positive effects of species diversity in yield

Brophy, Caroline; Finn, John A.; Lüscher, Andreas; Suter, Matthias; Kirwan, Laura; Sebastià, Maria-Teresa; Helgadóttir, Áslaug; Baadshaug, Ole H.; Bélanger, Gilles; Black, Alistair; Collins, Rosemary; op, Jure; Dalmannsdóttir, Sigridur; Delgado, Ignacio; Elgersma, Anjo; Fothergill, Michael; Frankow-Lindberg, Bodil E.; Ghesquiere, An; Golinska, Barbara; Golinski, Piotr

Published in:
Journal of Ecology

DOI:
10.1111/1365-2745.12754

Publication date:
2017

Citation for published version (APA):
Brophy, C., Finn, J. A., Lüscher, A., Suter, M., Kirwan, L., Sebastià, M-T., Helgadóttir, Á., Baadshaug, O. H., Bélanger, G., Black, A., Collins, R., op, J., Dalmannsdóttir, S., Delgado, I., Elgersma, A., Fothergill, M., Frankow-Lindberg, B. E., Ghesquiere, A., Golinska, B., ... Connolly, J. (2017). Major shifts in species' relative abundance in grassland mixtures alongside positive effects of species diversity in yield: a continental-scale experiment. *Journal of Ecology*, *105*(5), 1210-1222. https://doi.org/10.1111/1365-2745.12754

DR. JOHN ANTHONY FINN (Orcid ID : 0000-0003-3616-5563)

Received Date : 20-Jul-2016
Revised Date : 01-Dec-2016
Accepted Date : 12-Jan-2017
Article type : Standard Paper

Handling Editor: Hans Cornelissen

Major shifts in species' relative abundance in grassland mixtures alongside positive effects of species diversity in yield: a continental-scale experiment.

Caroline Brophy[1*], John A. Finn[2], Andreas Lüscher[3], Matthias Suter[3], Laura Kirwan[4], Maria-Teresa Sebastià[5,6], Áslaug Helgadóttir[7], Ole H. Baadshaug[8], Gilles Bélanger[9], Alistair Black[10,†], Rosemary P. Collins[11], Jure Čop[12], Sigridur Dalmannsdottir[7], Ignacio Delgado[13], Anjo Elgersma[14,††], Michael Fothergill[11], Bodil E. Frankow-Lindberg[15], An Ghesquiere[16], Barbara Golinska[17], Piotr Golinski[17], Philippe Grieu[18], Anne-Maj Gustavsson[19], Mats Höglind[20], Olivier Huguenin-Elie[3], Marit Jørgensen[20], Zydre Kadziuliene[21], Päivi Kurki[22], Rosa Llurba[5,6], Tor Lunnan[20], Claudio Porqueddu[23], Ulrich Thumm[24] and John Connolly[25]

[1]Department of Mathematics and Statistics, Maynooth University, Maynooth, Co Kildare, Ireland; [2]Teagasc, Environment Research Centre, Johnstown Castle, Wexford, Ireland; [3]Agroscope, Institute for Sustainability Sciences ISS, Reckenholzstrasse 191, 8046 Zurich, Switzerland; [4]UCD Institute of Food and Health, UCD, Belfield, Dublin 4, Ireland; [5]Group GAMES & Dept HBJ, ETSEA, Universitat de Lleida, Av. Rovira Roure 191, 25198 Lleida, Spain; [6]Laboratory ECOFUN, Centre Tecnologic Forestal de Catalunya, Ctra Sant Llorenç km 2, 25280 Solsona, Spain; [7]Agricultural University of Iceland, Árleyni 22, 112 Reykjavík, Iceland; [8]Department of

Plant Sciences, Norwegian University of Life Sciences, P.O. Box 5003, N-1432, Ås, Norway.; [9]Agriculture and Agri-Food Canada, 2560, Hochelaga Blvd, Québec G1V 2J3, Canada; [10]Teagasc, Beef Research Centre, Grange, Dunsany, Co. Meath, Ireland; [11]IBERS, Aberystwyth University, Plas Gogerddan, Aberystwyth, SY23 3EB, Wales, UK; [12]Biotechnical Faculty, University of Ljubljana, Jamnikarjeva 101, SI-1000 Ljubljana, Slovenia; [13]CITA-DGA, Av. Montañana 930, 50059 Zaragoza, Spain; [14]Plant Sciences Group, Wageningen University, PO Box 16, 6700, AA Wageningen, The Netherlands; [15]Department of Crop Production Ecology, Swedish University of Agricultural Sciences, Box 7043, SE-750 07 Uppsala, Sweden; [16]ILVO, Plant Science, Applied Genetics and Breeding, Caritasstraat 39, 9090, Melle, Belgium; [17]Department of Grassland and Natural Landscape Sciences, Poznan University of Life Sciences, Dojazd 11, 60-632, Poznan, Poland; [18]UMR AGIR, INP-ENSAT, University of Toulouse, F-31326, Castanet Tolosan, France; [19]Department of Agricultural Research for Northern Sweden, Swedish University of Agricultural Sciences, SE-901 83, Umeå, Sweden; [20]NIBIO, Norwegian Institute of Bioeconomy Research, P.O. Box 115, NO-1431 Ås, Norway; [21]Institute of Agriculture, Lithuanian Research Centre for Agriculture and Forestry, Akademija, LT-58344, Kedainiai, Lithuania; [22]Natural Resources Institute (Luke), Management and Production of Renewable Resources, Lönnrotinkatu 5, FI 50100, Mikkeli, Finland; [23]CNR-ISPAAM, Traversa la Crucca 3, località Baldinca, 07100, Sassari, Italy; [24]Department of Crop Science, University of Hohenheim, 70593, Stuttgart, Germany; [25]School of Mathematics and Statistics, University College Dublin, Dublin 4, Ireland; [†]Present address: Faculty of Agriculture and Life Sciences, PO Box 85084, Lincoln University, Lincoln 7647, Canterbury, New Zealand

[††]Present address: independent scientist, POB 323, 6700 AH, Wageningen, The Netherlands

*Correspondence author: Email: caroline.brophy@nuim.ie

Running heading: Species dynamics in grassland mixtures

Summary

1. Increased species diversity promotes ecosystem function; however, the dynamics of multi-species grassland systems over time and their role in sustaining higher yields generated by increased diversity are still poorly understood. We investigated the

development of species' relative abundances in grassland mixtures over three years to identify drivers of diversity change and their links to yield diversity effects.

2. A continental-scale field experiment was conducted at 31 sites using 11 different four-species mixtures each sown at two seed abundances. The four species consisted of two grasses and two legumes, of which one was fast establishing and the other temporally persistent. We modelled the dynamics of the four species mixtures, and tested associations with diversity effects on yield.
3. We found that species' dynamics were primarily driven by differences in the relative growth rates of competing species, and secondarily by density dependence and climate. The temporally persistent grass species typically had the highest relative growth rates and hence became dominant over time. Density dependence sometimes induced stabilising processes on the dominant species and inhibited shifts to monoculture. Legumes persisted at most sites at low or medium abundances and persistence improved with higher site annual minimum temperature.
4. Significant diversity effects were present at the majority of sites in all years and the strength of diversity effects was improved with higher legume abundance in the previous year. Observed diversity effects, when legumes had declined, may be due to (i) important effects of legumes even at low abundance, (ii) interaction between the two grass species or (iii) a store of N because of previous presence of legumes.
5. *Synthesis*. Alongside major compositional changes driven by relative growth rate differences, diversity effects were observed at most sites, albeit at reduced strength as legumes declined. This evidence strongly supports the sowing of diverse mixture swards that include legumes over the long-standing practice of sowing grass monocultures. Careful and strategic selection of the identity of the species used in mixtures is suggested to facilitate the maintenance of species diversity and especially

persistence of legumes over time, and to preserve the strength of yield increases associated with diversity.

Key-words: climate, determinants of plant community diversity and structure, diversity effects, grasses, legumes, multi-species communities, relative growth rate, species composition, stabilising processes, sward dynamics.

Introduction

The common practice of managing highly fertilised grassland monocultures has been often critiqued because of financial and environmental costs associated with high levels of fertiliser application, and the inability of a single species to fully utilise system resources. Thus, there is a need for productive systems that require less fertiliser and have positive environmental impacts (Tilman 1999; Lüscher *et al.* 2014; Suter *et al.* 2015). There is wide consensus that increasing species diversity often promotes ecosystem functions such as biomass yield, nitrogen yield, nutrient uptake and resistance to weed invasion (Sanderson *et al.* 2004; Hooper *et al.* 2005; Frankow-Lindberg 2012; Roscher *et al.* 2013; Suter *et al.* 2015). Complementarity among species, for example between grasses and legumes, can induce such synergistic effects in ecosystem functions (Finn *et al.* 2013). Nevertheless, over time, some species in a mixture may become dominant at the expense of other species and sward diversity may be reduced. In particular, legumes may decline or disappear (Beuselinck *et al.* 1994; Guckert & Hay 2001), thus improvements in yield due to diversity may be reduced (Carroll, Cardinale & Nisbet 2011). The extent to which declining species diversity (not only species and functional richness but also species evenness) will diminish diversity linked yield increases remains uncertain. The use of species with different rates of establishment in grass-legume mixtures may help maintain species diversity and provide

additional opportunities for complementarity along a temporal axis (Sanderson, Stout & Brink 2016). Here, we examine the dynamics of relative abundances of multiple species in agronomic grassland mixtures across a wide pedo-climatic scale of 31 coordinated multi-year experimental sites. The experimental species comprised four mutually complementary functional groups: a fast-establishing and a temporally-persistent grass, and a fast-establishing and a temporally-persistent legume. We test causes of changes in dynamics at the species level and also test the association between changing species relative abundances and diversity effects on yield. Identifying patterns in sward dynamics will contribute to our understanding of how to retain sward diversity and associated benefits over the lifetime of an agronomic system.

