

Frequency-dependent feedback constrains plant community coexistence

Maarten B. Eppinga ^{1*}, Mara Baudena ¹, Daniel J. Johnson ², Jiang Jiang³, Keenan M. L. Mack⁴, Allan E. Strand⁵ and James D. Bever⁶

Ecological theory suggests that coexistence of many species within communities requires negative frequency-dependent feedbacks to prevent exclusion of the least fit species. For plant communities, empirical evidence of negative frequency dependence driving species coexistence and diversity patterns is rapidly accumulating, but connecting these findings to theory has been difficult as corresponding theoretical frameworks only consider small numbers of species. Here, we show how frequency-dependent feedback constrains community coexistence, regardless of the number of species and inherent fitness inequalities between them. Any interaction network can be characterized by a single community interaction coefficient, I_c , which determines whether community-level feedback is positive or negative. Negative feedback is a necessary (but not sufficient) condition for persistence of the entire community. Even in cases where the coexistence equilibrium state cannot recover from perturbations, $I_c < 0$ can enable species persistence via cyclic succession. The number of coexisting species is predicted to increase with the average strength of negative feedback. This prediction is supported by patterns of tree species diversity in more than 200,000 deciduous forest plots in the eastern United States, which can be reproduced in simulations that span the observed range of community feedback. By providing a quantitative metric for the strength of negative feedback needed for coexistence, we can now integrate theory and empirical data to test whether observed feedback–diversity correlations are strong enough to infer causality.

Despite decades of research, it remains surprisingly challenging to explain how species competing for the same resources can coexist in the large numbers observed in plant communities^{1–7}. Ecological theory states that stabilizing mechanisms creating negative frequency-dependent feedbacks are needed to prevent competitive exclusion due to species differences in fitness^{6–8}. Recent empirical studies of species-rich plant communities are in qualitative agreement with the expected connection between negative frequency-dependent feedback and patterns of species abundance^{9,10} and species richness^{11–13}. Until now, utilization of these data for more direct tests of ecological theory is hampered by the lack of a theoretical analytical framework linking species interactions to community diversity and stability. In contrast with studies of evolutionary game dynamics¹⁴, theoretical ecological studies have focused either on coexistence of small numbers of species³ or on sufficiently large communities, enabling a distributional approach based on probabilities (of particular species interactions¹⁵). Alternatively, coexistence has been studied for equivalent species¹⁶, species exerting equal effects on all competitors¹⁷, (anti-)symmetric species interactions^{5,7} and species experiencing strictly hierarchical competition². These assumptions enable a detailed description of potential plant community dynamics, but do not consider asymmetric competitive interactions. In addition, the above approaches do not represent the forces driving negative frequency-dependent feedbacks at the community level.

Focusing on a specific mechanism through which frequency-dependent feedbacks may arise from asymmetric species interactions, plant–soil feedback theory has been developed by means of

models that can be directly parameterized with data from pot or field experiments, providing a promising means to integrate experiments with theory^{18–20}. Yet, theoretical development and empirical testing of plant–soil feedback theory has been limited to small numbers of interacting species^{18–23}. Here, we generalize the plant–soil feedback model¹⁸ (see Supplementary Information Section 1), yielding a generic phenomenological model of frequency-dependent interactions. When rewriting the species' density dynamics into frequency dynamics, this generalized model projects onto the well-known replicator equation originating from evolutionary game theory¹⁴. The replicator equation can be solved analytically for communities containing any number of species. This procedure does not require an assumption of species equivalence, meaning that equilibrium solutions retain species-specific parameters describing the growth for each species. We also derive a fundamental criterion for species persistence. Our results provide a framework to assess the importance of empirically observed frequency-dependent interactions in maintaining coexistence of highly diverse plant communities (Fig. 1).

Results and discussion

The proposed framework consists of a set of ordinary differential equations describing plant frequencies and hence community dynamics over time (Supplementary Information Section 1). The only parameters in this model are the coefficients σ_{ij} , which quantify the growth rate of species i in an environment dominated by species j (Fig. 1). Thus, when these coefficients are set to non-neutral values, species fitness becomes frequency dependent, creating a

¹Copernicus Institute of Sustainable Development, Environmental Sciences, Faculty of Geosciences, Utrecht University, Utrecht, the Netherlands.

