

## LETTERS

**Biodiversity and ecosystem multifunctionality**Andy Hector<sup>1</sup> & Robert Bagchi<sup>1†</sup>

**Biodiversity loss can affect ecosystem functions and services<sup>1–4</sup>. Individual ecosystem functions generally show a positive asymptotic relationship with increasing biodiversity, suggesting that some species are redundant<sup>5–8</sup>. However, ecosystems are managed and conserved for multiple functions, which may require greater biodiversity. Here we present an analysis of published data from grassland biodiversity experiments<sup>9–11</sup>, and show that ecosystem multifunctionality does require greater numbers of species. We analysed each ecosystem function alone to identify species with desirable effects. We then calculated the number of species with positive effects for all possible combinations of functions. Our results show appreciable differences in the sets of species influencing different ecosystem functions, with average proportional overlap of about 0.2 to 0.5. Consequently, as more ecosystem processes were included in our analysis, more species were found to affect overall functioning. Specifically, for all of the analysed experiments, there was a positive saturating relationship between the number of ecosystem processes considered and the number of species influencing overall functioning. We conclude that because different species often influence different functions, studies focusing on individual processes in isolation will underestimate levels of biodiversity required to maintain multifunctional ecosystems.**

The potential for diversity to affect ecosystem processes (or functions) was recognized by mid-twentieth-century researchers<sup>12</sup>, and originally dates back to Darwin<sup>13,14</sup>. The recent interest in this topic is due to the realization that sustained loss of biodiversity could affect ecosystem functioning and the services derived by humans<sup>1–4</sup>. Meta-analysis of the results of the first generation<sup>5</sup> of experimental research on biodiversity and ecosystem functioning has revealed that individual ecosystem processes generally show a positive but saturating relationship with increasing diversity, although the mechanisms underlying these relationships are still under debate<sup>5–8</sup>. The saturating relationship suggests that some species are redundant with respect to a single function. However, short-term studies only address the effect of biodiversity on ecosystem functioning at a given point in time and under a given set of conditions. Some experimental results<sup>15–19</sup> suggest that biodiversity can sometimes have an insurance value<sup>20</sup> by buffering ecosystem-level processes in a way analogous to that in which diverse investment portfolios spread financial risk and improve average performance in the longer term<sup>21–23</sup>. Nevertheless, all of the research to date considers ecosystem processes examined individually, despite the fact that most ecosystems are managed or valued for several ecosystem services or processes: so-called ecosystem ‘multifunctionality’<sup>24</sup>. If it is the case that a single species, or group of species, controls ecosystem functioning, then the remaining species are functionally redundant. Although it seems unlikely that a single species could control all ecosystem processes, it is possible that a single group of species may. However, if there is appreciable lack of overlap in the groups of species that influence different ecosystem processes, then higher levels of biodiversity will be required to maintain overall ecosystem functioning than indicated by analyses focusing on individual ecosystem processes in isolation.

To address the question of ecosystem multifunctionality, we used published data<sup>10,11</sup> on seven ecosystem processes and properties (for brevity, hereafter processes) measured by the BIODEPTH project. BIODEPTH comprised a consortium of eight co-ordinated biodiversity experiments that manipulated plant diversity at different European grassland sites and monitored the response of a variety of ecosystem processes<sup>9</sup>. The ecosystem variables used for this analysis were above- and below-ground net biomass production, the pool of nitrogen in above-ground vegetation, resource levels above- (light interception) and below-ground (soil mineral nitrogen) and decomposition of lignin and cellulose (see Methods).

In the first stage of this analysis, we identified sets of species that had effects on each ecosystem process, using an information theoretic approach. For each site, we performed a backward-deletion multiple regression that included contrasts for the presence/absence of each species, and used the Akaike Information Criterion (AIC) to identify the most parsimonious set of species that influenced each ecosystem process (our results were robust to the inclusion of pairwise interactions, the exclusion of data from monoculture plots, and the use of an alternative information criterion: see Methods and Supplementary Information). For each ecosystem process at each site, we then selected the subset of species with effects on processes that would usually be considered desirable from an ecosystem services perspective. These were species with positive effects for all functions, except for soil inorganic nitrogen and light availability at ground level, where negative effects are consistent with lower levels of unconsumed resources and therefore greater uptake and less potential for nitrate leaching. We then analysed the resulting species counts for differences between sites and processes.