Potential drivers of changes in the relative abundances of species over time in mixture communities include 1) the species selected for the mixture which will differ in functional traits that underpin variation in intrinsic relative growth rates and competitive hierarchies (Aarssen 1983), 2) their initial biomasses through density dependence (Suter *et al.* 2010), 3) attributes of the environment such as climatic conditions (Van der Putten, Macel & Visser 2010), 4) management (Hebeisen *et al.* 1997; Nyfeler *et al.* 2009), and 5) biotic factors. In a hypothetical two-species system where both have the same average relative growth rate when in competition with each other, relative abundances will remain unchanged over time (assuming no competitive asymmetry due to differences in foliage architecture) (Fig.1e), but if one species has a higher relative growth rate, it will become dominant over time (Fig. 1b). The effects of differences in relative growth rates may be reduced or reinforced by interspecific and intraspecific density dependence, defined here as when the initial biomass of a species is positively or negatively correlated to its own relative growth rate (intraspecific) or to those of other species (interspecific). For example, if species 1 has a greater relative

growth rate than species 2 but species 1 also has a negative intraspecific density dependence, the combination of these two effects may act as a stabilising process on the system (defined as "any mechanism that causes species to limit themselves more than they limit others" by Adler *et al*. (2007)). Such an intraspecific density dependence may help to maintain equilibrium of the relative abundances of the two species by reducing the relative growth rate competitive advantage species 1 has over species 2 (compare a and b in Fig. 1). In multi-species mixtures, understanding sward dynamics is a more complex task due to the large number of competitive forces potentially acting within the system; in a system with *s* species there are *s* intraspecific and *s*(*s*-1) interspecific effects, all of which may affect dynamics and stability (May 2001; Ramseier, Connolly & Bazzaz 2005). Disentangling relative growth rate differences and density dependence assists in determining the stabilising or destabilising processes within the system.

Strategically selecting the species in a mixture to span a wide niche space increases the propensity of mechanisms such as complementarity and facilitation to occur, which maximises resource use and promotes ecosystem function (Hooper *et al.* 2005; Wood *et al.* 2015). For example, combining grasses and legumes in a grassland system produces synergistic effects caused by the additional N input to the system by the N_2-fixing legumes (Boller & Nösberger 1987; Carlsson & Huss-Danell 2003; Nyfeler *et al.* 2011). While diversity can positively impact on ecosystem functions in the absence of legumes (van Ruijven & Berendse 2003), legume*non-legume interaction effects may be stronger than interactions between legumes or between non-legumes (Nyfeler *et al.* 2009). However, difficulties have been identified in maintaining legume proportions in swards over agronomically relevant time periods because of vulnerability to abiotic and biotic stress (Beuselinck *et al.* 1994; Guckert & Hay 2001). The use of multiple legume species that

specifically vary in their rates of establishment and in their persistence may facilitate resource use and help maintain the contribution of legumes over the lifetime of the system (Sanderson, Stout & Brink 2016).

Grassland production is expected to be higher in mild to warm climates compared to more extreme hot or cold climates, and short-term extreme climatic events might cause deviations from the expected productivity (Isbell *et al.* 2015; Hofer *et al.* 2016). However, less is known about how climate, or short-term weather events, affect the relative growth rates and persistence of individual species within multi-species mixtures. Wachendorf *et al*. (2001) and Lüscher *et al*. (2005) found that higher air temperature increased the relative abundance of white clover in parts of the growing season. The growth and relative abundance of legumes in two-species mixtures with grasses are generally favoured in milder environments (Newton *et al.* 1994; Nolan, Connolly & Wachendorf 2001) but uncertainty remains as to how this result extrapolates to more diverse mixtures.

Here, we report on a continental-scale multi-site experiment ("The Agrodiversity Experiment", Kirwan *et al.* 2014), where we tracked the plot-level annual yields of each of four sown species over three years. The species used included two grasses and two legumes, and within each, one species was fast establishing and the other was temporally persistent; it was anticipated that the fast establishing species would have higher RGRs initially with the temporally persistent species subsequently overtaking. At each site, four species with these traits that were high performing agricultural species for forage production (yield and forage quality), given the site conditions, were selected. We evaluated the dynamics of the four-species mixtures to highlight the drivers of changes in species proportions, including differences in the relative growth rates of competing species, initial biomasses (i.e. density

dependence), and climate, and we examined stabilising and destabilising patterns in the coexistence of the four species. Besides a decline of species richness, a decline in the evenness of the system might also occur, which can considerably influence ecosystem function (Kirwan *et al.* 2007). The experiment was designed with the intention to maintain the coexistence of both grasses and legumes in mixtures by selecting species with different temporal development traits, i.e. we expected a shift in species abundance across the fast establishing and temporally persistent species but hoped for good equilibrium across the grass and legume species. We address the following questions:

(1) Do the relative abundances of the four species in mixtures change over time and to what extent do differences in the relative growth rates of competing species drive the changes?
(2) Are the effects of differences in relative growth rates altered by density dependence? Do the combination of species' relative growth rates and density dependence induce stabilising processes?
(3) Is the persistence of legumes over the experimental period related to the rapidity of establishment and persistence traits of the selected legume species and/or the grass partner(s)?
(4) Are species dynamics related to site-specific climatic variables?
(5) Is there a relationship between species' dynamics in mixture and diversity effects on yield?

Materials and methods

EXPERIMENT

Experimental mixtures of grassland species from four functional types were established over a wide geographic and climatic gradient at 30 European sites and one Canadian site (Kirwan *et al.* 2007; Kirwan *et al.* 2014). The four functional types were chosen to span gradients in N acquisition and temporal establishment. At each site, we chose a fast establishing (denoted G_F) and a temporally persistent (G_P) grass species, and a fast establishing (L_F) and a temporally persistent (L_P) N_2-fixing legume species. The identity of the four species sown varied according to species group across the 31 sites (Table 1, Table S1 in Supporting Information), where species group was decided broadly according to the climatic conditions at the site. The species groups are called Mid-European (ME, 18 sites), North European (NE, 6 sites), Moist Mediterranean (MM, 1 site), Dry Mediterranean (DM, 2 sites), West European (WE, 2 sites), Other 1 (O1, 1 site) and Other 2 (O2, 1 site). There was considerable overlap in species across the species groups, resulting in a total of eleven unique species across all experimental sites (Table 1). Categorisation of species according to their rate of establishment is a relative measure that can be informed by evidence, i.e. while *Phleum pratense* is faster establishing and less persistent than *Poa pratensis,* it is slower and more persistent than *Lolium perenne* (Klapp 1971). At each site, eleven four-species mixture stands were established in experimental plots with the initial proportion of each species varying according to a simplex design (Cornell 2002); there were four stands with one of the four species dominant (70% of one species, 10% of each of the other three), six stands with a pair of species dominant (40% of each of two species, 10% of each of the two others), and a stand with each species equally present (25% of each species). The eleven stands were replicated at low and high seed abundance (high being the monoculture seeding rate recommended by local practice and low being 60% of the high seeding rate) giving 22 plots

at each site and a total of 682 plots across the 31 sites. Monocultures of each species were also sown at each seed abundance giving a further eight plots at each site. The experiment duration (number of post-seeding years) was three years for 24 sites, two years for six sites and one year for one site (Table S1) and plots were established between 2002 and 2005. Table S1 provides the following site-specific management information: the number of harvests (2 to 7 $year^{-1}$), size of experimental plots and of the sub-plots harvested for biomass, and N fertiliser applied to all plots at the site (0 to 150 kg N ha^{-1} $year^{-1}$). While N fertiliser was applied to all plots at many sites, even the highest level of application (150 kg N ha^{-1} $year^{-1}$ at five sites) is considerably below levels that are typically applied to intensively managed grassland monocultures. Plots were not weeded during the experimental years. Full experimental details are available in the Kirwan *et al*. (2014) data paper and the site numbering convention in Table S1 is according to this database.

MEASUREMENTS AND DATA

The herbage dry matter yield (t ha^{-1}) of the four sown species and pooled weed species was recorded at each harvest over a period of up to three years at each site where year 1 was the first full year after establishment, i.e. the first year post-seeding. Sub-samples of the yield were separated and used to estimate the yield of each species type (G_F, G_P, L_F, L_P) and pooled weed species. Values at each harvest within each year were then summed to give the annual yields for each species. Some sites did not separate at each harvest but did measure total yield; in these cases, the harvests that were separated were used to estimate the components of each species in the accumulated annual yield. At each site, daily precipitation (mm) and daily minimum, mean and maximum (air) temperature (^oC) were recorded. Daily precipitation values were summed and daily mean temperature values were averaged within each year for each site. The average of the lowest ten daily minimum temperature values and

average of the highest ten daily maximum temperature values within each year was also computed for each site. Daily mean temperature was not recorded for sites 35 and 43 and the averages of daily minimum and maximum temperature values were used instead. There was a small number of other missing daily values which were omitted from calculations.

STATISTICAL ANALYSIS

We computed relative growth rates (RGRs) for each species (G_F, G_P, L_F, L_P) in each mixture plot from sowing to year 1, year 1 to year 2 and year 2 to year 3. In the period from sowing to year 1, the RGR was calculated as the natural log of the ratio of the annual yield proportion of the species in year 1 to the sowing proportion of the species. For the second two comparison periods, the RGR was calculated as the natural log of the ratio of the annual yield of a species in the latter year to the former (following Connolly & Wayne 2005; Ramseier, Connolly & Bazzaz 2005; Suter *et al.* 2007). To avoid treating zero yield values as missing, they were replaced by the minimum observed non-zero value for each species prior to RGR calculations; of 1870 measurements for each species on mixture plots across the 31 sites and three years, this occurred in 3% of cases for G1, 1% for G2, 16% for L1, and 13% for L2. Such percentages are not expected to distort the outcome of the analyses (it is not "zero-inflated" data), but rather by assigning a small non-zero value our method respects the quantitative value that the zeros should represent.

Initially, we estimated the average RGR for each species across all 31 sites from sowing to year 1 for mixture communities, using a separate linear mixed model for each species, and where the average was assumed random from site to site. Then, for each species and period of comparison year 1 to 2 and year 2 to 3, we modelled the RGRs across the mixture communities as related to the biomass of each species in the community in the previous year, (using RGR_{GF12} to represent the RGR of G_F in year 1 to 2 as an example),

$$RGR_{GF12} = \log(y_{2GF} / y_{1GF}) = \mu_j + \alpha M + \gamma_1 y^c_{1GF} + \gamma_2 y^c_{1GP} + \gamma_3 y^c_{1LF} + \gamma_4 y^c_{1LP} + \varepsilon \quad (1)$$

where y_{2GF} and y_{1GF} are the plot yields of species G_F in years 2 and 1, and, for example, y^c_{1GP} is the plot yield of G_P in the first year, centred to its average value. M is the overall sown seed abundance for the plot, coded -1 for low and +1 for high. The term μ_j is the average relative growth rate for species G_F in the *j*th species group (Table 1) at average seed abundance and average initial biomass of all species, α is the effect of sown seed abundance, and the γs are coefficients of the initial biomasses (density dependence). So γ_2 represents the effect on the relative growth rate of G_F of a unit increase in G_P in the first year. There are eight such equations, one for each species in each period of comparison and the models were fitted as random coefficient models to respect the multi-site structure of the data (Appendix S1, note 1). A positive intercept (μ_j) indicates that, at average sown seed abundance and average initial biomass of each species, the average yield of the species increased over the time period, but for a given community this increase can be magnified or dampened by the combination of all other (positive or negative α and γ) coefficients. Allowing the μ_j coefficients to vary according to species group (as opposed to having a constant μ) assumes that a species' average relative growth rate is unique to its competitive setting. For example, *Trifolium pratense* is the L_F species in both the ME and NE species groups but its average relative growth rate was allowed to differ in the contrasting competitive settings. Climatic variables (annual rainfall, and minimum, mean and maximum temperature; centred) were tested by adding them to equation 1 one at a time. We tested for pairwise species differences among the coefficients in each RGR model using a relative growth rate difference (RGRD) approach (Connolly & Wayne 2005). RGR and RGRD models were fitted using either maximum likelihood (ML, when testing fixed effects) or restricted maximum likelihood (REML, for standard error estimation). Model comparisons were made using likelihood ratio tests.

