Supporting Information

Appendix S1: Supporting text and information on monoculture variation.

We view the yield of a multi-species mixture to be the sum of the proportional contribution of monocultures and the effects of interspecific interactions on yield. The greater the differences among monoculture yields, the stronger the overyielding that is required for a mixture to achieve transgressive overyielding. (However, the role of selection effects may also need to be considered.) Biodiversity studies in semi-natural grasslands have a relatively wide range in yield among the monocultures, and representative species are selected from the local species pool. Agronomic experiments use species that have been bred for high quantity and quality of yield; this 'screening' contributes to some reduction in the magnitude of the yield differences across monocultures of different species. For example, Nyfeler *et al*. (2009) calculated that the yield of the highest yielding monoculture was 1.93 times the yield of the average monoculture in Cardinale *et al*.'s (2007) meta-analysis. For total yield in Fig. 1 of our experiment, the corresponding average value across 31 sites was 1.24 (s.d. = 0.11, range 1.06 – 1.43). Note, however, that there was still a considerable range in yield among the monocultures in our experiment, as evidenced by the difference between the best-performing monoculture and the average of the monoculture performances (Fig. 1). Regardless of the degree of variation among the monocultures, considerable levels of complementarity were still required to achieve transgressive overyielding, especially of the magnitude that was generally recorded in this experiment.

References

Cardinale, B.J., Wrigh, J.P., Cadotte, M.W., Carroll, I.T., Hector, A., Srivastava, D.S., Loreau, M. & Weis, J.J. (2007) Impacts of plant diversity on biomass production increase through time because of species complementarity. *Proceedings of the National Academy of Sciences of the United States of America,* **104,** 18123-18128.

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.

Appendix S2: Supporting text and information on the methods and analyses

Our design resulted in four planned levels of initial evenness (Table S2, Supporting information). Species proportions at sowing were based on the locally-recommended seeding rates (kg ha^{-1}) used for monocultures for each species at a site, and the lower level of sown abundance was 60% of the higher. In general, weeding was not implemented at sites (see Table S1 in Supporting Information for the few exceptions), and were managed according to recommendations for agricultural grasslands under local conditions, which varied across sites (2 to 7 harvests year $^{-1}$; 0 to 150 kg ha^{-1} year^{-1} of nitrogen applied as mineral fertiliser) but were always consistent across plots within a site (see Table S1 for details). The first year of yield data was based on the first whole year of harvests in the year after sowing (for further details on plant species, methods and sites, see Kirwan *et al*. 2007 and Table S1, Supporting Information).

STATISTICAL METHODS

The statistical methods used to address the objectives 3 and 4 are described in further detail here.

The effect of community evenness on annual yield was modelled using the diversity-interaction models of Kirwan *et al.* (2009). This series of models relates ecosystem function to species identity and interaction effects. It is a flexible modeling system which facilitates identification of patterns in species interaction effects related to properties such as community evenness and species functional traits. In the analysis of the first year data from this experiment, Kirwan *et al*. (2007) found that the diversity effect (excess of mixture performance over that expected from monocultures) was related to community evenness.

Model specification – fixed effects:

The models are of the form y = ID+DE, where y is the ecosystem function, ID is the species identity effects and DE is the diversity effect. By testing for patterns in the pairwise interaction effects, we can test for the effects of community evenness (strength of interaction equal for all species) and functional traits (species that share functional traits have the same strength of interaction) on the diversity effect. The basic model for a single year:

$$\begin{aligned} & y=\sum_{i=1}^{11}\beta_i P_i && \text{identity effects (ID)}\\ & +\sum_{j\neq k}\delta_{jk}P_{\mathrm{j}}.P_k && \text{interaction effects (DE)}\\ & +\alpha M && \text{seed density}\\ & +\varepsilon && \text{residual}\end{aligned}$$

As different species were chosen to represent the four functional types at different geographical and climatic regions, there were 11 species in total (only four at any one site). The ID effects are specified by including species proportions. P_i is the proportion of the i^{th} species in the community. For a monoculture, P_i =1, and so the coefficient β_i is the estimated species performance in monoculture. For a mixture, $\Sigma\beta_i P_i$ represents the mixture performance expected from monoculture performances (i.e. in the absence of species interactions). The DE is then the excess of mixture performance over that expected from monoculture. The interactions among the four functional types are included in the model through the pairwise products of their proportions ($P_j P_k$, j,k = NF, NP, ZF, ZP), resulting in six pairwise interaction terms ($\delta_{NF.NP}$, $\delta_{ZF.ZP}$, $\delta_{NF.ZF}$, $\delta_{NP.ZP}$, $\delta_{NF.ZP}$, $\delta_{NP.ZF}$). Because we aim to test the general effect of the functional traits on yield, we include pairwise interactions among the four functional types in the model rather than species' pairwise interactions. For example, $\delta_{NF.NP}$, the coefficient of the product of the proportions of NF and NP, estimates the interaction effect of a fast-establishing N-fixing legume with a slow-establishing N-fixing legume, even though these functional types could be represented by different species across different sites. Seed density (M) is included as a factor in the model (with levels -1 and +1, so that all other terms are estimated at average density.