²Department of Earth and Environmental Sciences, Los Alamos National Laboratory, Los Alamos, NM, USA. ³Department of Soil and Water Conservation, Nanjing Forestry University, Nanjing, China. ⁴Department of Biology, Illinois College, Jacksonville, IL, USA. ⁵Department of Biology, Grice Marine Laboratory, College of Charleston, Charleston, SC, USA. ⁶Department of Ecology and Evolutionary Biology and Kansas Biological Survey, University of Kansas, Lawrence, KS, USA. *e-mail: M.B.Eppinga@uu.nl

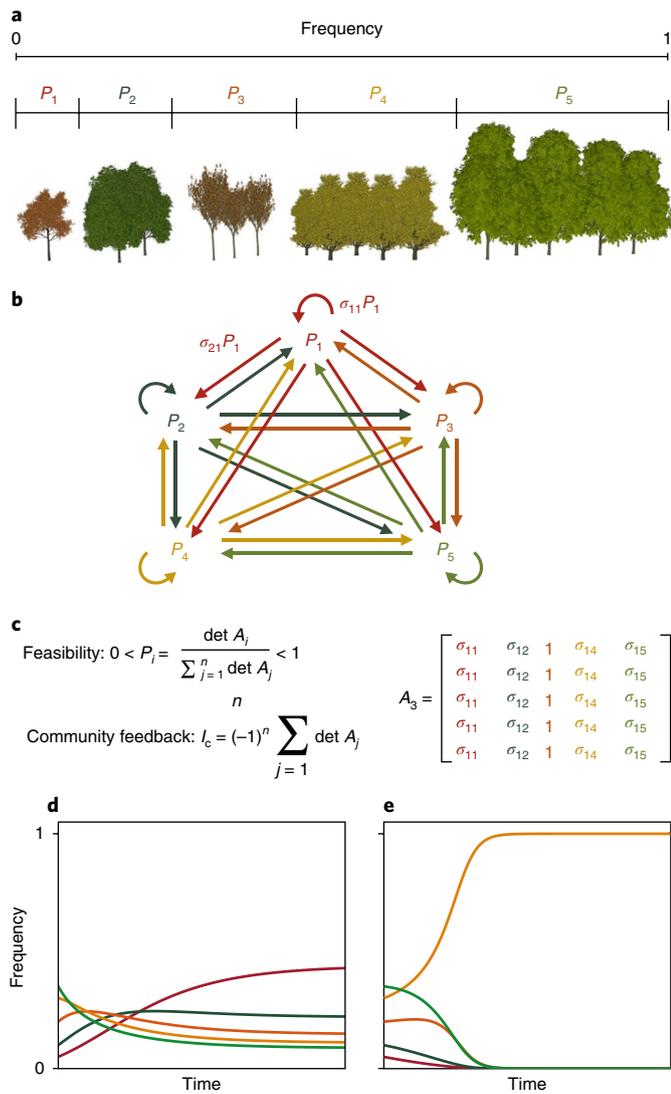


Fig. 1 | Conceptual representation of how the frequency-dependent feedback framework can be extended to the community level for any number of plant species (here, $n = 5$). **a**, Plant species (P_i) are modelled as frequencies within a community. The diagram includes trees from Laubwerk Plant Kit 1 (www.laubwerk.com). **b**, Plant species exert conspecific (σ_{ii}) and heterospecific (σ_{ij}) effects, the magnitude of which (per capita) depends on the frequency of the exerting species. **c**, The resulting interaction network determines whether the community coexistence equilibrium is feasible; equilibrium frequencies are calculated by reduced determinants of the interaction matrix, A_i , the sum of which yields the community interaction coefficient, I_c . **d**, $I_c < 0$ generates negative community-level feedback and is a necessary condition for persistence of all plant species within the community. **e**, $I_c > 0$ generates positive community-level feedback and leads to competitive exclusion.

feedback between community composition and the growth of each species. For a system of two species, net feedback is described by the two-species interaction coefficient I_s ^{18,19}:

$$I_s = \sigma_{11} - \sigma_{12} - \sigma_{21} + \sigma_{22} \quad (1)$$

Provided that the coexistence equilibrium is feasible (that is, both plant species have a positive frequency at equilibrium), $I_s < 0$ will generate negative feedback, and the equilibrium is stable.

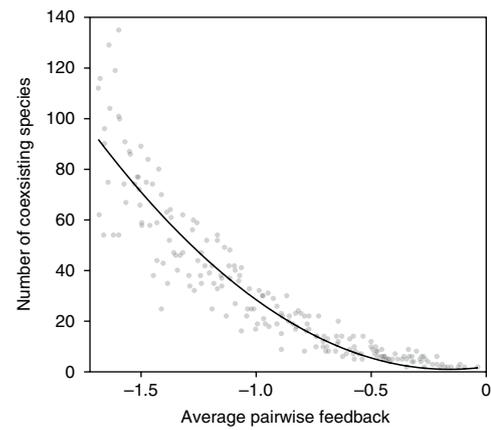


Fig. 2 | Two hundred runs of plant community assembly show that more species can stably coexist under stronger negative feedback. The solid line shows a quadratic fit assuming a minimum community size of two plant species ($y = 39.00x^2 + 12.52x$; $R_{adj}^2 = 0.86$).