To estimate diversity effects we fitted a Diversity-Interactions model (Kirwan *et al.*

2007; Kirwan *et al.* 2009; Connolly *et al.* 2013) separately to each site and year. The model is

$$y = \sum_i \beta_i P_i + \alpha \, \mathrm{M} + \delta \, \mathrm{E} + \varepsilon \qquad (2)$$

where y is the total yield (including weed biomass), P_i denotes the sown proportion of species i, for i = G_F, G_P, L_F, L_P; M is coded -1 for low and +1 for high sown seed abundance, E = $\frac{2s}{s-1} \sum_{\substack{i,j=1 \\ i<j}}^{s} P_i P_j$ and δ is the diversity effect coefficient. Each species was sown in monoculture at high and low seed abundance at each site and these plots were included in this analysis of yield, in addition to all mixture plots. To determine whether the strength of estimated diversity effects was related to legume abundance in the previous year, we scaled the estimated diversity effect coefficients by the estimated average monoculture performance at the site (because of the wide range in yields across sites) and fitted a repeated measures regression model to the scaled diversity effect coefficients in years 2 and 3, with the average legume percentage in the preceding year as the predictor. See Appendix S1 notes 2 and 3 for further details of these models. SAS software version 9.3 (SAS Institute Inc) was used to fit all models.

Results

THE FAST ESTABLISHING GRASS DOMINATED THE INITIAL DEVELOPMENT PHASE

The fast establishing grass (G_F) had the best average relative growth rate (RGR) of the four species from sowing to year 1. Across all sites, the estimated average RGRs from sowing to year 1 were: G_F: 0.57 ($P < 0.001$), G_P: -0.49 ($P = 0.015$), L_F: -0.51 ($P = 0.161$), and L_P: -1.63 ($P < 0.001$), where P-values test against zero. The average RGR of G_F was higher than each of the other species (G_F *vs.* G_P: $P < 0.001$, G_F *vs.* L_F: $P = 0.01$, G_F *vs.* L_P: $P <$

0.001), and the average RGRs of both G_P and L_F were higher than L_P (G_P *vs*. L_P: $P = 0.005$, L_F *vs*. L_P: $P = 0.03$). On average across all plots and all 31 sites, proportions of the four species G_F, G_P, L_F, L_P in year 1 were 0.42, 0.22, 0.27, 0.10 (Fig. S1a), respectively, while average sowing proportions were 0.25 for each species. Note that relative abundances are computed relative to sown species (i.e. weeds excluded) here and for the remainder of the paper. Thus, G_F clearly showed the best, and L_P the worst, relative performance during this initial phase.

DIFFERENCES IN SPECIES' RELATIVE GROWTH RATES DOMINATED DYNAMICS

The relative abundances of the four species sown in mixtures changed substantially over the three-year period (Figs. 2 and S1). On average across all 31 sites, the temporally persistent grass (G_P) became increasingly dominant over time and in year 3 its mean relative abundance was 67% (Fig. S1a). There was considerable variation from site to site in mean relative abundances of species (Figs. 2 and S1b,c, Table S2). For example, in year 3, site-level mean G_P relative abundance ranged from 5% at the French site (site 9) to 100% at one of the Lithuanian sites (site 20). The dominance of G_P was more extreme at ME sites, where G_P was *Dactylis glomerata*, than at North European (NE) sites, where G_P was *Poa pratensis*, (Fig. S1b,c). At individual sites, G_P had the highest relative abundance at 5 of the 31 sites in year 1, at 19 of 30 sites in year 2 and at 23 of 24 sites in year 3 (Table S2).

For both the ME and NE species groups at average seed abundance and average initial biomass of each species, the average yield of G_P increased and all other species decreased in both the year 1 to 2 and year 2 to 3 time periods (Table 2, intercepts shown in the first two rows, all 8 coefficients were significant for ME but many were not for NE). All the other species groups (MM, DM, WE, O1 and O2) only included 1 or 2 sites each; thus, these intercepts are based on fewer data and are shown in Table S3. Sown seed abundance affected

only the RGR for L_F in the comparison period year 1 to 2 (Table 2). Six of the eight intraspecific density dependence coefficients were significant and had negative estimates (Table 2, highlighted in grey). For example, high year 1 biomass of G_F negatively affected its own RGR from year 1 to year 2, indicated by the negative coefficient estimate (-0.16) of y^c_{1GF} for the RGR of G_F (Table 2, first column). There were also a number of significant interspecific density dependence coefficients, indicated by bold font in the off-diagonal initial biomass coefficients (Table 2). For example, high year 1 biomass of G_P negatively affected the year 1 to year 2 RGR of G_F, indicated by the negative coefficient estimate (-0.30) of y^c_{1GP} for the RGR of G_F (Table 2, first column). For most models, the inclusion of the four density dependence coefficients was highly significant (Table 2, likelihood ratio tests).

Using model estimates (Table 2), we predicted relative abundances for the ME and NE species groups for years 2 and 3 across a range of relative abundance combinations in years 1 and 2, respectively. Predictions were first made on the biomass scale, holding total biomass of all species in the initial year constant while manipulating the individual species' components, and then converted to relative abundances for ease of interpretation (Figs. 3 and S2). G_P (*D. glomerata*) was almost always dominant for the ME group for both periods of comparison (Fig. 3). In addition, as the initial relative abundance of each species was increased, the predicted relative abundance of *D. glomerata* remained somewhat constant (Fig. 3). G_P (*P. pratensis*) was also frequently dominant across the scenarios in the NE species group (Fig. S2), although increasing the initial relative abundance of G_F (*Phleum pratense*) improved its own performance considerably in both years, reducing or reversing the dominance of *P. pratensis* (Fig. S2a,b).

BY YEAR 3, GRASSES DOMINATED OVER LEGUMES AND TEMPORALLY PERSISTENT DOMINATED OVER FAST ESTABLISHING SPECIES

On average across all sites, grasses ($G_F + G_P$) were dominant over legumes ($L_F + L_P$) from year 1 and increased their dominance to 83% in year 3; legumes declined from 36% to 25% to 17% in years 1 to 2 to 3 (Fig. S1a). At individual sites, grasses were dominant over legumes at 24 of 31 sites in year 1, at 27 of 30 in year 2 and at 23 of 24 in year 3 (Table S2). Seven of 24 sites in year 3 had almost no legumes left (<2%) (Fig. 4, Table S4). Despite the average decline in legume abundances over the three years, there was medium to high legume persistence at many sites; in year 3, four sites of 24 had above 30% legume relative abundance and a further eight had between 15 and 30% (Fig. 4).

On average across sites, fast establishing species ($G_F + L_F$) were dominant over temporally persistent species ($G_P + L_P$) in year 1 with 69% of harvested yield. By year 2, however, the persistent species had a higher relative abundance and by year 3 were dominant with 78% of harvested yield (Fig. S1a). At individual sites, there was a clear trend of dominance by fast establishing species in year 1, mixed dominance in year 2, and temporally persistent species dominance in year 3 (Fig. 4). Specifically, temporally persistent species were dominant over fast establishing species at 5 of 31 sites in year 1, at 15 of 30 sites in year 2 and at 23 of 24 sites in year 3 (Table S2).

LEGUME ABUNDANCE WAS POSITIVELY CORRELATED WITH MINIMUM TEMPERATURE

Of the four climatic variables tested, minimum temperature had the strongest effects on RGRs, with higher minimum temperature increasing the RGRs of G_P and L_F from year 2 to 3 and of L_P in both comparison periods ($P < 0.05$ in all tests, Table S5). The net effect of a higher minimum temperature was an increase in the relative abundance of legumes, which

was most evident for the NE species group (Fig. 5). In year 2, the predicted legume abundance (L_F + L_P) for NE sites was 20% with minimum temperature of -22^{o}C, which shifted to 25% with an 8 degree increase in minimum temperature to -14 ^{o}C; in year 3, the two minimum temperatures -22 and -14°C resulted in 8% and 18% legume abundance, respectively. In addition, mean site-level legume percentage (L_F + L_P) was positively correlated with minimum temperature in years 2 and 3 (Fig. S3, without outliers), further supporting that the persistence of legumes was generally favoured by higher minimum temperature and hampered by extremely cold temperatures.

DESPITE DECLINED LEGUME PROPORTION, DIVERSITY EFFECTS STILL EXISTED, BUT WITH REDUCED STRENGTH

Diversity effects were significant at 90% of 31, 83% of 30 and 75% of 24 sites in years 1, 2 and 3 respectively (Table S4), where diversity effects were estimated in each year based on sown proportions (eqn 2). While the distribution of legume abundances across sites changed considerably over the three years (Fig. 4), there was no identifiable pattern between categories of legume abundance and presence of a significant diversity effect within each year (Fig. S4). However, the strength of the diversity effect was related to legume abundance in the preceding year; thus, the magnitude of the scaled diversity effect in years 2 and 3 was higher for those sites that had higher legume abundance in the preceding year (Fig. 6, $P <$ 0.001). There was no difference between the slopes in the two years ($P = 0.06$); however, both intercepts were significantly greater than zero ($P < 0.001$ in each case), i.e. estimated average diversity effects were positive even when legume abundance in the previous year was zero.

Discussion

Substantial changes in the relative abundances of these agronomic grassland mixtures over time were primarily driven by differences in the average relative growth rates of the four species in their competitive environments and, secondarily, by climate and density dependence. On average across sites, the temporally persistent grass became dominant by year 3 and evenness among the four species was substantially reduced. Legume abundance in year 3 was low or medium at most sites but was generally greater under higher minimum temperatures. Despite considerable variation across sites and declines in overall species evenness and in legume proportions, significant diversity effects were still found at the majority of sites in all years, while the strength of observed diversity effects was higher at sites with higher proportions of legumes in the preceding year.