Model selection

A series of models were fitted to identify patterns in the interaction effects (Kirwan *et al.* 2009). Model 0 contains only the ID effects and seed density (dens). If the interaction coefficients are not significantly different from each other, then a single interaction term will describe the diversity effect ($\sum_{j\neq k}\delta_{jk}P_{\mathrm{j}}.P_k$ becomes $\delta\sum_{j\neq k}P_{\mathrm{j}}.P_k$). Kirwan *et al.* (2007) have shown that this interaction term is a function of community evenness (where E = (2s/(s-1))$\sum_{j\neq k}P_j P_k$). Model 1 contains ID effects and this evenness (E) term. Kirwan et al (2007) found that the relationship between E and the diversity effect was not linear and that a quadratic E term was required. Model 2 tests for this quadratic relationship. Model 3 is as described above with ID and two-species (pairwise) interaction terms. Model 4 is more complex and tests whether three-species interactions are required.

Model 0	ID + dens
Model 1	ID + E + dens
Model 2	ID + E + E^2 + dens
Model 3	ID + pairwise + dens
Model 4	ID + pairwise + three-species + dens

Model specification - random effects

The model was fitted to the data using a random coefficients approach, in which we estimate the general or population relationship, while allowing for variation about this relationship across sites. The identity and interaction effects are included in the model as both fixed and random terms. The fixed effects test for the general effect of the terms across sites. The random term estimates the variation in the effect across sites. Including these random coefficients, the model then becomes:

$$
\begin{aligned}
y &= \sum_{i=1}^{11} (\beta_i + b_i) P_i \\
&+ \sum_{j \neq k} (\delta_{jk} + d_{jk}) P_{\mathrm{j}} . P_k \\
&+ (\alpha + a) M \\
&+ \varepsilon
\end{aligned}
$$

Here, there are fixed and random terms for each effect. For example, β_i is the general identity effect of species i, and b_i estimates the variation in this effect across sites. Here, $b_i \sim \mathrm{N}(0,\sigma_{\mathrm{bi}}^2)$, $d_{jk} \sim \mathrm{N}(0,\sigma_{\mathrm{djk}}^2)$, $a_t \sim \mathrm{N}(0,\sigma_{\mathrm{a}}^2)$, $\varepsilon \sim \mathrm{N}(0,\sigma^2)$. Due to the large number of random coefficients that require estimation, some constraints were added. The variance components for the species identity effects (σ_{bi}^2) were constrained to be equal for grasses and equal for legumes. The variance components for the interaction effects depended on the model being fitted. For example, if an evenness model was fitted with a single interaction coefficient, then there was a single random coefficient for that evenness term.

Model specification – estimation of effects across three years

To generalize this model across the three years of the experiment, we add another layer of fixed and random effects. The interaction of each fixed term with year (specified as a factor) assesses how the species identity and interaction effects change through time. Year is also included as a random effect, to account for the repeated measurements taken on plots.

$$\begin{aligned} y &= \sum_{t=1}^{3}\sum_{i=1}^{11}(\beta_{it}+b_{it})\,P_{it}*year_t \\ &+\sum_{t=1}^{3}\sum_{j\neq k}(\delta_{jkt}+d_{jkt})\,P_{\mathrm{jt}}.P_{kt}*year_t \\ &+(\alpha_t+a_t)\,M*year_t \\ &+\boldsymbol{\varepsilon} \end{aligned}$$

Here, $b_{it}\sim\mathrm{N}(0,\sigma_{\mathrm{bit}}^2)$, $d_{jkt}\sim\mathrm{N}(0,\sigma_{\mathrm{djkt}}^2)$, $a_t\sim\mathrm{N}(0,\sigma_{\mathrm{at}}^2)$, $\boldsymbol{\varepsilon}\sim\mathrm{N}(0,\mathbf{R})$. Similar constraints are added as above, with the variance components for the species identity effects (σ_{bit}^2) constrained to be equal for grasses and equal for legumes, but allowed to differ for each year. The variance in each year and the year-to-year correlations (**R** matrix) were estimated as an unstructured variance-covariance matrix.

An additional complication to the generalisation of this model across time is the fact that the community composition also changes through time, and the species proportions P_{it} need careful specification. One approach is to use the planned sowing proportions for each of the three years. The planned proportions are the treatment that was applied at the time of sowing. They are also the proportions that a farmer would sow. Because community composition can change through time, however, the observed effects may include both selection and diversity effects.

An alternative is to specify the P_{it} as the species proportion of annual biomass in the previous year (sowing proportions for year1). We refer to these as ‘realised proportions’. This approach reduces the likelihood of diversity effects being due to changes in community composition from the intended or initial composition. It also has practical relevance as farmers can apply adaptive management to maintain a particular community composition or level of evenness to optimize diversity effects. As the plots were not weeded, there will be a proportion of weed included in the ID effects for years 2 and 3. The interpretation of this term is that the greater the weed proportion in year 1, the greater the contribution of the weed identity effect to biomass in year 2.