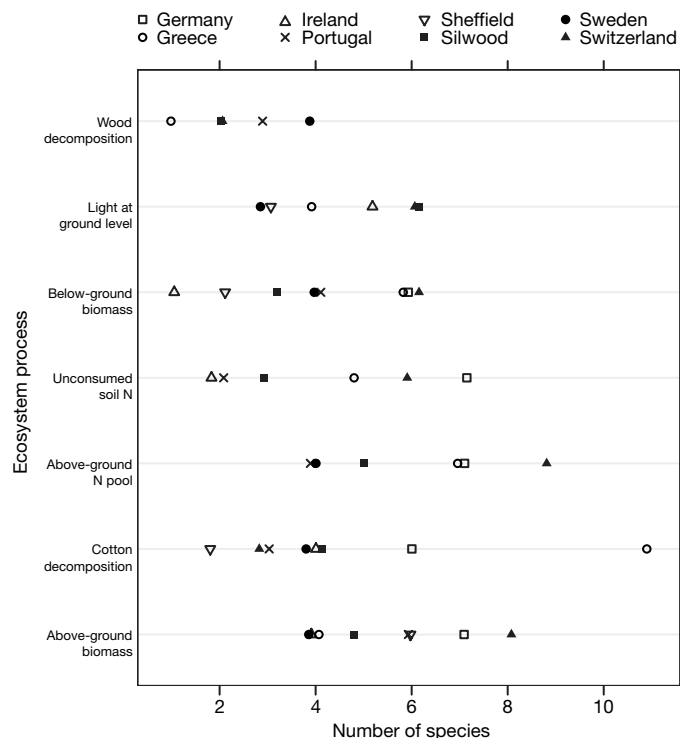
There were some differences between sites and processes in the number of species,  $x$ , that affected functioning (Fig. 1), although analysis of the results of the AIC multiple regressions revealed that the differences between processes were not strong enough to be conventionally significant (Poisson analysis of deviance:  $\chi^2_6 = 10.5$ ,  $P = 0.10$ ). The average number of species affecting a single ecosystem process,  $\bar{x}$ , ranged from 3.2 to 6.6 species depending on experimental site (Table 1), although these differences between experiments were also not conventionally significant ( $\chi^2_7 = 13.3$ ,  $P = 0.07$ ). The total number of species with effects on one or more ecosystem process at a site ranged from 8 to 18 (Table 1). This variation between experiments probably arises from a mixture of biological and design differences: they may partly reflect variation in the size of the species pools at the different sites, as those with the highest numbers of species with effects on functioning were also those with the largest species pools (see Methods and Supplementary Information).

We next looked at the overlap in these sets of species at each site that had AIC-detectable effects in the desirable direction on the different ecosystem processes. To do this we calculated the overlap,  $o$ , in the sets of species influencing each pair of processes, using Sørensen’s index:

$$o = \frac{|E_i \cap E_j|}{0.5(|E_i| + |E_j|)} \quad (1)$$

where  $|E_i|$  is the number of species contributing to process  $i$  and

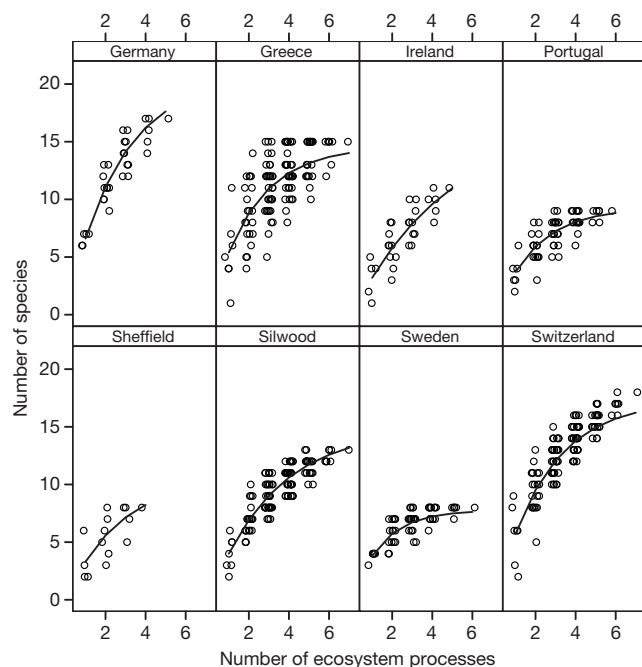
<sup>1</sup>Institute of Environmental Sciences, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland. †Present address: Department of Zoology, University of Oxford, Tinbergen Building, South Parks Road, Oxford OX1 3PS, UK.