DIFFERENCES IN RELATIVE GROWTH RATES DROVE COMPOSITIONAL SHIFTS, WHILE DENSITY DEPENDENCE PLAYED A STABILISING ROLE

Differences in the relative growth rates of competing species was seemingly the strongest driver of species' relative abundance over the three years. Selection effects (Weigelt, Steinlein & Beyschlag 2002; Mokany, Ash & Roxburgh 2008) contribute to this mechanism: a species with a higher relative growth rate than its competitors will dominate the other species over time, ignoring further potential influences or important differences in foliage architecture (e.g. leaf angle and relative placement along plant height). The experimental species were selected according to four functional groups that we intended to be mutually complementary: combining grasses and legumes is well known to produce synergistic effects (Boller & Nösberger 1987; Carlsson & Huss-Danell 2003; Nyfeler *et al.* 2011) and using species with different rates of establishment within the grass and legume species may induce complementarity along a temporal axis (Sanderson, Stout & Brink 2016).

As anticipated, the average relative growth rates of the fast establishing grass and legume species were generally negative from years 1 to 2 and 2 to 3 and the rates of the temporally persistent grasses were positive (Tables 2 and S3, intercepts). Unexpectedly, the average relative growth rate performance of the temporally persistent legume species was generally low.

In addition to differences in average relative growth rates, there were various intra- and inter-specific density-dependent patterns in the dynamics of our multi-species communities; their joint interpretation (as in Suter *et al.* 2010) showed evidence of stabilising processes acting on the system (Adler, HilleRisLambers & Levine 2007). To demonstrate this, we generated versions of Figs. 3 and S2 in which we omitted the density dependence; these predictions are shown for just the dominant G_P species in grey in Figs. 3 and S2. The most notable result was that in year 2, the flat response of G_P remaining at approximately 70% when G_P in year 1 was increased, was not repeated in the absence of density dependence; rather, G_P continually increased in relative abundance towards monoculture (Fig. 3c, G_P black line *vs*. grey line). Thus, for the Mid-European (ME) temporally persistent grass *D. glomerata* (G_P) had a strong positive average relative growth rate in year 1 to 2, which would lead to a monoculture of *D. glomerata* if density dependence was absent. However, when its own initial biomass was larger, the species limited itself with a negative intraspecific density dependence (Table 2, G_P from year 1 to year 2) and predicted G_P relative abundance in year 2 stayed almost constant (at around 70%) as its relative abundance in year 1 varied between 20% to 60% (Fig. 3c). In an experimental system of twelve wetland species, Suter *et al*. (2010) similarly found evidence of density dependence in mixtures, but effects of individual species dominated overall trends. While the interpretation of relative growth rate models in a two-species system is relatively straightforward (Fig. 1), it is considerably more complex with more than two species (Goldberg & Fleetwood 1987). Thus, while we can

identify many patterns through model coefficients (intercepts and density dependence coefficients), the combined effects of all coefficients (Figs. 3 and S2) are particularly useful for identifying net compositional dynamics in the ecosystem (May 2001). Overall, differences in the relative growth rates of competing species primarily drove dynamics, while density dependence played an additional role, in particular through inducing stabilising processes for the ME group.

DIVERSITY EFFECTS ON YIELD SIGNIFICANT BUT REDUCED IN STRENGTH AS LEGUMES DECLINED

We found three major results related to our estimated diversity effects on yield: (1) diversity effects were significant at the majority of sites in each year (Fig. S4), (2) estimated diversity effects were reduced in strength with legume decline in the previous year across sites (positive and common slopes in Fig. 6a,b), and (3) the estimated average diversity effect was positive for at least one year after the disappearance of legumes (Fig. 6; positive intercepts). We have also clearly demonstrated several drivers of change in these four-species mixtures (differences in relative growth rates, density dependence, and climate) which led to a decline in the evenness of the relative abundances of the four species. How can we explain the significant diversity effects in all three years, despite general reductions over time in species evenness and specifically in legume proportions?

(1) Even if legume proportions are low to medium, there can be a maximum benefit to yield through positive grass-legume interactions. Only recently, Suter *et al*. (2015) demonstrated that mixtures with approximately one-third of legumes attained close to the maximum N yield, and a further increase in legumes did not improve the N output. This can be explained by strong stimulation of symbiotic N_2 fixation in

legumes by N-demand from the co-occurring grasses (Høgh-Jensen & Schjoerring 1997), with rates of symbiotic N_2 fixation in legumes and competitive N uptake in grasses being particularly high when grasses are dominant in mixtures (Nyfeler *et al.* 2011). In addition, the contributions of species' interactions to yield may be non-linearly related to species' relative abundances (Connolly *et al.* 2013), i.e. a species with low relative abundance may contribute disproportionately and substantially to diversity effects.

(2) The benefits of having had legumes in a grassland mixture may persist even after legumes decline or disappear. Legumes were sown in all mixtures and most sites had medium or high legume abundance in year 1 (Figs. 4 and S4). Even when legumes subsequently declined to very low proportions or disappeared, diversity effects in the following year were positive (Fig. 6; positive intercepts). Thus, 'legacy benefits' of legumes may partly explain diversity effects with low or no legumes. Several complex biological and environmental processes can produce legacy benefits. Legumes in mixtures contribute to increased N in the system, which may be retained and/or recycled over time, for example through root degradation, making the system more fertile even after legumes have declined.

(3) Even if legumes reduce in proportion or disappear, complementarity from grass*grass interactions may compensate in part for the reduction in grass*legume complementarity. Increases in yield due to diversity have been observed in the absence of legumes (van Ruijven & Berendse 2003).

Whatever the underlying processes, we found clear evidence of continued diversity effects on yield for three years alongside declines of species diversity (both evenness and richness); thus, having legumes initially preserved diversity effects on yield to a certain degree.

Whether diversity effects would continue for longer than three years or at what strength is unknown but of considerable importance.

LEGUME PERSISTENCE IS AFFECTED BY CLIMATE

The growth and persistence of legumes were found to be favoured in milder (Newton *et al.* 1994; Nolan, Connolly & Wachendorf 2001) and Mediterranean (Dumont *et al.* 2015) environments. This agrees with the increased persistence of legumes with higher minimum temperature, which was particularly evident in the NE species group, where an increase from -22 ^{o}C to -14^{o}C in minimum temperature predicted more than a doubling of legume relative abundance in year 3 (Fig. 5). While we showed both a positive effect of increasing minimum temperature on legume persistence (Fig. 5) and observed correlations between legume percentage and site minimum temperature in years 2 and 3 (Fig. S3b,c), the outliers in the correlation analysis show that it is possible to have good legume persistence in cold conditions (sites 7 and 33) and poor legume persistence in milder conditions (site 15). The unusually good persistence at sites 33 and 7 is likely attributed to snow cover which can facilitate winter survival of legumes in northern areas (Belanger *et al.* 2006), or grass persistence may have been hampered by fungal diseases under snow cover (Wali *et al.* 2006). These outliers may also have been caused by biotic and abiotic factors not considered here (see Suter *et al.* 2015 for a discussion). For example, inorganic N content in the soil, which depends on N mineralization being itself driven by temperature, can influence legume performance and persistence. Indeed, sudden frost (Elgersma & Schlepers 1997) or repeated freeze-thaw cycles (Brandsæter *et al.* 2002) may have detrimental effects at various stages of the legume development.

RELATIVE COMPETITIVE ABILITIES A KEY FACTOR TO IMPROVE LEGUME PERSISTENCE

There is a need to carefully select the species for use in mixtures. It is important to consider what species are suitable for the local environment, the value of the outputs of each species, and how the traits and growing patterns of the species will complement or hinder one another. We expected that strategically selecting the functional traits of the species in our agronomic mixtures across N acquisition and temporal establishment would lead to the dominance of temporally persistent species over time and would also contribute to a strong presence of legumes alongside grasses. It was not expected that grass species would become quite so dominant over legume species; nevertheless, medium (> 15%) or high (> 30%) legume abundance was observed in year 3 for 12 of the 24 sites and at these sites, the temporally persistent legume (L_P) was typically more abundant in year 3 than the fast establishing legume (L_F) (Fig. S5). Legume persistence was stronger at the NE sites than at the ME sites (Figs. S1b,c and S5). Due to its stoloniferous growth pattern, *T. repens* is a weak competitor for light, and across the ME sites it was generally not able to compete with *D. glomerata* (G_P) which has a tussock-like and tall growing pattern conducive to outcompeting *T. repens*. However, across the NE sites, *T. repens* was better able to compete against *P. pratensis* (G_P) which is also a creeping plant. Conversely, *Medicago sativa* (Lp) with its erect *habitus* showed ability to compete with *D. glomerata* in the two DM sites (Fig. S5, Table S2). Thus, the use of four species that combined two functional contrasts (grass-legume by fast establishing-temporally persistent) and the identity of the selected species within each species group contributed to legume persistence to a certain degree at many sites, but the grass-legume balance was far from a 50:50 representation. These results demonstrate the clear potential to improve our knowledge of the relative competitive abilities of grasses and legumes, and to better ensure their persistence and complementarity in mixture: a different

choice of less competitive grass species at some sites may have led to a better balance in composition of grasses and legumes over time. Recent work has also shown *D. glomerata* to be an aggressive competitor but with promising results for legume persistence when less aggressive sowing partners were selected (Malisch *et al.* 2017).

A general categorisation of legumes according to fast establishing and temporally persistent groups could be confounded by site-specific conditions. For example, a legume that is classified as temporally persistent at a site with mild climate might not persist well at an extremely cold site. Legume persistence in mixed swards could be enhanced by selecting legume species and cultivars with proven persistence abilities for the local environment and that have good competitive abilities, and by matching legumes with grasses that are less aggressive in competition (Annicchiarico & Proietti 2010). Genotypic diversity has successfully increased yields of *L. perenne* (Pollnac, Smith & Warren 2014) suggesting that enhancing functional trait diversity by combining mixtures of cultivars with mixtures of species may lead to additional diversity effects to ecosystem function than mixing species alone. Further research is needed to explore this.

In our plots, nitrogen fertiliser application varied across sites from 0 to 150 kg N ha^{-1} (Table S1). While N fertiliser is associated with higher yields, it can also reduce the competitive ability and hence persistence of legumes (Nyfeler *et al.* 2009) and may have contributed to the poor legume persistence observed at some sites (e.g. site 15). However, intensively managed grassland monocultures typically have substantially higher levels of N fertiliser applied than our maximum level of N fertiliser, and more importantly, Nyfeler *et al.* (2009; 2011) showed that fertiliser rates up to 150 kg N ha^{-1} did not negatively affect positive mutualistic interactions between grasses and legumes in mixture. Therefore, our low levels of

fertiliser application are unlikely to saturate the system with N, thereby still permitting additional N to enter through symbiotic N_2 fixation of legumes (as shown for a subset of these sites in Suter *et al.* 2015). This allows the possibility of a legume legacy effect to occur, as discussed earlier, if legumes decline or disappear.