Alternatively, $I_s > 0$ will generate positive feedback and the equilibrium is unstable, meaning that one of the plant species will be excluded^{18,19} (Supplementary Information Section 2). I_s can thus be interpreted as a pairwise feedback: the sum of two species-specific feedbacks, which are described by the difference between a species' conspecific and its heterospecific responses (thus, the species-specific feedback of plant species 1, for example, is given by²⁴: $\sigma_{11} - \sigma_{21}$). For the two-species case, the equilibrium frequency for each plant species is then given by the ratio between its competitor's species-specific feedback and the pairwise feedback¹⁸, yielding:

$$\hat{P}_1 = \frac{\sigma_{22} - \sigma_{12}}{I_s}, \quad \hat{P}_2 = \frac{\sigma_{11} - \sigma_{21}}{I_s} \quad (2)$$

in which \hat{P}_i indicates the species' equilibrium frequencies. These results can be generalized from two to any number of species (Supplementary Information Section 3). General results can be obtained using the interaction matrix:

$$A = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1n} \\ \sigma_{21} & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \sigma_{n1} & \dots & \dots & \sigma_{nn} \end{bmatrix} \quad (3)$$

As in the two-species case, the frequency of a given species at the coexistence equilibrium is the ratio between the feedback effects that this species experiences from all competitors and the sum of all feedback effects within the community:

$$\hat{P}_i = \frac{\det A_i}{\sum_{j=1}^n \det A_j} \quad (4)$$

where A_i indicates the interaction matrix A , but with the i^{th} column being replaced with a column vector of ones (Fig. 1). Thus, based on the interaction structure within the plant community (equation (3)), the equilibrium frequencies for species at coexistence can be calculated (equation (4)). Stability of this equilibrium can then be assessed by standard eigenvalue analysis of the system's Jacobian matrix. We used this approach to study community assembly and the resulting plant species richness (quantified as the maximum number of species that were found to stably coexist) under various regimes of negative feedback (Fig. 2; see also Methods

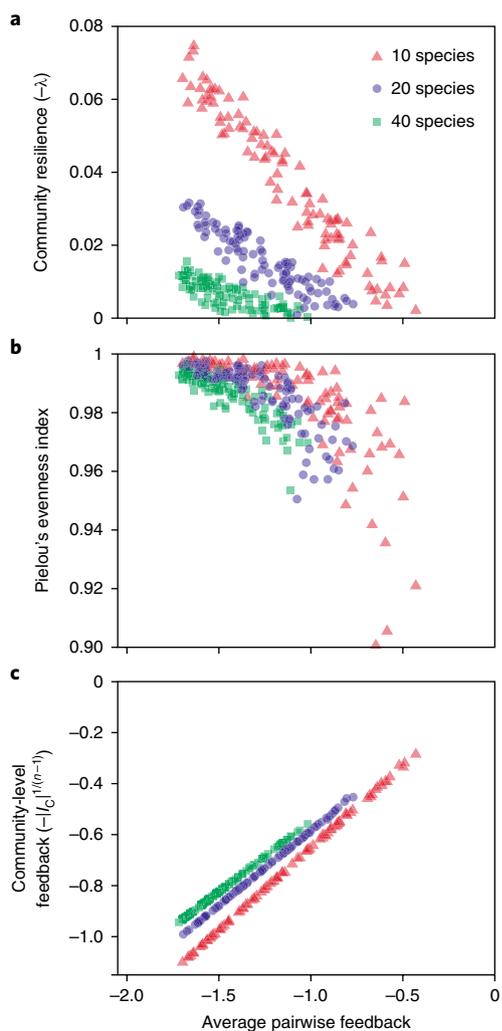


Fig. 3 | Effects of frequency-dependent feedback on plant community structure. **a**, Increasingly negative feedback increases the resilience of the coexistence equilibrium (measured as the absolute real part of the leading (and negative) eigenvalue²⁷). **b**, Increasingly negative feedback increases species evenness within the stably coexisting community measured with Pielou's evenness index (that is, the ratio between the actual and the maximum value of the Shannon–Wiener index). **c**, Even though community-level feedback is determined by an analytical expression of the order σ_{ij}^{n-1} , the strength of average pairwise feedback (that is, of the order σ_{ij}^1) correlates well with the (rescaled) strength of this community-level feedback.

and Supplementary Information Section 6). We found that under more negative feedback, indicated by the system's average pairwise feedback^{10,19,23}, communities with higher species richness could be assembled (Fig. 2). For a plant community in which the number of species was fixed, more negative feedback increased the resilience of the coexistence equilibrium (Fig. 3a, in which resilience is measured as the return time after perturbation²⁵). In addition, more negative feedback increased community evenness (Fig. 3b). The fact that average pairwise feedback correlates well with these community-level properties is surprising because pairwise descriptors provide no formal (that is, mathematically formulated) conditions for coexistence in communities of more than two species⁷. Therefore, to establish whether pairwise descriptors can be used as an indicator of community-level feedback, a formal measure of community-level feedback is also needed.