**Figure 1** | Number of species with desirable effects on the suite of ecosystem processes measured in the different BIODEPTH project experiments. The number of species was identified by the AIC-based multiple regression (and species with effects with undesirable signs were then excluded).

$|E_i \cap E_j|$  is the number of species that contribute to both processes  $i$  and  $j$ . The average proportional overlap between the sets of species influencing a pair of processes,  $\bar{o}$ , ranged from 0.19 to 0.49 (Table 1) although these differences in  $\bar{o}$  between sites were not significant (quasi-maximum likelihood binomial analysis of deviance:  $F_{7,111} = 1.08, P = 0.38$ ). We identified the single most important species affecting each process from the multiple regression minimally adequate model estimates, and calculated a simple measure of turnover as the proportion of ecosystem processes with different most-important species (the number of processes divided by the number of different species; where the maximum possible value of 1 indicates that the identity of the most important species differs for each process). Turnover in the most important species was generally high, ranging from 0.67 to 1, with an average of 0.78 (Table 1).

In the second stage of the analysis, we took the multiple regression results and examined how the number of species affecting functioning changed as more processes were considered. We took the subsets of species that had detectable and desirable effects on the ecosystem process (or combination of processes) in question at each site, and calculated how many species affected functioning for every possible combination of ecosystem processes (that is, all pair-wise combinations, each combination of three, and so on). Figure 2 plots all possible combinations of



**Figure 2** | Positive relationship between the range of ecosystem processes considered and the number of species that affect one or more aspect of ecosystem functioning. The points (jittered for clarity) show numbers of species required for all possible combinations of ecosystem processes. Lines are theoretical predictions from the model based on the average number of species required for a single process,  $\bar{x}$ , and the average overlap in the sets of species required for each pair of processes,  $\bar{o}$ , using equation (2).

ecosystem processes, to explore the complete range of behaviour observed in our experiments. At all sites, the envelope of points forms a concave-up curved wedge, with the apex determined by the number of species required for all functions, and the shortest side of the wedge given by the range of values observed for single ecosystem processes (Table 1). We also predicted how many species on average affect ecosystem functioning as numbers of processes increase, based on only the mean number of species with effects on a single ecosystem process,  $\bar{x}$ , and the mean average pairwise overlap,  $\bar{o}$ . The average number of species,  $S_E$ , required for  $E$  processes was predicted as (see Methods for a derivation):

$$S_E = \sum_{i=1}^E \binom{E}{i} \bar{x}(-\bar{o})^{(i-1)} \quad (2)$$

Predictions of  $S_E$  for each site are plotted as curves in Fig. 2. These curves are the average predictions given the simplifying assumptions made by using  $\bar{x}$  and  $\bar{o}$ . The spread of the points around the curves reflects differences in  $x$  for different processes and the variation in  $o$  for different pairs of processes. Two null hypotheses illustrate some extreme cases within the range of possible behaviour. At one extreme, if each species influences only a single process ( $\bar{o} = 0$ ), and each process is only affected by a single species, it would generate a positive

**Table 1** | Summary of ecosystem multifunctionality analysis

Site	Number of processes	Number of species	Number of species for all processes	Turnover	Mean number of species per process, $\bar{x}$	Mean overlap between process pairs, $\bar{o}$
Germany	5	31 (24)	17	0.80	6.6 (4.6–9.1)	0.32 (0.19–0.47)
Greece	7	23 (19)	15	0.57	5.4 (3.9–7.3)	0.37 (0.26–0.48)
Ireland	5	12 (11)	11	1.00	3.2 (1.9–5.0)	0.19 (0.06–0.39)
Portugal	6	14 (11)	9	0.67	3.7 (2.3–5.4)	0.40 (0.25–0.57)
Sheffield	4	12 (12)	8	1.00	3.3 (1.8–5.3)	0.31 (0.10–0.59)
Silwood	7	34 (24)	13	0.86	4.0 (2.7–5.7)	0.24 (0.17–0.40)
Sweden	6	12 (12)	8	0.67	3.8 (2.5–5.6)	0.49 (0.33–0.65)
Switzerland	7	48 (33)	18	0.71	5.7 (4.1–7.7)	0.33 (0.23–0.44)