LESSONS FOR THE DESIGN OF MULTI-SPECIES MIXTURES

Although it has been suggested that an optimal benefit from legumes in a mixture can be achieved by 20 to 50% proportional contribution to dry matter yield (Thomas 1992), our results provide evidence that positive diversity effects can still be present, albeit at reduced strength, when legume contributions have dropped below this. More importantly, we provide evidence that sowing legumes in a mixture sward can be beneficial for ecosystem function for some time (a year or two) after strong declines in legume proportion (Finn *et al.* 2013; Suter *et al.* 2015). Moreover, at the Northern European (NE) sites, positive diversity effects on yield were not accompanied by a reduction in herbage digestibility and crude protein concentration that is usually observed with increased yield (Sturludóttir *et al.* 2014). Our results are particularly relevant for one- to three-year grassland leys where it is evident that strong diversity effects on yield can be maintained. Adaptive management practices to enhance legume persistence or regeneration may be useful for short term (2-3 year) and longer term (3-5 year) managed grasslands or for intensive 'permanent' grasslands.

This continental-scale field experiment demonstrated the importance of the relative growth rates of competing species for community dynamics and species shift over time. Thus, considerable care is needed in selecting the identity of the species for use in grassland mixtures, paying particular attention to their traits and competitive abilities relative to each other. This fits well to the overwhelming evidence on benefits from pre-selection of species

with traits to enhance ecosystem functions such as biomass yield, N yield, weed suppression and drought resistance in managed grassland systems (e.g. Sanderson 2010; Roscher *et al.* 2013; Lüscher *et al.* 2014; Dooley *et al.* 2015; Hofer *et al.* 2016). While we have shown the presence of significant diversity effects on yield with reduced legume abundance, we have also shown that the strength of diversity effects were improved with higher legume proportions. Thus, we strongly recommend the use of mixture swards as an alternative to the long-standing practice of monocultures in agronomy and that legume species are included in the mixture. We also recommend choosing species with comparable competitive abilities to avoid rapid dominance of one species. Ideally the species will have high agronomic performance (proven already in monoculture) to maximise yields, and diverse traits that span functional axes (Suter, Hofer & Lüscher 2017), in addition to grass-legume, to maximise diversity effects; this could be manipulation of temporal development over years as was done here, temporal development within growing season (Husse *et al.* 2016) or rooting depth (Hofer *et al.* 2016). Our results support that mixing four agronomic species can generate strong positive diversity effects if they are selected accordingly. It is likely that increasing the number of species beyond four would further enhance diversity effects through additional system resource usage, but saturation of mixture performance could occur fast as discussed by Lüscher, Soussana & Huguenin (2011); research is urgently needed to determine this. As advocated elsewhere (Lüscher, Soussana & Huguenin-Elie 2011; Litrico & Violle 2015), we show the need for simultaneously considering the traits of candidate species for multi-species mixtures that govern their population dynamics in mixture, as well as their traits that govern their agronomic properties (yield and digestibility) and enhance resource use efficiency (through synergistic interactions).

Author's contributions

C.B. led the development of ideas, the data analysis, and the writing of the manuscript. J.Con., J.F., A.L. and M.S. assisted substantially with the ideas, methodology and writing of the manuscript. A.H., J.Con., A.L. and M.T.S. designed and coordinated the multi-site experiment. L.K., J.Con., C.B. and J.F. led the compilation of the database. All authors carried out a local experiment and / or contributed to the preparation of the manuscript. All authors gave final approval for publication.

Acknowledgements

We thank many colleagues, too numerous to list, for assistance with this work. We thank two anonymous reviewers and the associate editor for their useful comments. C.B., J.F., L.K. and J.C. were supported by Science Foundation Ireland, grant number 09/RFP/EOB2546. M.S. and J.C. were supported by the Animal Change project from the European's Seventh Framework Programme (FP7/2007-2013; grant agreement no. 266018). Co-ordination of this project was supported by the EU Commission through COST Action 852 'Quality legume based forage systems for contrasting environments'.

Data Accessibility

The data used in this paper are published in the Kirwan *et al*. (2014) data paper (doi: 10.1890/14-0170.1).

References

Aarssen, L.W. (1983) Ecological combining ability and competitive combining ability in plants: Toward a general evolutionary theory of coexistence in systems of competition. *The American Naturalist,* **122,** 707–731.

Adler, P.B., HilleRisLambers, J. & Levine, J.M. (2007) A niche for neutrality. *Ecology Letters,* **10,** 95-104.

Annicchiarico, P. & Proietti, S. (2010) White clover selected for enhanced competitive ability widens the compatibility with grasses and favours the optimization of legume content and forage yield in mown clover-grass mixtures. *Grass and Forage Science,* **65,** 318-324.

Bélanger, G., Castonguay, Y., Bertrand, A., Dhont, C., Rochette, P., Couture, L., Drapeau, R., Mongrain, D., Chalifour, F.P. & Michaud, R. (2006) Winter damage to perennial forage crops in eastern Canada: Causes, mitigation, and prediction. *Canadian Journal of Plant Science,* **86,** 33-47.

Beuselinck, P.R., Bouton, J.H., Lamp, W.O., Matches, A.G., McCaslin, M.H., Nelson, C.J., Rhodes, L.H., Sheaffer, C.C. & Volenec, J.J. (1994) Improving legume persistence in forage crop systems. *Journal of Production Agriculture,* **7,** 311-322.

Boller, B.C. & Nösberger, J. (1987) Symbiotically fixed nitrogen from field-grown white and red clover mixed with ryegrasses at low levels of ^{15}N-fertilization. *Plant and Soil,* **104,** 219-226.

Brandsæter, L.O., Olsmo, A., Tronsmo, A.M. & Fykse, H. (2002) Freezing resistance of winter annual and biennial legumes at different developmental stages. *Crop Science,* **42,** 437-443.

Carlsson, G. & Huss-Danell, K. (2003) Nitrogen fixation in perennial forage legumes in the field. *Plant and Soil,* **253,** 353-372.

Carroll, I.T., Cardinale, B.J. & Nisbet, R.M. (2011) Niche and fitness differences relate the maintenance of diversity to ecosystem function. *Ecology,* **92,** 1157-1165.

Connolly, J., Bell, T., Bolger, T., Brophy, C., Carnus, T., Finn, J.A., Kirwan, L., Isbell, F., Levine, J., Lüscher, A., Picasso, V., Roscher, C., Sebastia, M.T., Suter, M. & Weigelt,

A. (2013) An improved model to predict the effects of changing biodiversity levels on ecosystem function. *Journal of Ecology,* **101,** 344-355.

Connolly, J. & Wayne, P. (2005) Assessing determinants of community biomass composition in two-species plant competition studies. *Oecologia*, **142,** 450-457.

Cornell, J.A. (2002) *Experiments with mixtures: Designs, models, and the analysis of mixture data,* 3rd Edition edn. Wiley, Chichester.

Dooley, Á., Isbell, F., Kirwan, L., Connolly, J., Finn, J.A. & Brophy, C. (2015) Testing the effects of diversity on ecosystem multifunctionality using a multivariate model. *Ecology Letters,* **18,** 1242-1251.

Dumont, B., Andueza, D., Niderkorn, V., Lüscher, A., Porqueddu, C. & Picon-Cochard, C. (2015) A meta-analysis of climate change effects on forage quality in grasslands: specificities of mountain and Mediterranean areas. *Grass and Forage Science,* **70,** 239-254.

Elgersma, A. & Schlepers, H. (1997) Performance of white clover/perennial ryegrass mixtures under cutting. *Grass and Forage Science,* **52,** 134-146.

Finn, J.A., Kirwan, L., Connolly, J., Sebastià, M.T., Helgadóttir, A., Baadshaug, O.H., Bélanger, G., Black, A., Brophy, C., Collins, R.P., Čop, J., Dalmannsdóttir, S., Delgado, I., Elgersma, A., Fothergill, M., Frankow-Lindberg, B.E., Ghesquiere, A., Golinska, B., Golinski, P., Grieu, P., Gustavsson, A.M., Höglind, M., Huguenin-Elie, O., Jørgensen, M., Kadziuliene, Z., Kurki, P., Llurba, R., Lunnan, T., Porqueddu, C., Suter, M., Thumm, U. & Lüscher, A. (2013) Ecosystem function enhanced by combining four functional types of plant species in intensively managed grassland mixtures: a 3-year continental-scale field experiment. *Journal of Applied Ecology,* **50,** 365–375.

Frankow-Lindberg, B.E. (2012) Grassland plant species diversity decreases invasion by increasing resource use. *Oecologia,* **169,** 793-802.

Goldberg, D.E. & Fleetwood, L. (1987) Competitive effect and response in four annual plants. *The Journal of Ecology,* **1,** 1131-1143.

Guckert, A. & Hay, R.K.M. (2001) The overwintering, spring growth, and yield in mixed species swards, of white clover in Europe. *Annals of Botany,* **88,** 667-668.

Halíř, R. & Flusser, J. (1998) Numerically stable direct least squares fitting of ellipses. *6th International Conference in Central Europe on Computer Graphics and Visualization.*, pp. 125-132.

Hebeisen, T., Lüscher, A., Zanetti, S., Fischer, B., Hartwig, U., Frehner, M., Hendrey, G. & Blum, H. (1997) Growth response of *Trifolium repens* L. and *Lolium perenne* L. as monocultures and bi-species mixture to free air CO_2 enrichment and management. *Global Change Biology,* **3,** 14-160.

Hofer, D., Suter, M., Haughey, E., Finn, J.A., Hoekstra, N.J., Buchmann, N. & Lüscher, A. (2016) Yield of forage species of intensively managed temperate grassland is either largely resistant or resilient to experimental summer drought. *Journal of Applied Ecology,* **53,** 1023–1034.

Hooper, D.U., Chapin, F.S., Ewel, J.J., Hector, A., Inchausti, P., Lavorel, S., Lawton, J.H., Lodge, D.M., Loreau, M., Naeem, S., Schmid, B., Setala, H., Symstad, A.J., Vandermeer, J. & Wardle, D.A. (2005) Effects of biodiversity on ecosystem functioning: A consensus of current knowledge. *Ecological Monographs,* **75,** 3-35.