We found that for the model framework analysed, the two-species interaction coefficient I_S can be generalized to the community level, yielding an exact measure for community-level feedback. More specifically, the generalized community interaction coefficient, I_C , is given by (Supplementary Information Section 3):

$$I_C = (-1)^n \sum_{j=1}^n \det A_j \quad (5)$$

A pattern generally observed in empirical studies is that conspecific effects (and variation therein) are relatively strong compared with heterospecific effects (and variation therein)^{11–13}. Under these conditions, our measure of community-level feedback correlated well with the average strength of pairwise feedback (equation (1); Fig. 3c and Supplementary Information Section 6). This result did not only hold for asymmetric interaction matrices (Fig. 3c), but also for (anti)-symmetric interaction matrices (Supplementary Information Section 3). This monotonic relation between average pairwise feedback and community-level feedback thus provides an explanation for recent empirical observations of average pairwise feedback correlating well with community diversity^{11–13}.

We could prove that negative community-level feedback, $I_C < 0$, presents a necessary condition for persistence of all species (Supplementary Information Section 3). Interestingly, the value of I_C is not affected by inherent fitness inequalities⁶ between species (Supplementary Information Section 7). Following the community assembly approach (see Methods), we found that negative feedback could still be present in communities for which the real part of the leading eigenvalue of the coexistence equilibrium was positive (Supplementary Information Section 4). Thus, even when small perturbations would drive the community away from the coexistence equilibrium point, negative community-level feedback may still prevent species extinction. In this case, species abundances fluctuate, but dominance is passed on from one species to another in ways that create repeating loops (for example, $A \rightarrow B \rightarrow C \rightarrow A$), referred to as cyclic succession^{22,26} (Supplementary Information Section 4). However, $I_C < 0$ is not a sufficient condition for persistence of all species within a community (Supplementary Information Section 3). These observations highlight the importance of knowing the exact structure of the right-hand side of equation (3), as the loops creating the cyclic succession dynamics are not conserved when all the heterospecific effects that a species experiences have been averaged (for example, Supplementary Information Section 4).

Parameterizing the right-hand side of equation (3) requires full-factorial experimental designs, the size of which increases rapidly with increasing plant community size. However, even in cases where information is incomplete, our theoretical framework can be used to infer plant community characteristics by constraining the right-hand side of equation (3) within a realistic range of parameter values (rather than providing the exact value of each parameter). Observational data, as recently gathered in diverse forest ecosystems^{9,11–13}, can be used for this aim. To illustrate this potential, we analysed data from North American deciduous forests east of the hundredth meridian, aggregating surveys from 207,444 plots to $2^\circ \times 2^\circ$ grid cells¹¹ (Fig. 4a). Previous measures of conspecific and averaged heterospecific frequency dependence¹¹ can be translated into the model parameters used in the current framework (equations (3) and (4), and Supplementary Information Section 5). Focusing on grid cells containing sufficient data for five or more tree species¹¹ (Supplementary Information Section 5), we distinguished four community diversity classes (5–15, 15–25, 25–35 and ≥ 35 species occurring in a grid cell; Fig. 4a). We used the previously derived measures for conspecific and averaged heterospecific density dependence¹¹ to calculate for each of these cells the average pairwise feedback experienced by the tree species present.