The significance range of values between fitted evenness curves and the maximum diversity effect (Figs. 4 and S2) were calculated based on the approach of Johnson & Neyman (1936), as applied by Suter *et al*. (2007).

RESULTS

Models 0-4 were fitted to annual total yield and annual yield of sown species across all available years and all sites. The goodness-of-fit of the candidate models was compared using Akaike's Information Criterion (AIC) and results are presented in Tables S3 and S5. The lower the value of the AIC statistic, the better the model fit. For both annual total yield and annual yield of sown species, the best fitting model was Model 2, in which yield has a quadratic relationship with evenness. This is consistent with the model found to fit best in a previous analysis of the first year's data (Kirwan *et al.* 2007).

References

Johnson, P.O. & Neyman, J. (1936) Tests of certain linear hypotheses and their application to some educational problems. *Statistical Research Memoirs*, **1**, 57-93.

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

Kirwan, L., Luescher, A., Sebastia, M.T., Finn, J.A., Collins, R.P., Porqueddu, C., Helgadottir, A., Baadshaug, O.H., Brophy, C., Coran, C., Dalmannsdottir, S., Delgado, I., Elgersma, A., Fothergill, M., Frankow-Lindberg, B.E., Golinski, P., Grieu, P., Gustavsson, A.M., Hoglind, M., Huguenin-Elie, O., Iliadis, C., Jorgensen, 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.

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.

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.

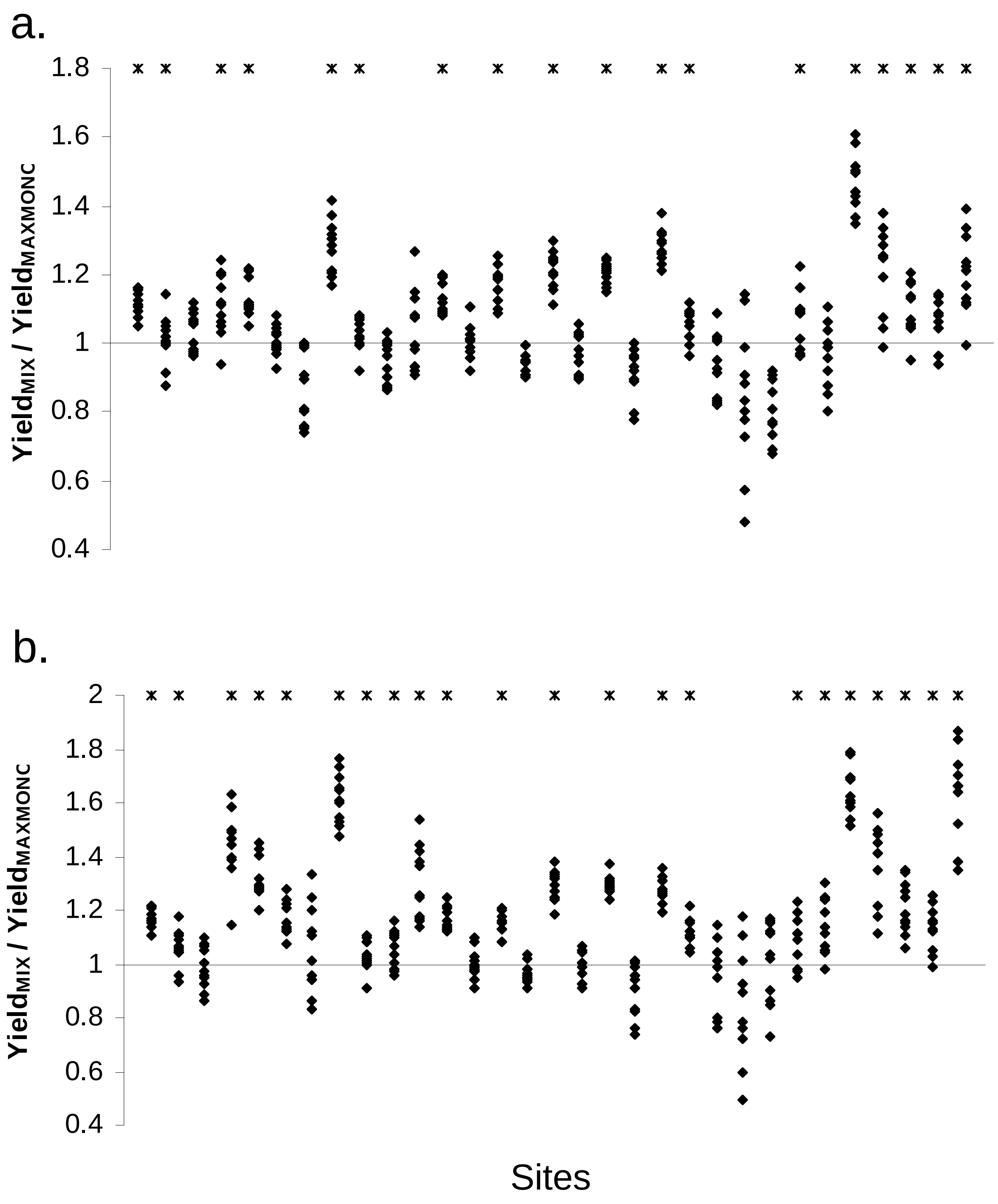

Fig. S1. Ratio of yield of each mixture community to yield of the best-performing monoculture (across years) at each site based on a) total yield and b) yield of sown species.