Husse, S., Huguenin-Elie, O., Buchmann, N. & Lüscher, A. (2016) Larger yields of mixtures than monocultures of cultivated grassland species match with asynchrony in shoot growth among species but not with increased light interception. *Field Crops Research,* **194,** 1-11.

Høgh-Jensen, H. & Schjoerring, J.K. (1997) Interactions between white clover and ryegrass under contrasting nitrogen availability: N_2 fixation, N fertilizer recovery, N transfer and water use efficiency. *Plant and Soil,* **197,** 187-199.

Isbell, F., Craven, D., Connolly, J., Loreau, M., Schmid, B., Beierkuhnlein, C., Bezemer, T.M., Bonin, C., Bruelheide, H., de Luca, E., Ebeling, A., Griffin, J.N., Guo, Q., Hautier, Y., Hector, A., Jentsch, A., Kreyling, J., Lanta, V., Manning, P., Meyer, S.T., Mori, A.S., Naeem, S., Niklaus, P.A., Polley, H.W., Reich, P.B., Roscher, C., Seabloom, E.W., Smith, M.D., Thakur, M.P., Tilman, D., Tracy, B.F., van der Putten, W.H., van Ruijven, J., Weigelt, A., Weisser, W.W., Wilsey, B. & Eisenhauer, N. (2015) Biodiversity increases the resistance of ecosystem productivity to climate extremes. *Nature,* **526,** 574-577.

Kirwan, L., Connolly, J., Brophy, C., Baadshaug, O., Belanger, G., Black, A., Carnus, T., Collins, R., Čop, J., Delgado, I., De Vliegher, A., Elgersma, A., Frankow-Lindberg, B., Golinski, P., Grieu, P., Gustavsson, A.-M., Helgadóttir, Á., Höglind, M., Huguenin-Elie, O., Jørgensen, M., Kadžiulienė, Ž., Lunnan, T., Lüscher, A., Kurki, P., Porqueddu, C., Sebastià, M.T., Thumm, U., Walmsley, D. & Finn, J. (2014) The Agrodiversity Experiment: three years of data from a multisite study in intensively managed grasslands. *Ecology,* **95,** 2680.

Kirwan, L., Connolly, J., Finn, J.A., Brophy, C., Luscher, A., Nyfeler, D. & Sebastià, M.T. (2009) Diversity-interaction modeling: estimating contributions of species identities and interactions to ecosystem function. *Ecology,* **90,** 2032-2038.

Kirwan, L., Lüscher, A., Sebastià, M.T., Finn, J.A., Collins, R.P., Porqueddu, C., Helgadóttir, A., Baadshaug, O.H., Brophy, C., Coran, C., Dalmannsdóttir, S., Delgado, I., Elgersma, A., Fothergill, M., Frankow-Lindberg, B.E., Golinski, P., Grieu, P., Gustavsson, A.M., Höglind, M., Huguenin-Elie, O., Iliadis, C., Jørgensen, M.,

Kadziuliene, Z., Karyotis, T., Lunnan, T., Malengier, M., Maltoni, S., Meyer, V., Nyfeler, D., Nykanen-Kurki, P., Parente, J., Smit, H.J., Thumm, U. & Connolly, J. (2007) Evenness drives consistent diversity effects in intensive grassland systems across 28 European sites. *Journal of Ecology*, **95,** 530-539.

Klapp, E. (1971) *Wiesen und Weiden: Eine Grünlandlehre (Meadows and pastures: a grassland doctrine),* 4th Edition edn. Paul Parey Verlag: Berlin und Hamburg.

Litrico, I. & Violle, C. (2015) Diversity in plant breeding: a new conceptual framework. *Trends in plant science*, **20,** 604-613.

Lüscher, A., Fuhrer, J. & Newton, P.C.D. (2005) Global atmospheric change and its effect on managed grassland systems. *Grassland: a global resource*, pp. 251-264. Wageningen Academic Publishers.

Lüscher, A., Mueller-Harvey, I., Soussana, J.F., Rees, R.M. & Peyraud, J.L. (2014) Potential of legume-based grassland–livestock systems in Europe: a review. *Grass and Forage Science,* **69,** 206-228.

Lüscher, A., Soussana, J.-F. & Huguenin-Elie, O. (2011) Role and impacts of legumes in grasslands for high productivity and N gain from symbiotic N_2 fixation. *Grassland Productivity and Ecosystem Services* (eds G. Lemaire, J. Hodgson & A. Chabbi), pp. 101-107. CABI, Wallingford, UK.

Malisch, C.S., Suter, D., BrunoStuder & Lüscher, A. (2017) Multifunctional benefits of sainfoin mixtures: Effects of partner species, sowing density and cutting regime. *Grass and Forage Science*, doi: 10.1111/gfs.12278.

May, R.M. (2001) *Stability and complexity in model ecosystems*. Princeton University Press, Princeton and Oxford.

Mokany, K., Ash, J. & Roxburgh, S. (2008) Functional identity is more important than diversity in influencing ecosystem processes in a temperate native grassland. *Journal of Ecology,* **96,** 884-893.

Newton, P.C.D., Clark, H., Bell, C.C., Glasgow, E.M. & Campbell, B.D. (1994) Effects of elevated CO_2 and simulated seasonal changes in temperature on the species composition and growth rates of pasture turves. *Annals of Botany,* **73,** 53-59.

Nolan, T., Connolly, J. & Wachendorf, M. (2001) Mixed grazing and climatic determinants of white clover (*Trifolium repens* L.) content in a permanent pasture. *Annals of Botany,* **88,** 713-724.

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

Nyfeler, D., Huguenin-Elie, O., Suter, M., Frossard, E. & Lüscher, A. (2011) Grass-legume mixtures can yield more nitrogen than legume pure stands due to mutual stimulation of nitrogen uptake from symbiotic and non-symbiotic sources. *Agriculture, Ecosystems and Environment,* **140,** 155-163.

Pollnac, F.W., Smith, R.G. & Warren, N.D. (2014) Cultivar diversity as a means of ecologically intensifying dry matter production in a perennial forage stand. *Ecosphere,* **5,** 1-12.

Ramseier, D., Connolly, J. & Bazzaz, F.A. (2005) Carbon dioxide regime, species identity and influence of species initial abundance as determinants of change in stand biomass composition in five-species communities: an investigation using a simplex design and RGRD analysis. *Journal of Ecology,* **93,** 502-511.

Roscher, C., Schumacher, J., Lipowsky, A., Gubsch, M., Weigelt, A., Pompe, S., Kolle, O., Buchmann, N., Schmid, B. & 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,** 139-149.

Sanderson, M.A. (2010) Stability of production and plant species diversity in managed grasslands: A retrospective study. *Basic and Applied Ecology,* **11,** 216-224.

Sanderson, M.A., Skinner, R.H., Barker, D.J., Edwards, G.R., Tracy, B.F. & Wedin, D.A. (2004) Plant species diversity and management of temperate forage and grazing land ecosystems. *Crop Science,* **44,** 1132-1144.

Sanderson, M.A., Stout, R. & Brink, G. (2016) Productivity, botanical composition, and nutritive value of commercial pasture mixtures. *Agronomy Journal,* **108,** 93-100.

SAS Institute Inc Version 9.3 of the SAS System for Windows. Copyright © 2002-2010 SAS Institute Inc. SAS and all other SAS Institute Inc. product or service names are registered trademarks of SAS Institute Inc., Cary, NC, USA.

Sturludóttir, E., Brophy, C., Bélanger, G., Gustavsson, A.M., Jørgensen, M., Lunnan, T. & Helgadóttir, Á. (2014) Benefits of mixing grasses and legumes for herbage yield and nutritive value in Northern Europe and Canada. *Grass and Forage Science, **69,*** 229-240.

Suter, M., Connolly, J., Finn, J., Loges, R., Kirwan, L., Sebastià, M. & Lüscher, A. (2015) Nitrogen yield advantage from grass–legume mixtures is robust over a wide range of legume proportions and environmental conditions. *Global Change Biology,* **21,** 2424-2438.

Suter, M., Hofer, D. & Lüscher, A. (2017) Weed suppression enhanced by increasing functional trait dispersion and resource capture in forage ley mixtures. *Agriculture Ecosystems & Environment*, doi: 10.1016/j.agee.2017.01.007.

Suter, M., Ramseier, D., Connolly, J. & Edwards, P.J. (2010) Species identity and negative density dependence lead to convergence in designed plant mixtures of twelve species. *Basic and Applied Ecology,* **11,** 627-637.

Suter, M., Ramseier, D., Guesewell, S. & Connolly, J. (2007) Convergence patterns and multiple species interactions in a designed plant mixture of five species. *Oecologia,* **151,** 499-511.

Thomas, R. (1992) The role of the legume in the nitrogen cycle of productive and sustainable pastures. *Grass and Forage Science,* **47,** 133-142.

Tilman, D. (1999) Global environmental impacts of agricultural expansion: The need for sustainable and efficient practices. *Proceedings of the National Academy of Sciences,* **96,** 6.

Van der Putten, W.H., Macel, M. & Visser, M.E. (2010) Predicting species distribution and abundance responses to climate change: why it is essential to include biotic interactions across trophic levels. *Philosophical Transactions Royal Society B,* **365,** 2025-2034.

van Ruijven, J. & Berendse, F. (2003) Positive effects of plant species diversity on productivity in the absence of legumes. *Ecology Letters,* **6,** 170-175.

Wachendorf, M., Collins, R.P., Elgersma, A., Fothergill, M., Frankow-Lindberg, B.E., Ghesquiere, A., Guckert, A., Guinchard, M.P., Helgadottir, A., Lüscher, A., Nolan, T., Nykänen-Kurki, P., Nösberger, J., Parente, G., Puzio, S., Rhodes, I., Robin, C., Ryan, A., Stäheli, B., Stoffel, S., Taube, F. & Connolly, J. (2001) Overwintering and

growing season dynamics of *Trifolium repens* L. in mixture with *Lolium perenne* L.: A model approach to plant-environment interactions. *Annals of Botany,* **88,** 683-702.

Wali, P.R., Helander, M., Nissinen, O. & Saikkonen, K. (2006) Susceptibility of endophyte-infected grasses to winter pathogens (snow molds). *Botany,* **84,** 1043-1051.

Weigelt, A., Steinlein, T. & Beyschlag, W. (2002) Does plant competition intensity rather depend on biomass or on species identity? *Basic and Applied Ecology,* **3,** 85-94.

Wood, S.A., Karp, D.S., DeClerck, F., Kremen, C., Naeem, S. & Palm, C.A. (2015) Functional traits in agriculture: agrobiodiversity and ecosystem services. *Trends in Ecology and Evolution,* **30,** 531-539.