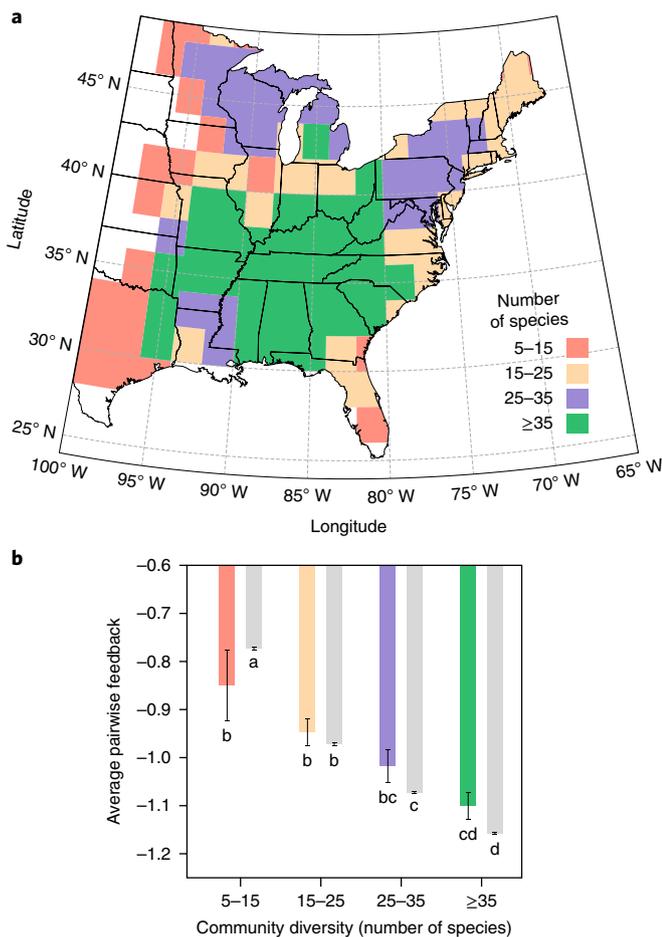


Fig. 4 | Application of the theoretical framework reveals that observed patterns of negative frequency dependence in over 200,000 forest plots in the eastern United States are consistent with predictions from theoretical communities in which community-level feedback determines species richness. a, Map showing the locations of the $2^\circ \times 2^\circ$ grid cells and the community diversity observed in each cell (divided into four diversity classes: 5–15 species ($n=24$), 15–25 species ($n=36$), 25–35 species ($n=22$) and ≥ 35 species ($n=31$)). **b**, Grid cells with higher community diversity are characterized by stronger negative feedback (coloured bars; Kruskal-Wallis H test, $\chi^2_{3,109} = 15.34$, $P = 0.0015$), the variation in average pairwise feedback strength being consistent with the range occurring in theoretical communities (grey bars; $n_{5-15} = 1,059$; $n_{15-25} = 426$; $n_{25-35} = 272$; $n_{\geq 35} = 355$). Bars indicate means ± 1 s.e.m., and different letters indicate differences at the $P = 0.05$ significance level (Kruskal-Wallis H test, $\chi^2_{7,2217} = 1,574$, $P < 1.10^{-323}$, followed by Dunn-Sidak post hoc tests).

Consistent with previous observations of feedback–diversity correlations^{11,12}, the average pairwise feedback became more strongly negative with increasing community diversity (see the coloured bars in Fig. 4b). A hitherto unanswered question, however, is whether these observed differences in feedback are indeed strong enough to drive (rather than correlate with) the variation in community diversity. Therefore, we compared the pattern obtained from these observations with independently assembled theoretical communities that spanned the same range of diversity classes. In these theoretical communities, the differences in species diversity were entirely driven by variation in the strength of community-level feedback. The conspecific and heterospecific coefficients in these theoretical communities were varied over the same range as observed in the field data (Supplementary Information Section 5). The community

assembly process was carried out as before (Fig. 2; see Methods). We found that the variation in pairwise feedback strength between these theoretical communities (grey bars in Fig. 4b) was similar to the variation empirically observed (Fig. 4b). Although a correlation between feedback and diversity has been previously observed in northern American deciduous forests^{11,12}, application of the framework developed here verifies that the observed variation in feedback is in theory strong enough to be a causal factor driving the observed variation in community diversity (Fig. 4).

Our results further suggest that in the absence of inherent fitness differences between species, a species' experienced feedback will positively correlate with its relative abundance (equation (4)), as has been commonly observed in the field^{10,11,13,27}. Although inherent fitness differences may decouple this link between experienced feedback and relative abundance, we found that such differences would not undermine the stabilizing force of community-level negative feedback (Supplementary Information Section 7). Together, these results provide an explanation for recent observations of consistent signatures of negative feedback in tropical and temperate forest communities but declining slopes of feedback–abundance relationships with increasing latitude¹³.