Each point represents the sum of aboveground biomass over multiple years. Significant transgressive overyielding is indicated by an asterisk over a site at the top of each panel. Sites arranged in order of decreasing total yield of the best-performing monoculture (as in Fig. 1 and Table S1).

Table S1. Information on sites, choice of sown species, basic management information and weather information. Sites are distinguished by their identifier number in the database from which these data are derived.

Sites are arranged in order of declining total yield of the best-performing monoculture over the three years (as in Fig. 1).

Site # in database	Country	Site	Latitude	Longitude	Altitude (m a.s.l.)	Species* (ZF, ZP, NF, NP)	Nitrogen fertiliser (kg ha^{-1} per annum)	Harvests per annum (year 1,2,3)	Size of plots (m^2)	Annual rainfall (mm)	Annual mean temp (°C)	No. of years
1	Belgium	Merelbeke	50°59'N	3°49E	11	Lp, Pp, Tp, Tr	150	4,3,4	8.4	780	9.9	3
40	Slovenia	Ljubljana	46°3'N	14°28'E	300	Lp, Dg, Tp, Tr	120	4	8.6	1147	10.7	2
15	Ireland	Wexford	52°16'N	6°30'W	54	Lp, Dg, Tp, Tr	150	5	16.0	1033	10.1	3
21	Netherlands	Wageningen	51°58'N	5°40'E	7	Lp, Dg, Tp, Tr	0,108,108	5	6.0	760	9.6	3
10	Germany	Renningen	48°46’N	9°11’E	460	Lp, Dg, Tp, Tr	150	4,5,5	18.0	693	8.2	3
11	Germany	St. Johann	48°28'N	9°18E	700	Lp, Dg, Tp, Tr	150	4	18	1046	7.4	2
9	France	Auzeville Tolosane	43°05'N	1°43'E	162	Lp, Dg, Tp, Ms	120	3,2,3	6.0	680	13.0	3
34	Switzerland	Zurich-Reckenholz**	47°26'N	8°32'E	491	Lp, Dg, Tp, Tr	150	5	18.0	1031	9.4	3
24	Norway	Ås	59°40’N	10°51’E	95	Lp, Dg, Tp, Tr	135	3	12.0	785	5.3	3
35	Wales	Aberystwyth	52°26'N	4°01'W	30	Lp, Dg, Tp, Tr	90	4	6.0	1038	9.7	3
43	Ireland	Athenry	53°17'N	8°44’W	40	Lp, Pp, Tr, Ta	75	7	10	885	10.4	2
36	Wales	Bronydd Mawr	51°57'N	3°37'W	323	Lp, Dg, Tp, Tr	93	4,3,4	6.0	1500	8.2	3
27	Poland	Brody	52°26'N	16°18'E	91.4	Lp, Dg, Tp, Tr	90	4,3	9.0	587	8.0	2
44	Ireland	Moorepark	52°8'N	8°16'W	48	Lp, Pp, Tr, Ta	100	7	10	1207	9.5	2
26	Poland	Brody	52°26'N	16°18'E	94.2	Lp, Dg, Tp, Tr	120	4,4,3	9.0	587	8.0	3
22	Norway	Saerheim	58°46'N	5°39'E	90	Lp, Dg, Tp, Tr	0	3	12.0	1180	7.1	3
33	Sweden	Öjebyn (Piteå)	65°19'N	21°24'E	5	Pp, Poa, Tp, Tr	60	2,3,2	19.0	539	2.1	3

23	Norway	Tromsø	69°40'N	18°56'E	15	Pp, Poa, Tp, Tr	60	2	21.0	1031	3.1	3
19	Lithuania (b)	Dotnuva	55°24'N	23°50'E	71	Lp, Dg, Tp, Tr	120	3,2,2	6.5	650	6.1	3
7	Finland	Mikkeli***	61°40'N	27°13'E	107	Pp, Fa, Tp, Tr	60	3	16.0	643	3.1	2
25	Norway	Løken	61°07'N	9°04'E	435	Pp, Poa, Tp, Tr	80	2	10.5	576	1.6	3
18	Lithuania (a)	Dotnuva	55°24'N	23°50'E	71	Lp, Dg, Tp, Tr	120	3	47.5	650	6.1	3
30	Spain	Gosol	42°13'N	1°39'E	1410	Lp, Dg, Tp, Tr	0	2	8.25	948	7.9	1
16	Italy	Ottava	40°44'N	8°32'E	80	Lr, Dg, Mp, Ms	31, 57, 61	4,5,5	9.0	547	16.2	3
20	Lithuania (c)	Dotnuva	55°24'N	23°50'E	71	Lp, Dg, Tp, Tr	120	3,3,2	24.0	650	6.1	3
52	Canada	Lévis	46°46'N	71°12'W	43	Pp, Poa, Tp, Tr	60	2	12	1175	5.26	3
31	Sweden (a)	Svalöv	55°55'N	13°07'E	55	Lp, Dg, Tp, Tr	0	3	8.8	700	7.7	3
32	Sweden (b)	Svalöv	55°55'N	13°07'E	55	Lp, Dg, Tp, Tr	0	3	8.8	700	7.7	3
13	Iceland (a)	Korpa	64°09'N	21°45'W	35	Pp, Poa, Tp, Tr	40	2	6.0	900	4.5	3
14	Iceland (b)	Korpa	64°09'N	21°45'W	35	Pp, Poa, Tp, Tr	80	2	10.0	900	4.5	3
28	Spain	Zaragoza**	41°44'N	2°53'E	225	Lr, Dg, Mp, Ms	61	2,3,1	9.0	409	14.3	3