'Number of processes' is the number of ecosystem processes measured per site. 'Number of species' is the number of species originally sown; shown in parentheses is the number present in biomass samples in year three of the experiment. 'Number of species for all processes' is the total (cumulative) number of unique species with AIC-detectable effects on at least one ecosystem process. 'Turnover' is the proportion of ecosystem processes with different most important species, as explained in the main text. Definitions of  $\bar{x}$  and  $\bar{o}$  are given in the main text; numbers in parentheses are the 95% confidence intervals.

linear relationship with a slope of  $\bar{x} = 1$  (one species to one process). On the other hand, if all ecosystem processes are influenced by the same subset of species ( $\bar{\sigma} = 1$ ), the slope would be zero with the intercept indicating the size of the group of species that affects ecosystem functioning. All observed relationships lie between these two extremes, showing a positive but saturating relationship between increasing numbers of ecosystem processes and the total number of species that influence ecosystem functioning (Fig. 2; Table 1).

Our study of ecosystem multifunctionality has limitations that should be addressed by future research. For example, the types of ecosystem stocks and flows analysed by the first generation of research on biodiversity and ecosystem functioning can be seen in some ways as the more-accessible ones. It would be interesting to repeat a similar analysis for an ecosystem where attempts had been made to identify and measure all of the ecosystem services considered important from a management perspective, and where the effects of all species could be precisely estimated with a well-replicated balanced design. Our study also manipulates only plant diversity, and our analysis therefore almost certainly underestimates levels of diversity needed for ecosystem multifunctionality. Our results will reflect the effects of other groups of organisms on ecosystem processes to the degree that the manipulation of plant species had knock-on effects to other groups, such as bacteria and mycorrhiza, but our study does not directly address the functional role played by these and other important groups. Better assessments of ecosystem multifunctionality could be made by manipulating multiple trophic levels simultaneously, and especially by studies that include groups such as microbes that are known to drive many ecosystem processes.

To our knowledge, our study is the first to quantitatively address the effects of biodiversity on multiple ecosystem processes simultaneously. Analysis of single ecosystem processes in isolation generally reveals a positive but saturating relationship with increasing biodiversity. Our analysis of ecosystem multifunctionality reveals that different processes are not affected by exactly the same species. Because different species affect different processes, maintaining multifunctional ecosystems will require greater biodiversity than suggested by studies focusing on single ecosystem processes in isolation.

## METHODS SUMMARY

**Ecosystem processes.** The ecosystem process data used in our analysis are available online (<http://www.esapubs.org/archive/mono/M075/001/suppl-1.htm>) from Ecological Archives<sup>10,11</sup> (see Methods).

**Identifying species effects on ecosystem processes.** At each site, we identified the most parsimonious set of species influencing each ecosystem function using a backward-elimination multiple regression analysis to identify a minimally adequate model<sup>25</sup> based on AIC. We used the stepAIC function in the MASS library<sup>26</sup> implemented in R 2.4.1<sup>27</sup>. Models including pairwise interactions between species or excluding data from monocultures produced qualitatively similar results (Supplementary Information). We used an information-theoretic approach rather than one based on probability because information criteria have several advantages for the type of complex multiple regression analysis performed here<sup>28</sup> (Supplementary Information). Many alternative information criteria exist<sup>29,30</sup>. We present results for AIC, as it is reputed to have the most solid foundation in theory and the best analytical behaviour<sup>28</sup>, but analysis using the Bayesian Information Criteria (BIC) produced qualitatively similar results (Supplementary Information).