Quantitative studies of ecological networks have mainly focused on food webs and mutualistic networks^{15,28}, or have focused on competition within a trophic level by studying networks that are symmetric (for example, mimicking overlap in resource use⁷) or anti-symmetric (where there is a winner and a loser in each two-way interaction³; see also ref. ³). Frequency-dependent feedback, as studied here, considers a community in which plants interact through a fully connected network but where interactions between species can be asymmetric (for example, species A strongly affects species B, but B only weakly affects A). Empirical evidence for specific mechanisms creating this kind of interaction structure is rapidly increasing. Examples include density-dependent seedling mortality^{9,10}, plant–soil feedbacks initiated by invertebrate fauna²⁹, soil pathogens²¹ or mycorrhizal associations²⁷, and autotoxic effects of extracellular self-DNA originating from plant litter³⁰. Understanding the contribution of these mechanisms to plant community coexistence will be contingent on moving beyond the study of interactions between small numbers of species^{18–20}. The theoretical framework presented here provides a crucial link between empirical estimates of frequency-dependent feedback and its consequences for community stability and species diversity patterns.

Methods

Strength of negative frequency-dependent feedback and species richness of communities. A random interaction matrix for 1,000 species was generated. From this matrix, communities were assembled, starting with two randomly selected species. Analytical calculation of the coexistence equilibrium (see equations (3) and (4)) and the relevant eigenvalues of the corresponding Jacobian matrix (Supplementary Information Section 3) were then calculated to assess whether this two-species coexistence equilibrium was stable. If coexistence was feasible and stable, another species was randomly selected and added to the community, after which the procedure was repeated. This process continued until the community was no longer locally stable, or until a predetermined species number was reached, at which point the community resilience (by means of the magnitude of the relevant eigenvalue with the maximal real part²⁵) and evenness (using Pielou's index) were assessed. Average feedback strengths in the 1,000-species matrix were varied by manipulating the strengths of conspecific effects relative to the strengths of heterospecific effects. In the simulations, conspecific coefficients ranged between 0.05 and 0.3, and heterospecific coefficients ranged between 0.3 and 0.9 (refs 10,23). The subsets considered within these ranges varied along the feedback gradient (see Supplementary Information Section 6 for details). However, we verified that alternative parameterizations led to qualitatively similar feedback–community diversity relations (Supplementary Information Section 6).

Plant–soil feedbacks and the Forest Inventory and Analysis dataset. We analysed tree species diversity of forest plots located east of the hundredth meridian, which are part of the US Forest Service's Forest Inventory and Analysis (FIA) dataset. These data were previously extracted and preprocessed¹⁴. In this previous work, the strength of density dependence was inferred from (maximum likelihood) fitting

a negative exponential relation between observed seedling densities and tree density in the surrounding plot. Here, we used the dataset described in ref.³¹ to obtain exponential coefficients, using cumulative basal area as the predictor variable for adult density. The dataset analysed in ref.³¹ differed from ref.¹¹ in that it also included ecologically relevant joint absences of seedlings and adult trees within forest plots, which tends to increase coefficient values³². The exponential coefficients were used to calculate values for the conspecific frequency-dependent effect parameters, σ_{ij} , and for the average heterospecific effect, σ_{ij} , on each species (Supplementary Information Section 5). Theoretical results showed that average pairwise feedback (equation (1)) correlates well with community-level feedback (see Fig. 3c and Supplementary Information Sections 3 and 6). Average pairwise feedback could also be calculated for the FIA data selection, as it only requires information on the average of heterospecific effects that a tree species experiences¹¹ (Supplementary Information Section 5). We calculated average pairwise feedback strengths aggregating observations from different plots at the level of grid cells spanning $2^\circ \times 2^\circ$ (latitude \times longitude)¹¹ (Supplementary Information Section 5). In total, average pairwise feedback was calculated for 113 grid cells in which the average pairwise feedback could be calculated for five or more tree species (up to 51 tree species). These calculated feedbacks were compared with a null prediction in which the associations between seedlings and adults within each grid cell were randomized (Supplementary Information Section 5). Patterns of feedback strength and observed community diversity in the corresponding grid cells were then compared with patterns in theoretical communities in which community diversity was determined by the strength of community-level feedback. As the data provided row-averaged values of heterospecific effects (that is, the off-diagonal entries of the community matrix; see equation (3)), the model parameters in simulations could be constrained; that is, varied over the same range as observed in the data, including the ranges assumed for within-row and between-row variation in heterospecific effects (Supplementary Information Section 5).

Reporting Summary. Further information on experimental design is available in the Nature Research Reporting Summary linked to this article.

Code availability. The MATLAB and R scripts written for this study only contain standard functions and analysis methods. Scripts are available from the corresponding author upon reasonable request.

Data availability. The US Forest Service's FIA data that support the findings of this study are available from www.fia.fs.fed.us/tools-data/ under FIA Data Mart. The database subset that was used in the current study is available from the corresponding author upon reasonable request.