*Species codes as follows. Non-fixing grasses: Dg = *Dactylis glomerata*, Fa = *Festuca arundinaceum*, Lp = *Lolium perenne*, Lr = *Lolium rigidum*, Pp = *Phleum pratense*, Poa = *Poa pratensis.* N$_2$-fixing legumes: Mp = *Medicago polymorpha*, Ms = *Medicago sativa*, Ta = *Trifolium ambiguum*, Tp = *Trifolium pratense*, Tr = *Trifolium repens*.

**The plots at Switzerland were weeded for the first two harvests in the establishment year and the plots at Spain-Zaragoza were weeded only in year 1 of harvesting.

***Barley was used on all plots as a nurse crop at establishment and harvested in the establishment year, as per conventional practice at this site.

Table S2. The simplex design with the proportions of the four functional types used at each site of the experiment and the corresponding level of species richness and evenness (*E*) at sowing.

These fifteen communities were repeated at two levels of seed density, resulting in 30 plots per site. The lower level of seed density was 60% that of the high level; the high level of seed density was based on the local recommended sowing rate for each monoculture. See Methods for codes for functional types.

Community	ZF	ZP	NF	NP	Species richness	Initial evenness
1	0.25	0.25	0.25	0.25	4	1
2	0.4	0.4	0.1	0.1	4	0.88
3	0.4	0.1	0.4	0.1	4	0.88
4	0.4	0.1	0.1	0.4	4	0.88
5	0.1	0.4	0.4	0.1	4	0.88
6	0.1	0.4	0.1	0.4	4	0.88
7	0.1	0.1	0.4	0.4	4	0.88
8	0.7	0.1	0.1	0.1	4	0.64
9	0.1	0.7	0.1	0.1	4	0.64
10	0.1	0.1	0.7	0.1	4	0.64
11	0.1	0.1	0.1	0.7	4	0.64
12	1.0	0.0	0.0	0.0	1	0
13	0.0	1.0	0.0	0.0	1	0
14	0.0	0.0	1.0	0.0	1	0
15	0.0	0.0	0.0	1.0	1	0

Table S3. Summary of fitted models of annual total yield (including weed biomass), regressed on 'realised' species proportions.

Model fit is presented for the series of models, arranged in increasing order of complexity. The best fitting model (in bold) was the quadratic evenness model.

Model	Fixed Terms	Fixed parameters	Random parameters	AIC
Model 0	(ID + Dens)*Year	39	15	9677
Model 1	(ID + E+ Dens)*Year	42	18	8888.1
Model 2	**(ID + E+ E^2 + Dens)*Year**	**45**	**18**	**8769.9**
Model 3	(ID+pairwise+Dens)*year	57	18	8942.2
Model 4	(ID+pairwise+E^2+Dens)*year	60	18	8852.2
Model 5	(ID+pairwise+3-sp+Dens)*year	69	18	8896.7

Table S4. Analysis of annual total yield (including weed biomass), regressed on 'realised' species proportions. Estimated coefficients for the quadratic evenness model are presented here (model 2 in Table S3).