**Ecosystem multifunctionality analysis.** The mean number of species required per ecosystem process,  $\bar{x}$ , was estimated for each site using a Poisson generalized linear model analysing the main effects of site and process. The mean proportional overlap between the sets of species influencing each pair of processes,  $\bar{\sigma}$ , was estimated for each site using a quasi-maximum likelihood version of the binomial generalized linear model to account for under- or overdispersion<sup>25</sup>. The relationship between the number of processes and the average number of species with effects on functioning was then predicted using equation (2).

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

Received 13 April; accepted 15 May 2007.

1. Loreau, M. *et al.* Biodiversity and ecosystem functioning: current knowledge and future challenges. *Science* **294**, 804–809 (2001).

2. Kinzig, A., Tilman, D. & Pacala, S. (eds) *The Functional Consequences of Biodiversity: Empirical Progress and Theoretical Extensions* (Princeton Univ. Press, Princeton, 2002).
3. Loreau, M., Naeem, S. & Inchausti, P. (eds) *Biodiversity and Ecosystem Functioning: Synthesis and Perspectives* (Oxford Univ. Press, Oxford, 2002).
4. Hooper, D. U. *et al.* Effects of biodiversity on ecosystem functioning: a consensus of current knowledge and needs for future research. *Ecol. Monogr.* **75**, 3–36 (2005).
5. Srivastava, D. S. & Velland, M. Biodiversity-ecosystem function: is it relevant to conservation? *Annu. Rev. Ecol. Evol. System.* **36**, 267–294 (2005).
6. Balvanera, P. *et al.* Quantifying the evidence for biodiversity effects on ecosystem functioning and services. *Ecol. Lett.* **9**, 1146–1156 (2006).
7. Cardinale, B. J. *et al.* Effects of biodiversity on the functioning of trophic groups and ecosystems. *Nature* **443**, 989–992 (2006).
8. Worm, B. *et al.* Impacts of biodiversity loss on ocean ecosystem services. *Science* **314**, 787–790 (2006).
9. Hector, A. *et al.* Plant diversity and productivity experiments in European grasslands. *Science* **286**, 1123–1127 (1999).
10. Hector, A. *et al.* Ecosystem effects of the manipulation of plant diversity in European grasslands: data from the BIODEPTH project. *Ecol. Archives* M075–001, S1 (2005); (<http://esapubs.org/Archive/mono/M075/001/suppl-1.htm>).
11. Spehn, E. M. *et al.* Ecosystem effects of the manipulation of plant diversity in European grasslands. *Ecol. Monogr.* **75**, 37–63 (2005).
12. Carlander, K. D. The standing crop of fish in lakes. *J. Fish. Res. Bd Can.* **12**, 543–570 (1955).
13. Darwin, C. & Wallace, A. On the tendency of species to form varieties; and on the perpetuation of varieties and species by natural means of selection *Linn. Soc. J.* **3**, 45–62 (1858).
14. Hector, A. & Hooper, R. E. Darwin and the first ecological experiment. *Science* **295**, 639–640 (2002).
15. McCann, K. S. The diversity and stability of ecosystems. *Nature* **405**, 228–233 (2000).
16. Cottingham, K. L., Brown, B. L. & Lennon, J. T. Biodiversity may regulate the temporal variability of ecological systems. *Ecol. Lett.* **4**, 72–85 (2001).
17. Loreau, M. *et al.* in *Biodiversity and Ecosystem Functioning: Synthesis and Perspectives* (eds Loreau, M., Naeem, S. & Inchausti, P.) 79–91 (Oxford Univ. Press, Oxford, 2002).
18. Steiner, C. F., Long, Z. T., Krumins, J. A. & Morin, P. Population and community resilience in multitrophic communities. *Ecology* **87**, 996–1007 (2006).
19. Tilman, D., Reich, P. B. & Knops, J. M. H. Biodiversity and ecosystem stability in a decade-long grassland experiment. *Nature* **441**, 629–632 (2006).
20. Yachi, S. & Loreau, M. Biodiversity and ecosystem productivity in a fluctuating environment: The insurance hypothesis *Proc. Natl Acad. Sci. USA* **96**, 1463–1468 (1999).
21. Doak, D. F. *et al.* The statistical inevitability of stability-diversity relationships in community ecology. *Am. Nat.* **151**, 264–276 (1998).
22. Tilman, D., Lehman, C. L. & Bristow, C. E. Diversity-stability relationships: Statistical inevitability or ecological consequence? *Am. Nat.* **151**, 277–282 (1998).
23. Lehman, C. L. & Tilman, D. Biodiversity, stability, and productivity in competitive communities. *Am. Nat.* **156**, 534–552 (2000).
24. Sanderson, M. A. *et al.* Plant species diversity and management of temperate forage and grazing land ecosystems. *Crop Sci.* **44**, 1132–1144 (2004).
25. McCullagh, P. & Nelder, J. A. *Generalized Linear Models* (Chapman and Hall, London, 1989).
26. Venables, W. N. & Ripley, B. D. *Modern Applied Statistics with S* (Springer, Berlin, 2002).
27. R Development Core Team. *A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna, Austria, 2007).
28. Burnham, K. P. & Anderson, D. R. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach* (Springer, New York, 2002).
29. Ripley, B. D. in *Methods and Models in Statistics: In Honour of Professor John Nelder FRS* (eds Adams, N., Crowder, M., Hand, D. J. & Stephens, D.) 155–170 (Imperial College Press, London, 2004).
30. Taper, M. L. in *The Nature of Scientific Evidence* (eds Taper, M. L. & Lele, S. R.) 488–501 (Univ. Chicago Press, Chicago, 2005).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank the BIODEPTH project members for releasing this data on Ecological Archives, E. Bazeley-White and L. Wacker for data management, M. Weilenmann for help with the manuscript, L. Jackson and the participants of the DIVERSITAS-BioMERGE 2006 workshop for discussion; W. Venables for advice on stepAIC; and M. Loreau, B. Schmid and L. Turnbull for comments.