Received: 2 February 2018; Accepted: 27 June 2018;
Published online: 30 July 2018

References

- MacArthur, R. & Levins, R. The limiting similarity, convergence and divergence of coexisting species. *Am. Nat.* **101**, 377–385 (1967).
- Tilman, D. Competition and biodiversity in spatially structured habitats. *Ecology* **75**, 2–16 (1994).
- Tilman, D. *Resource Competition and Community Structure* (Princeton Univ. Press, Princeton, 1982).
- Clark, J. S. Individuals and the variation needed for high species diversity in forest trees. *Science* **327**, 1129–1132 (2010).
- Allesina, S. & Levine, J. A competitive network theory for species diversity. *Proc. Natl Acad. Sci. USA* **108**, 5638–5642 (2011).
- Chesson, P. Mechanisms of maintenance of species diversity. *Ann. Rev. Ecol. Evol. Syst.* **31**, 343–366 (2000).
- Barabás, G., Michalska-Smith, M. J. & Allesina, S. The effect of intra- and interspecific competition on coexistence in multispecies communities. *Am. Nat.* **188**, E1–E12 (2016).
- Adler, P., HilleRisLambers, J. & Levine, J. A niche for neutrality. *Ecol. Lett.* **10**, 95–104 (2007).
- Comita, L. S., Muller-Landau, H. C., Aguilar, S. & Hubbell, S. P. Asymmetric density dependence shapes species abundances in a tropical tree community. *Science* **329**, 330–332 (2010).
- Mangan, S. A. et al. Negative plant–soil feedback predicts tree-species relative abundance in a tropical forest. *Nature* **466**, 752–756 (2010).
- Johnson, D. J., Beaulieu, W. T., Bever, J. D. & Clay, K. Conspecific negative density dependence and forest diversity. *Science* **336**, 904–907 (2012).
- LaManna, J. A. et al. Negative density dependence mediates biodiversity–productivity relationships across scales. *Nat. Ecol. Evol.* **1**, 1107–1115 (2017).
- LaManna, J. A. et al. Plant diversity increases with the strength of negative density dependence at the global scale. *Science* **356**, 1389–1392 (2017).
- Hofbauer, J. & Sigmund, K. *Evolutionary Games and Population Dynamics* (Cambridge Univ. Press, Cambridge, 1998).
- Allesina, S. & Tang, S. Stability criteria for complex systems. *Nature* **483**, 205–208 (2012).
- Bascompte, J., Jordano, P. & Olesen, J. M. Asymmetric coevolutionary networks facilitate biodiversity maintenance. *Science* **312**, 431–433 (2006).
- Chisholm, R. A. & Muller-Landau, H. A theoretical model linking interspecific variation in density dependence to species abundances. *Theor. Ecol.* **4**, 241–253 (2010).
- Bever, J. D., Kristi, M. W. & Antonovics, J. Incorporating the soil community into plant population dynamics: the utility of the feedback approach. *J. Ecol.* **85**, 561–573 (1997).
- Bever, J. D. Soil community feedback and the coexistence of competitors: conceptual frameworks and empirical tests. *New Phytol.* **157**, 465–473 (2003).
- Kulmatiski, A., Heavilin, J. & Beard, K. H. Testing predictions of a three-species plant–soil feedback model. *J. Ecol.* **99**, 542–550 (2011).
- Van der Putten, W. H., Van Dijk, C. & Peters, B. A. M. Plant-specific soil-borne diseases contribute to succession in foredune vegetation. *Nature* **362**, 53–56 (1993).
- Revilla, T., Veen, G. F., Eppinga, M. B. & Weissing, F. J. Plant–soil feedbacks and the coexistence of competing plants. *Theor. Ecol.* **6**, 99–113 (2013).
- Mack, K. M. L. & Bever, J. D. Coexistence and relative abundance in plant communities are determined by feedbacks when the scale of feedback and dispersal is local. *J. Ecol.* **102**, 1195–1201 (2014).
- Levins, R. The qualitative analysis of partially specified systems. *Ann. NY Acad. Sci.* **231**, 123–138 (1974).
- Wissel, C. A universal law of the characteristic return time near thresholds. *Oecologia* **65**, 101–107 (1984).
- Bennincà, E., Ballantine, B., Ellner, S. P. & Huisman, J. Species fluctuations sustained by cyclic succession at the edge of chaos. *Proc. Natl Acad. Sci. USA* **112**, 6389–6394 (2015).
- Klironomos, J. N. Feedback with soil biota contributes to plant rarity and invasiveness in communities. *Nature* **417**, 67–70 (2002).
- Neutel, A. M. et al. Reconciling complexity with stability in naturally assembling food webs. *Nature* **449**, 599–602 (2007).
- De Deyn, G. B. et al. Soil invertebrate fauna enhances grassland succession and diversity. *Nature* **422**, 711–713 (2003).
- Mazzoleni, S. et al. Inhibitory and toxic effects of extracellular self-DNA in litter: a mechanism for negative plant–soil feedbacks? *New Phytol.* **205**, 1195–1210 (2015).
- Johnson, D. J., Beaulieu, W. T., Bever, J. D. & Clay, K. Response to comment on “Conspecific negative density dependence and forest diversity”. *Science* **338**, 469 (2012).
- Dickie, I. A., Hurst, J. M. & Bellingham, P. J. Comment on “Conspecific negative density dependence and forest diversity”. *Science* **338**, 469 (2012).