Parameter	Year of harvest	Estimate	Standard error	DF	No. of sites	t-value	P
Lolium perenne	1	8.35	0.65	56	22	12.79	<.0001
	2	7.87	0.75	56	21	10.55	<.0001
	3	7.36	0.92	56	16	7.97	<.0001
Phleum pratense	1	6.80	0.97	24	10	7.04	<.0001
	2	7.26	1.07	24	10	6.75	<.0001
	3	5.77	1.38	24	7	4.19	0.0003
Lolium rigidum	1	3.53	2.15	3	2	1.64	0.1992
	2	3.77	2.40	3	2	1.57	0.2144
	3	3.90	2.66	3	2	1.46	0.2396
Dactylis glomerata	1	8.19	0.67	55	21	12.28	<.0001
	2	9.08	0.80	55	20	11.42	<.0001
	3	7.51	0.90	55	17	8.35	<.0001
Poa Pratensis	1	4.59	1.25	15	6	3.68	0.0022
	2	5.01	1.42	15	6	3.53	0.0031
	3	4.00	1.49	15	6	2.68	0.017
Festuca annua	1	6.53	3.06	0	1	2.14	0.2783
	2	7.57	3.39	0	1	2.23	0.2684
Trifolium pratense	1	9.78	0.64	72	27	15.32	<.0001
	2	8.74	0.68	72	26	12.77	<.0001
	3	7.33	0.93	72	22	7.86	<.0001
Medicago polymorpa	1	2.51	2.33	2	2	1.08	0.3941
	2	4.99	2.62	2	2	1.91	0.1967
	3	13.99	5.68	2	1	2.46	0.1327
Trifolium repens	1	7.28	0.63	73	28	11.62	<.0001
	2	7.07	0.72	73	27	9.84	<.0001
	3	6.34	0.99	73	21	6.4	<.0001
Medicago sativa	1	4.50	1.91	6	3	2.36	0.056
	2	7.60	2.08	6	3	3.66	0.0106
	3	9.38	2.28	6	3	4.12	0.0062
Trifolium ambiguum	1	6.38	2.34	2	2	2.73	0.1123
	2	9.60	4.57	2	2	2.1	0.1707
Weed	2	8.18	1.02	52	30	8	<.0001
	3	6.29	0.80	52	24	7.82	<.0001
Seed density	1	0.17	0.08	2066	31	2.01	0.0442
	2	0.12	0.07	2066	30	1.76	0.0789
	3	-0.03	0.07	2066	24	-0.46	0.6431
E	1	5.95	0.51	82	31	11.64	<.0001
	2	7.04	0.66	82	30	10.69	<.0001
	3	3.97	0.68	82	24	5.84	<.0001
E^2	1	-3.27	0.46	82	31	-7.05	<.0001
	2	-5.22	0.60	82	30	-8.69	<.0001
	3	-3.06	0.59	82	24	-5.23	<.0001

Table S5. Summary of fitted models of yield of sown species (excluding weed biomass), regressed on 'realised' species proportions.

Model fit is presented for the series of models, arranged in increasing order of complexity. The best fitting model (in bold) was the quadratic evenness model.

Model	**Fixed Terms**	**Fixed parameters**	**Random parameters**	**AIC**
Model 0	(ID + Dens)*Year	39	15	10287.3
Model 1	(ID + E+ Dens)*Year	42	18	9308.6
Model 2	**(ID + E+ E^2 + Dens)*Year**	**45**	**18**	**9157.6**
Model 3	(ID+pairwise+Dens)*year	54	15	9292
Model 4	(ID+pairwise+E^2+Dens)*year	60	18	8897.2
Model 5	(ID+pairwise+3-sp+Dens)*year	69	18	9221.2

Table S6. Analysis of yield of sown species (excluding weeds), regressed on ‘realised’ species proportions. The estimated coefficients for the quadratic evenness model are presented (model 2 in Table S5).

Parameter	Year of harvest	Estimate	Standard error	DF	No. of sites	t-value	P
Lolium perenne	1	7.28	0.67	56	22	10.88	<.0001
	2	6.77	0.75	56	21	9.04	<.0001
	3	6.07	0.92	56	16	6.6	<.0001
Phleum pratense	1	5.91	0.99	24	10	5.96	<.0001
	2	6.76	1.08	24	10	6.27	<.0001
	3	5.24	1.37	24	7	3.82	0.0008
Lolium rigidum	1	2.44	2.21	3	2	1.1	0.3500
	2	1.09	2.41	3	2	0.45	0.6819
	3	1.39	2.66	3	2	0.52	0.6378
Dactylis glomerata	1	6.84	0.68	55	21	9.99	<.0001
	2	8.40	0.80	55	20	10.47	<.0001
	3	6.89	0.90	55	17	7.69	<.0001
Poa Pratensis	1	3.26	1.28	15	6	2.55	0.0222
	2	4.95	1.43	15	6	3.46	0.0035
	3	3.67	1.48	15	6	2.47	0.0258
Festuca annua	1	5.69	3.13	1	1	1.82	0.3199
	2	7.32	3.40	1	1	2.15	0.2772
Trifolium pratense	1	8.31	0.65	72	27	12.82	<.0001
	2	6.29	0.77	72	26	8.15	<.0001
	3	4.85	0.93	72	22	5.19	<.0001
Medicago polymorpa	1	1.44	2.37	2	2	0.61	0.6051
	2	0.54	2.95	2	2	0.18	0.872
	3	-3.90	5.84	2	1	-0.67	0.5726
Trifolium repens	1	5.50	0.64	73	28	8.65	<.0001
	2	4.47	0.81	73	27	5.55	<.0001
	3	4.02	0.99	73	21	4.05	0.0001
Medicago sativa	1	2.79	1.93	6	3	1.44	0.1996
	2	5.16	2.35	6	3	2.2	0.0703
	3	8.19	2.29	6	3	3.58	0.0117
Trifolium ambiguum	1	3.12	2.37	2	2	1.32	0.3184
	2	-1.56	4.94	2	2	-0.32	0.7823
Weed	2	4.00	0.94	52	30	4.24	<.0001
	3	1.54	0.58	52	24	2.65	0.0106
Seed Density	1	0.27	0.09	2066	31	3.02	0.0025
	2	0.21	0.08	2066	30	2.78	0.0055
	3	0.08	0.07	2066	24	1.04	0.2973
E	1	7.77	0.56	82	31	13.85	<.0001
	2	-9.30	0.72	82	30	12.93	<.0001
	3	-5.12	0.72	82	24	7.08	<.0001
E^2	1	-4.35	0.52	82	31	-8.3	<.0001
	2	-6.47	0.67	82	30	-9.69	<.0001
	3	-3.44	0.64	82	24	-5.41	<.0001