Supporting Information

Appendix S1. Additional details on the statistical analysis.

Table S1. Information on the 31 experimental sites.

Table S2. The average relative abundance (%) of each species at each site and year.

Table S3. The estimated relative growth rate model coefficients for all species and periods of comparison.

Table S4. For each site and year (a) the average percentage of legumes in mixture and (b) the estimated diversity effect coefficient (δ) for yield (total plus weed).

Table S5. Estimates, standard errors and likelihood ratio tests for climatic variables in the models of relative growth rate responses.

Figure S1. The relative abundance of yield for each species (G_F, G_P, L_F and L_P) and each year averaged over (a) all 31 sites, (b) the 18 Mid-European (ME) sites, and (c) the six North European (NE) sites.

Figure S2. Predicted relative abundances for the North European (NE) sites in years 2 and 3 as affected by the relative abundance of each species in years 1 and 2, respectively.

Figure S3. Average legume percentage *versus* minimum temperature for each site and year.

Figure S4. The percentage of sites with average legume abundance in each of the categories low (0-15%), medium (15-30%), and high (30-100%), for each year, and the number of sites with significant diversity effects in each category.

Figure S5. Average relative abundance in each year for the twelve sites that had medium or high legume abundance in year 3.

Table 1. Species identities by species group and functional group classifications along with the number of sites per species group.

	Species functional group classifications				
	Grass		Legume		
	Fast establishing	Temporally persistent	Fast establishing	Temporally persistent	
Species group	G_F	G_P	L_F	L_P	No. of sites
1 Mid-European (ME)	*Lolium perenne* L.	*Dactylis glomerata* L.	*Trifolium pratense* L.	*Trifolium repens* L.	18
2 North European (NE)	*Phleum pratense* L.	*Poa pratensis* L.	*Trifolium pratense* L.	*Trifolium repens* L.	6
3 Moist Mediterranean (MM)	*Lolium perenne* L.	*Dactylis glomerata* L.	*Trifolium pratense* L.	*Medicago sativa* L.	1
4 Dry Mediterranean (DM)	*Lolium rigidum* L.	*Dactylis glomerata* L.	*Medicago polymorpha* L.	*Medicago sativa* L.	2
5 West European (WE)	*Lolium perenne* L.	*Phleum pratense* L.	*Trifolium repens* L.	*Trifolium ambiguum* M. Bieb.	2
6 Other 1 (O1)	*Lolium perenne* L.	*Phleum pratense* L.	*Trifolium pratense* L.	*Trifolium repens* L.	1
7 Other 2 (O2)	*Phleum pratense* L.	*Festuca arundinacea* Schreb.	*Trifolium pratense* L.	*Trifolium repens* L.	1

Table 2. Relative growth rate (RGR) model coefficients for each species and each period of comparison. Intraspecific density dependence coefficients are highlighted in grey. The final row provides likelihood ratio test (LRT) *P*-values for the inclusion of the four initial biomasses.

	Year 1 to year 2								Year 2 to year 3							
	G_F		G_P		L_F		L_P		G_F		G_P		L_F		L_P	
Intercepts																
ME	**-0.90**	a	**0.97**	b	**-1.37**	a	**-1.22**	a	**-1.60**	a	**0.22**	b	**-3.09**	c	**-1.71**	a
NE	**-0.87**	a	0.22	b	**-1.46**	a	-0.69	ab	-0.92	a	0.05	a	-1.22	a	-1.06	a
Seed abundance	0.03	a	0.01	a	**0.13**	a	0.02	a	-0.02	a	-0.01	a	-0.11	a	-0.09	a
Initial biomass (density dependence)																
y^c_{1GF}	**-0.16**	a	**-0.08**	b	-0.06	b	-0.06	ab	y_{2GF} -0.08	a	**0.06**	bc	0.10	b	0.04	ac
y^c_{1GP}	**-0.30**	ab	**-0.58**	a	**-0.19**	b	**-0.19**	ab	y_{2GP} -0.07	a	**-0.09**	a	0.01	a	-0.06	a
y^c_{1LF}	-0.06	a	**-0.05**	a	**-0.13**	b	-0.10	ab	y_{2LF} 0.03	ab	**0.05**	ab	0.08	a	-0.05	b
Y^c_{1LP}	0.00	a	-0.05	a	**-0.24**	b	**-0.55**	b	Y_{2LP} 0.02	a	**0.09**	a	-0.05	ac	**-0.46**	bc
LRT *P*-value	<0.001		<0.001		<0.001		<0.001		<0.001		<0.001		0.059		0.050	

The Mid-European (ME) and North European (NE) values are the intercept coefficients for these species groups; they estimate the mean relative growth rate for each species in their respective competitive setting and period of comparison, at average seed abundance and average initial biomass of each species. Intercepts for the MM, DM, WE, O1 and O2 species groups were included in the model but are displayed in Table S3, given the small number of sites in these groups. The estimated density dependence coefficients shown here hold for all species groups (i.e. they were not crossed with species group as the intercept was). Bold indicates significantly different from 0 at α=0.05. Within each row and each period of comparison, coefficients that share a letter are not significantly different from one another. It follows, that letters on coefficients are only comparable within but not across rows.

Fig. 1. Hypothetical scenarios of how the relative abundances of two species (Sp 1 and 2) may change due to varying the average relative growth rate of species 2 relative to species 1 (*y*-axis) and varying the intraspecific density dependence of species 1 (*x*-axis). Displayed in each pie chart is the final composition of the community after a fixed period of time assuming the initial composition of each community to have an equal share of the two species. For illustrative purposes, it is further assumed that the average relative growth rate of species 1 is fixed and that its interspecific density dependence on species 2 is neutral.

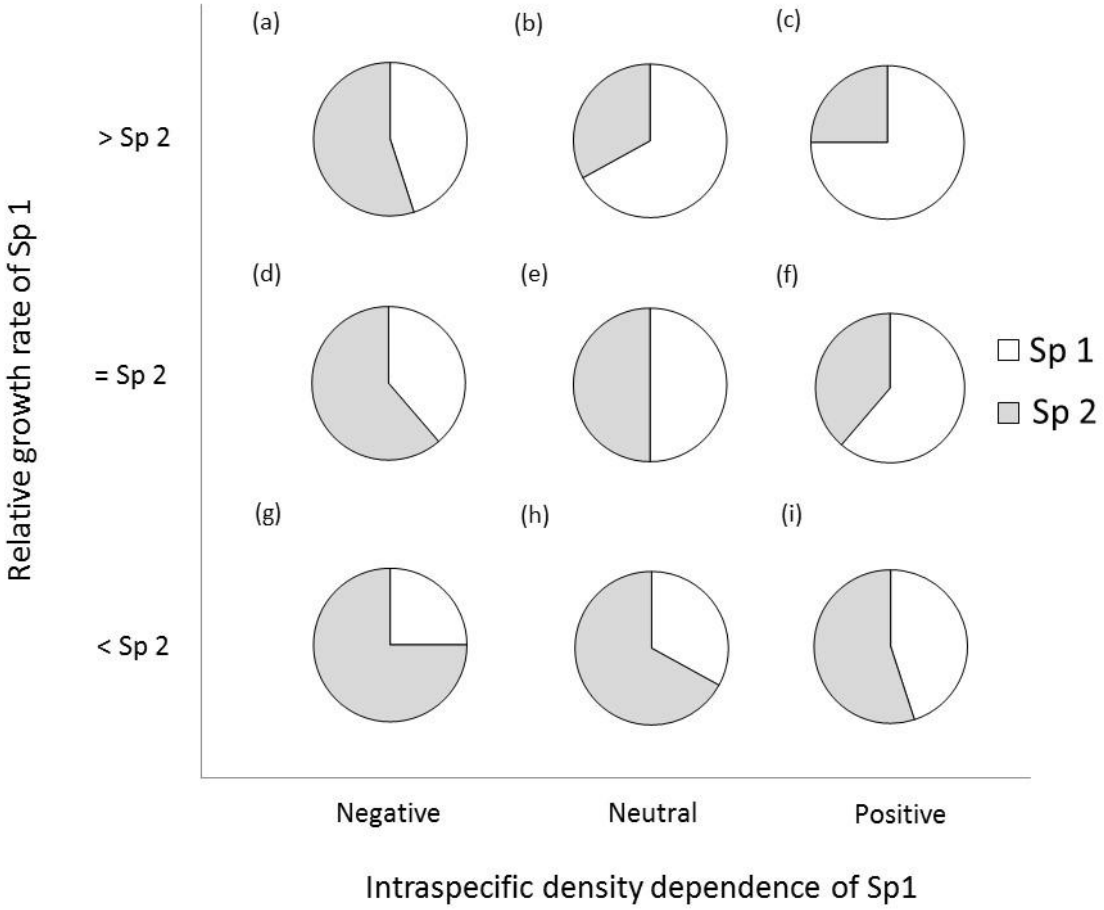

Fig. 2. The distribution of site-level relative abundance means for each species (G_F, G_P, L_F and L_P) and year across all sites; each box displays the upper quartile, median and lower quartile of the site-level means.

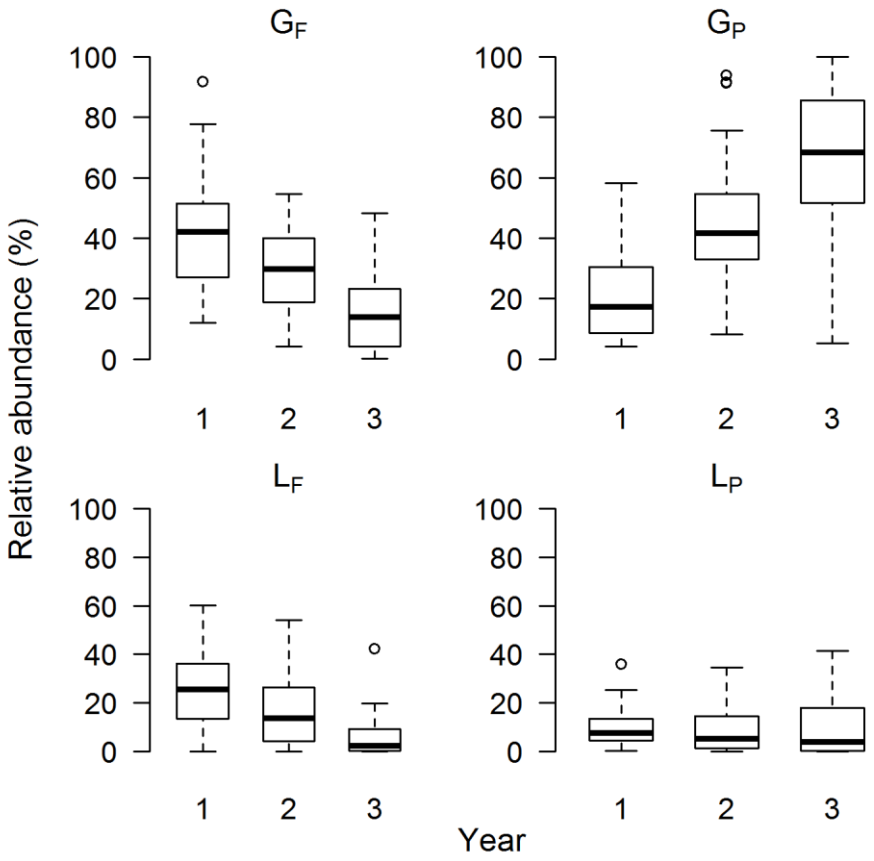

Fig. 3. Predicted species relative abundances for the Mid-European (ME) sites in years 2 and 3 as affected by the relative abundance of each species in years 1 and 2, respectively. The total biomass in the initial year is kept constant at the ME average of 11.9 (year 1) and 11.3 (year 2); likewise, the relative abundances of the three species other than the target species (on the x-axis) are kept equal. Predictions are made at average seed abundance and respect the ranges of the predictor variables in the observed data. *D. glomerata* predictions in grey are ignoring density dependence.