**Author Contributions** The two-stage analysis of multifunctionality was conceived by A.H., the mathematical formulation by R.B., and the analysis and writing was performed jointly.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to A.H. (ahector@uwinst.uzh.ch).

## METHODS

**The BIODEPTH experiments.** The BIODEPTH experiments were primarily designed to look at the influence of species and functional group richness on ecosystem processes. Random selection of species to form these gradients resulted in unequal replication for different species meaning that their effects are estimated with differing precision. The size of the species pool at each site reflected variation in background levels of diversity at the different grassland sites that provided the models for the experimental systems, and the highest level of diversity used in the species richness gradients was intended to approximate average levels of diversity in these surrounding natural grasslands<sup>31</sup>. Experiments at sites with larger pools tended to have a smaller proportion of species that were present in many plots and a larger proportion with lower replication because the experiments were of similar size regardless of the differences in the sizes of the species pool. The analyses presented in the main text use the whole diversity gradients, including monocultures, but the results are robust to their exclusion (see below).

**Ecosystem processes.** The data for the seven ecosystem processes used in our analysis are from year three of the BIODEPTH project, and comprise measurements of: (1) above-ground (shoot) biomass; (2) below-ground (root) biomass; (3) total nitrogen pools in above-ground vegetation; (4) soil mineral nitrogen (nitrate + ammonium); (5) percentage of transmitted photosynthetically active radiation (PAR) at ground level; (6) decomposition (per cent per day) of wooden sticks (lignin); and (7) decomposition (per cent per day) of cotton strips (cellulose).

**Identifying sets of species that affect ecosystem processes.** Before model fitting, the response variables were transformed as indicated by the Box-Cox function in the MASS library for S-Plus and R<sup>26</sup>. All ecosystem process variables (for brevity we use ecosystem processes to refer to both stocks and flows) were natural log transformed, except above-ground biomass (cube root) and soil inorganic N (untransformed). We performed single degree of freedom contrasts for the main effects of individual species at a given site within a backward-elimination multiple regression analysis, based on minimization of AIC, using the stepAIC function in the MASS library<sup>26</sup> implemented in R 2.4.1<sup>27</sup>. Starting with a full model that contained a main effect for the presence/absence of every species at a given site, each species was removed from the model in turn and the AIC calculated. The AIC values of the resulting models were compared, and the species whose exclusion led to the greatest improvement (reduction) in the AIC value was permanently excluded. This was repeated until dropping any of the remaining species increased the AIC, at which point a minimally adequate model<sup>25</sup> had been selected based on AIC which identified the most parsimonious set of species influencing the ecosystem process in question. Models that included pairwise interactions between species or which excluded data from monocultures produced qualitatively similar general results (Supplementary Information). The AIC is defined as:

$$AIC = -2\ln(L) + 2k$$

where  $L$  is the model maximum likelihood (and  $-2\ln(L)$  the deviance) and  $k$  is the number of model parameters. Schwartz's Bayesian Information Criterion (BIC), which has a harsher penalty for complexity and therefore tends to select simpler models, is defined as:

$$BIC = -2\ln(L) + \ln(n)k$$

where  $n$  is the sample size. We chose to repeat the analysis with the BIC because it is probably the most widely used alternative to AIC, and because the BIC and AIC come from different classes of information criteria (the dimensional-consistent and efficient classes, respectively).

**Ecosystem multifunctionality analysis.** The mean number of species required per ecosystem process,  $\bar{x}$  (with 95% confidence intervals), was estimated for each site using a Poisson generalized linear model analysing the main effects of site and process. The mean proportional overlap between the sets of species influencing each pair of processes,  $\bar{o}$  (and 95% confidence intervals), was estimated for each site using a quasi-maximum-likelihood version of the binomial generalized linear model to account for under- or overdispersion<sup>25</sup>. Once  $\bar{x}$  and  $\bar{o}$  had been estimated, the relationship between the number of processes and the average number of species with effects on functioning was predicted using equation (2).

**Predicting the number of species required for a given number of ecosystem processes.** We estimated the expected number of species required for a given number of ecosystem processes using set theory. Given  $E$  sets (each comprising the species providing a given ecosystem process), the number of elements of the union of the sets can be defined according to the inclusion-exclusion principle<sup>32</sup> as:

$$|A_1 \cup A_2 \cup \dots \cup A_E| = \sum_{1 \leq i \leq E} |A_i| - \sum_{1 \leq i_1 < i_2 \leq E} |A_{i_1} \cap A_{i_2}| + \sum_{1 \leq i_1 < i_2 < i_3 \leq E} |A_{i_1} \cap A_{i_2} \cap A_{i_3}| - \dots + (-1)^{(E-1)} |A_1 \cap A_2 \cap \dots \cap A_E| \quad (3)$$

We simplified equation (3) by making two assumptions. First, that all processes were controlled by the same number of species (that is,  $|A_i| = x$ , for all  $i$  where  $x$  is a constant). Second, that the overlap between the species contributing to pairs of processes,  $o$ , was also the same for all species pairs (that is,  $|A_i \cap A_j| = o$ , for all  $i$  and  $j$ , where  $o$  is a constant). Under these assumptions, and setting  $S_E = |A_1 \cup A_2 \cup \dots \cup A_E|$ , equation (3) simplifies to:

$$S_E = \binom{E}{1}x - \binom{E}{2}x \cdot o + \binom{E}{3}x \cdot o^2 - \dots + (-1)^{(E-1)} \binom{E}{E}x \cdot o^{(E-1)} \quad (4)$$

For each site, we estimated the constants  $x$  and  $o$  from the data as the mean number of species per ecosystem process ( $\bar{x}$ ) and the mean proportion of species shared between pairs of processes ( $\bar{o}$ ), respectively. Simplifying equation (4) and replacing  $x$  and  $o$  with  $\bar{x}$  and  $\bar{o}$  results in equation (2) in the main text. Note that equation (2) in the main text can be simplified algebraically to:

$$S_E = \frac{\bar{x} \cdot (1 - (1 - \bar{o})^E)}{\bar{o}} \quad (5)$$

This form can be used to infer the number of species that would theoretically provide all ecosystem processes by setting  $E$  to  $\infty$ :

$$S_{\max} = \lim_{E \rightarrow \infty} \frac{\bar{x} \cdot (1 - (1 - \bar{o})^E)}{\bar{o}} = \frac{\bar{x}}{\bar{o}} \quad (6)$$

31. Hector, A. *et al.* Biodiversity and ecosystem functioning: reconciling the results of experimental and observational studies. *Funct. Ecol.* (in the press).

32. Comtet, L. *Advanced Combinatorics: The Art of Finite and Infinite Expansions* 176–177 (Reidel, Dordrecht, 1974).