Table S7. Contributions of the separate pairwise interactions among the four functional types to the diversity effect.

These results are from an analysis of annual average of total yield (including weed biomass) regressed on the initial design proportions of species. The analysis was based on 24 sites with three years of data. The identity effects are presented for each species, and the number of sites at which a species occurred is in brackets. The pairwise interaction effects for the four functional types are indicated by the relevant combination of trait levels (N, Z, F, P) for N_2-fixing, non-fixing, fast-establishment and temporal persistence respectively, and the text in brackets indicates the trait levels being tested. In the column of model estimates (± s.e.), the pairwise interaction effects are represented by the values of the δ coefficients. The pairwise interaction effects are also scaled for the mixture with equal proportions of all four species ($\delta_{ij}p_i.p_j$, where $p_i = 0.25$), for which the diversity effect is the sum of the pairwise interaction effects = 4.308 t ha^{-1} $year^{-1}$.

Model term	**Species type**	**Model estimates (t ha^{-1} $year^{-1}$)**	**Interaction effects on yield for equi-proportional mixture (t ha^{-1} $year^{-1}$)**	**P**
Lolium perenne (n = 16)	ZF	6.82 ± 0.76	-	<.0001
Phleum pratense (n = 7)	ZF*	6.05 ± 1.15	-	0.0015
Lolium rigidum (n= 2)	ZF	1.52 ± 2.16	-	0.5944
Dactylis glomerata (n= 17)	ZP	6.67 ± 0.74	-	<.0001
Poa pratensis (n= 6)	ZP	3.82 ± 1.25	-	0.0231
Trifolium pratense (n= 22)	NF	5.96 ± 0.68	-	<.0001
Medicago polymorpha (n = 2)	NF	0.87 ± 2.26	-	0.7575
Trifolium repens (n = 21)	NP	4.29 ± 0.70	-	<.0001
Medicago sativa (n = 3)	NP	3.6 ± 1.84	-	0.179
Seed density	-	0.16 ± 0.08	-	0.0581
ZF*ZP (fast, persistent)	-	5.26 ± 1.86	0.329	0.0086
NF*NP (fast, persistent)	-	15.66 ± 1.86	0.979	<.0001
NF*ZF (N_2-fixing, non-fixing)	-	9.84 ± 1.86	0.615	<.0001
NP*ZP (N_2-fixing, non-fixing)	-	13.19 ± 1.86	0.824	<.0001
NP*ZF (all four trait levels)	-	12.25 ± 2.08	0.766	<.0001
NF*ZP (all four trait levels)	-	12.72 ± 2.08	0.795	<.0001

*In this analysis, an exception occurred at site 1, where *P. pratense* was used as a non-fixing, temporally persistent species (ZP).

Table S8. Incidence of switching in the identity of the best-performing functional type of monoculture within each year at each site (based on total yield).

Indicated are the identities of the functional type of monoculture in each year, and whether switching occurred (0= no, 1 = yes) from year 1 to year 2, and from year 2 to year 3. Hyphens indicate years when data was unavailable. Codes: N= N_2-fixing legume, Z = non-fixing grass, F = fast-establishing, and P = temporally persistent. Sites are ordered as in Table S1.

Site	Country	Site	Year 1	Year 2	Year 3	Switch year 1 to 2	Switch year 2 to 3
1	Belgium	Merelbeke	NF	NF	NF	0	0
40	Slovenia	Ljubljana	NF	NF	-	0	-
15	Ireland	Wexford	ZP	ZP	ZP	0	0
21	Netherlands	Wageningen	NF	NF	ZP	0	1
10	Germany	Renningen	NF	ZP	ZP	1	0
11	Germany	St. Johann	NF	ZP	-	1	-
9	France	Auzeville Tolosane	NF	NF	NF	0	0
34	Switzerland	Zurich-Reckenholz	NF	NF	ZP	0	1
24	Norway	Ås	NF	ZP	ZP	1	0
35	Wales	Aberystwyth	NF	NF	NF	0	0
43	Ireland	Athenry	NF	NP	-	1	-
36	Wales	Bronydd Mawr	NF	ZP	NP	1	1
27	Poland	Brody	ZP	NF	-	1	-
44	Ireland	Moorepark	NF	ZP	-	1	-
26	Poland	Brody	ZF	ZP	ZP	1	0
22	Norway	Saerheim	NF	NF	ZF	0	1
33	Sweden	Öjebyn (Piteå)	NF	NF	NF	0	0
23	Norway	Tromsø	NF	ZF	NF	1	1
19	Lithuania	Dotnuva	ZF	ZP	ZP	1	0
7	Finland	Mikkeli	ZF	ZP	-	1	-
25	Norway	Løken	ZF	ZF	NF	0	1
18	Lithuania	Dotnuva	ZF	ZP	ZP	1	0
30	Spain	Gosol	NF	-	-	-	-
16	Italy	Ottava	NF	NP	NP	1	0
20	Lithuania	Dotnuva	ZP	ZP	NP	0	1
52	Canada	Lévis	NF	ZF	NP	1	1
31	Sweden	Svalöv	NF	NP	NP	1	0
32	Sweden	Svalöv	NF	NP	NP	1	0
13	Iceland	Korpa	NF	ZP	NP	1	1
14	Iceland	Korpa	ZF	ZF	ZF	0	0
28	Spain	Zaragoza	ZP	ZF	ZP	1	1