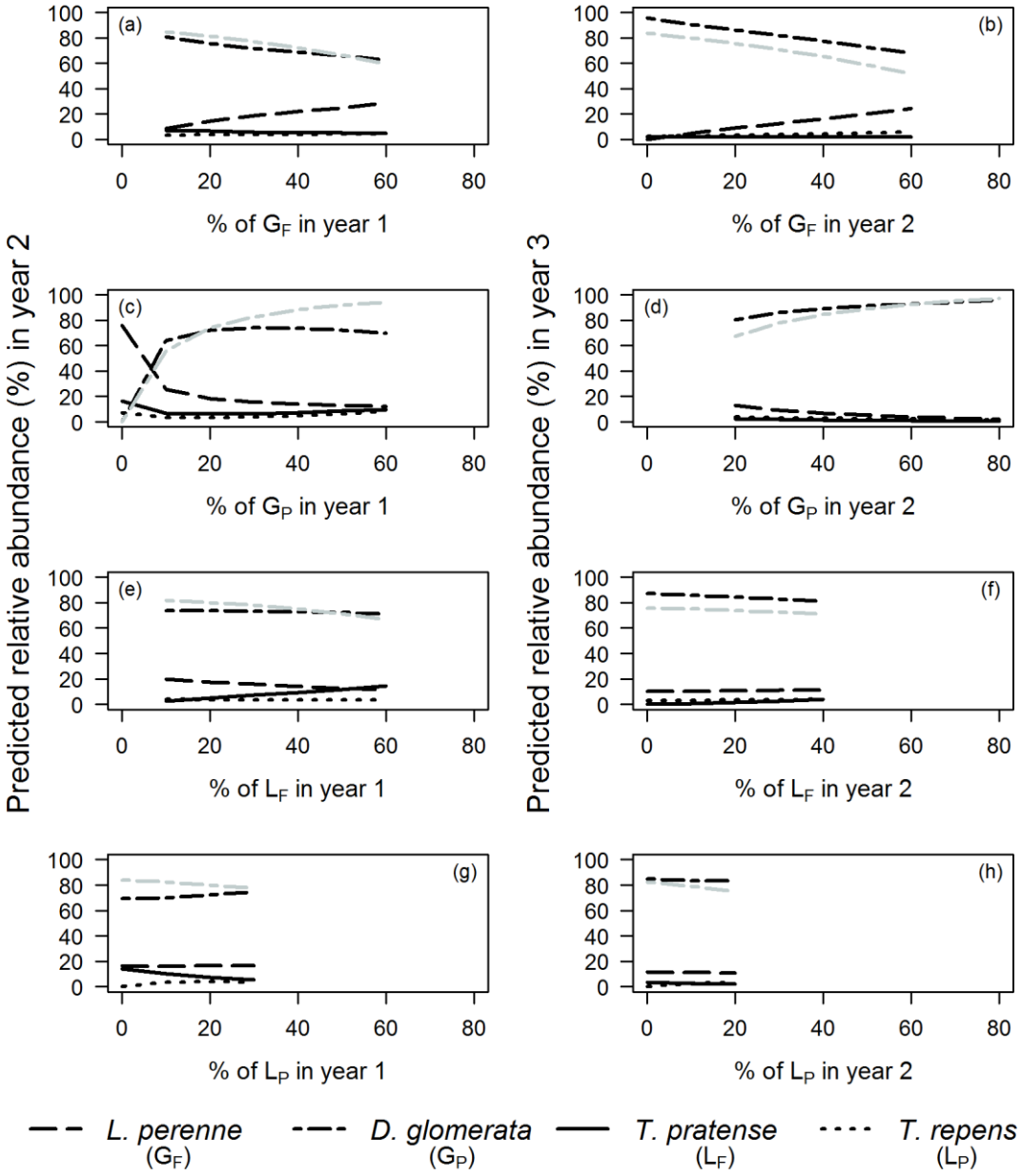

Fig. 4. Legume percentage ($L_F + L_P$) *versus* temporally persistent ($G_P + L_P$) percentage on average for each year (year 1 ○, year 2 ●, year 3 ●) and each site. The dotted horizontal lines indicate low (0-15%), medium (15-30%) and high (>30%) legume abundance categories. The left side of the plot indicates dominance (>50%) of fast establishing species, the right side dominance of temporally persistent species. The ellipses highlight the general trend in each year and were estimated using the methods of Halíř and Flusser (1998).

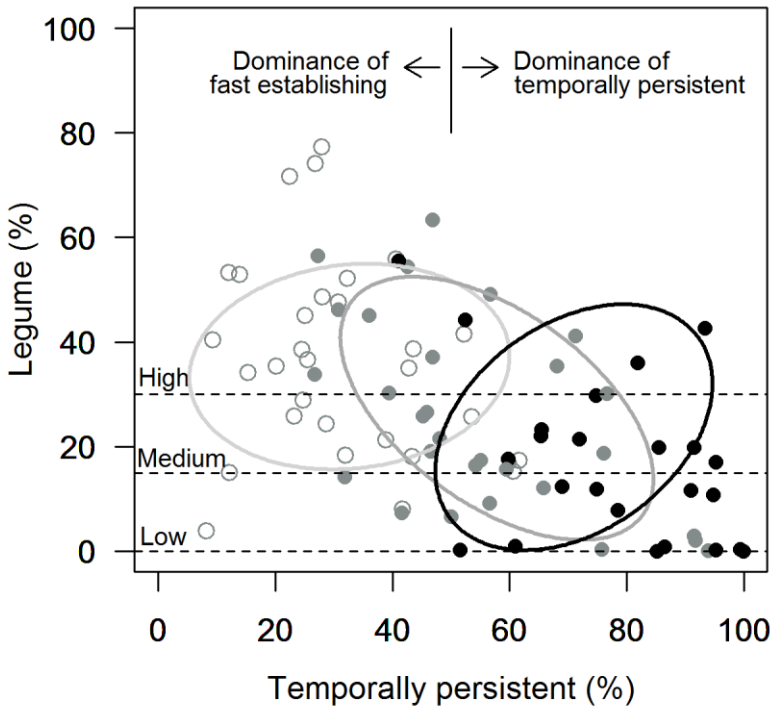

Fig. 5. Predicted relative species abundances for the Mid-European (ME) and the North European (NE) sites in years 2 and 3 as affected by minimum temperatures in years 1 and 2, respectively. Predictions are made at average sown seed abundance and average species biomasses in the preceding year. Annual minimum temperature for each site was computed as the average of the lowest ten daily minimum temperature values.

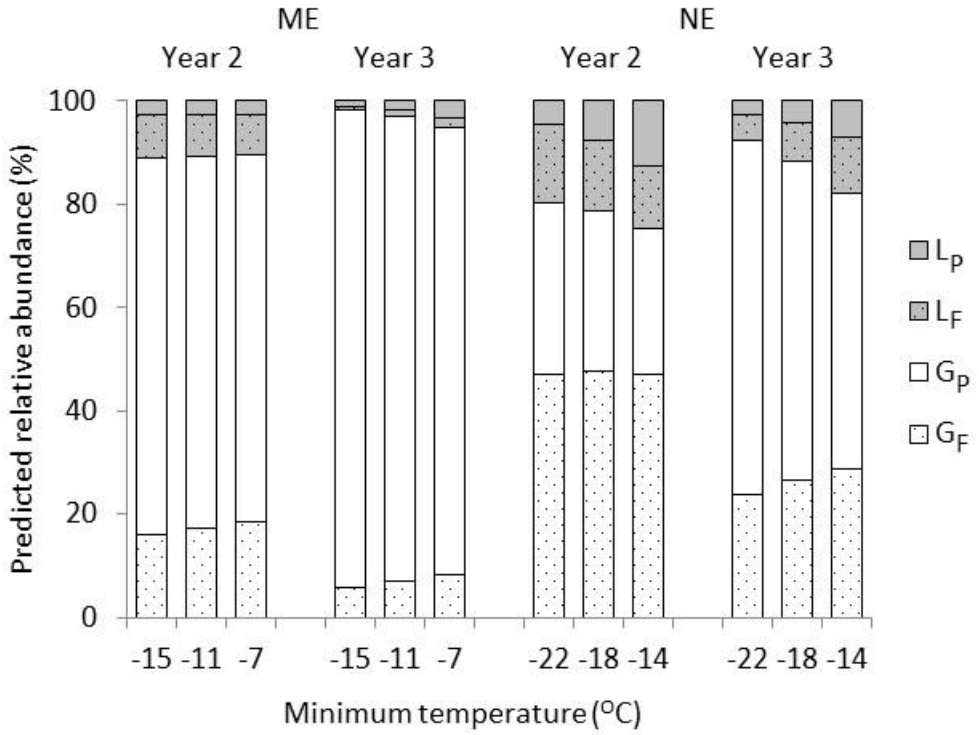

Fig. 6. The estimated diversity effect on yield coefficient scaled by the average monoculture performance in (a) year 2 and (b) year 3 *versus* average legume percent in the preceding year at each site. The diversity effect coefficient (δ) and monoculture performances (β_i's) were estimated at each site using sown proportions (eqn 2). Sites where the diversity effect was not significant are indicated by an unfilled circle (with site number indicated below the *x*-axis label). The vertical dashed grey lines indicate legume abundance categories (low=0-15%, medium=15-30% and high=30-100%).