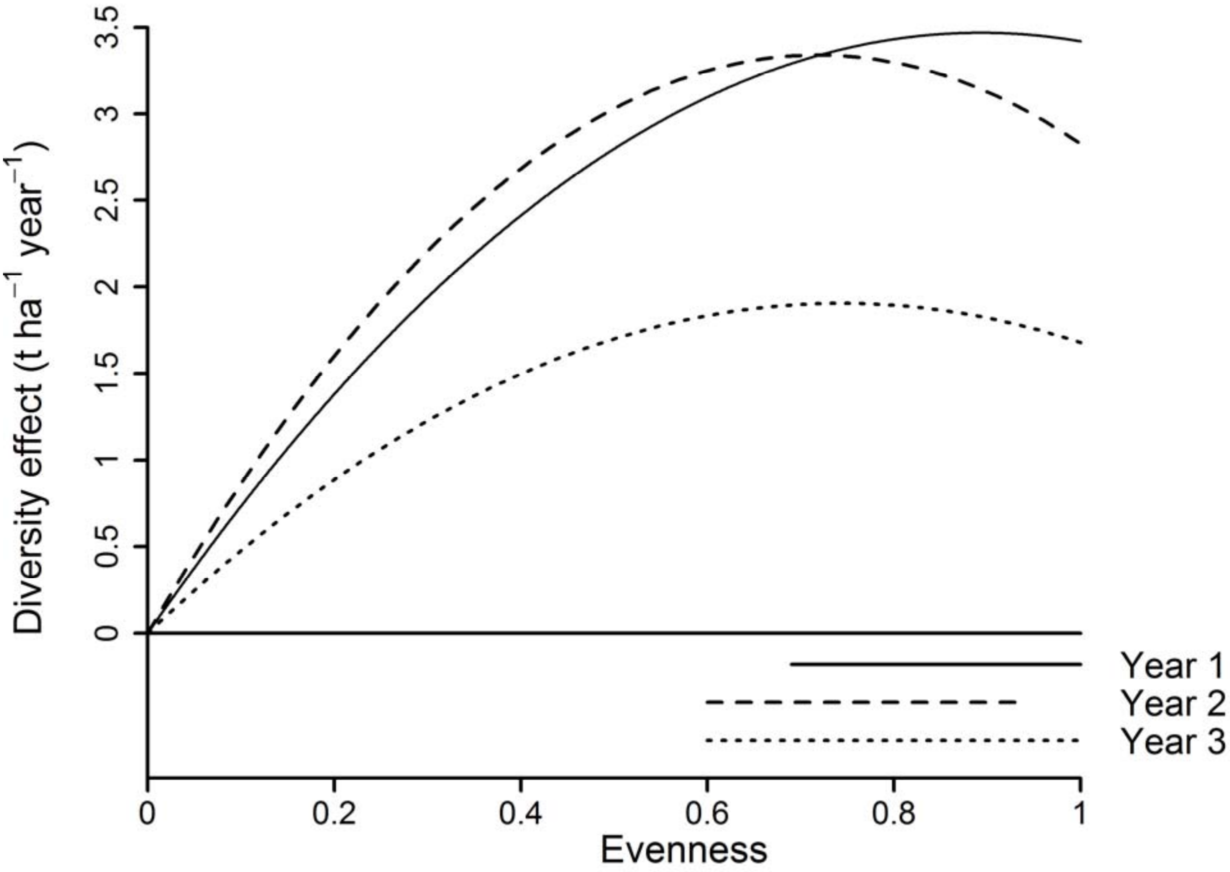

Figure S2. Predicted diversity effect (yield of sown species) from quadratic evenness model over three years and across all sites.

Evenness values in year 1 are based on sown proportions, whereas evenness values in years 2 and 3 are based on proportions of annual biomass in years 1 and 2, respectively. Horizontal lines below the regression curves indicate the range of evenness over which the diversity effect is not significantly ($p < 0.05$) smaller than the maximum diversity effect. (Based on data in Table S6.)