Acknowledgements

This work was conducted as part of the ‘Plant–soil feedback theory’ Working Group at the National Institute for Mathematical and Biological Synthesis, sponsored by the National Science Foundation through NSF Award DBI-1300426, with additional support from The University of Tennessee, Knoxville. The work was further supported by NSF grant DEB-1738041, awarded to J.D.B. We thank K. C. Abbott, J. Bauer, L. S. Comita, K. M. Crawford, S. A. Mangan, J. Umbanhowar and K. N. Suding for discussions of this work during working group meetings, J. Aning for assistance with the FIA data analysis, and T. Markus for assistance with the preparation of the figures. We also thank F. May for comments that improved the manuscript. Finally, we thank all of the FIA employees who collected the data that enabled us to test model predictions.

Author contributions

M.B.E., D.J.J., K.M.L.M., A.E.S. and J.D.B. designed the study. M.B.E. and M.B. performed the analytical model analyses. M.B.E., M.B., J.J., K.M.L.M. and A.E.S. developed the code, and performed and analysed the model simulations. M.B.E., M.B., D.J.J. and J.D.B. analysed the field observation data. M.B.E. and J.D.B. wrote the first draft, after which all authors contributed to improving the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41559-018-0622-3>.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to M.B.E.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a | Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated
- Clearly defined error bars
State explicitly what error bars represent (e.g. SD, SE, CI)

Our web collection on [statistics for biologists](#) may be useful.

Software and code

Policy information about [availability of computer code](#)

Data collection

The data used for the analyses presented in the manuscript are publicly available from www.fia.fs.fed.us/tools-data/, under FIA Data Mart. A .csv file of the specific subset from the original database that was used in this study is available from the corresponding author upon reasonable request.

Data analysis

The Matlab scripts and R scripts written for this study only contain standard functions and analysis methods. Scripts are available from the corresponding author upon reasonable request.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

As noted above, the data used for the analyses presented in the manuscript are publicly available. A .csv file of the specific subset from the original database that was used in this study is available from the corresponding author upon reasonable request.

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/authors/policies/ReportingSummary-flat.pdf](https://www.nature.com/authors/policies/ReportingSummary-flat.pdf)

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	We expanded the two-species plant-soil feedback model of Bever et al. (Journal of Ecology 85: 561-573, 1997) to a general model of frequency-dependent interactions among any number (n) of plant species. We then derived a necessary (but not sufficient) condition for plant community coexistence. We also found that the number coexisting species is predicted to increase with the average strength of negative feedback. This prediction was supported by patterns of tree species diversity in deciduous forest plots in the eastern United States, which can be reproduced in simulations that span the observed range of community feedback.
Research sample	We used publicly available data collected by the US Forest Service, which were previously extracted and pre-processed, as described in the study by Johnson et al. (Science 336: 904-907, 2012). As noted in this previous study, the US Forest Service's Forest Inventory and Analysis (FIA) program is a five-year cyclical inventory of forests in the US in which 20% of the plots in each state are sampled every year. Trees larger than 12.7 cm diameter at breast height (dbh) are recorded in the tree plot (i.e. a subplot) and the abundance of seedlings by species that are smaller than 2.54 cm dbh and taller than 30.5 cm for hardwoods or 12.24 cm for conifers are recorded in the seedling plot (i.e. a microplot).
Sampling strategy	Following Johnson et al. (2012), we treated fully forested plots as independent sample units because distance-dependent effects tend to lose significance 30 m from the tree.
Data collection	We only utilized the previously collected data described above; no additional data was collected for this study.
Timing and spatial scale	As described in Johnson et al. (2012), the selected subset of the FIA dataset comprised data collected between 2004 and 2009, in plots that were east of the 100th meridian longitude.
Data exclusions	As described by Johnson et al. (2012), forest plots that had been disturbed by humans (2% of plots) or fire in the previous five years, which is the time period between sampling within the FIA program, were excluded to limit the influence of human management on tree community structure and diversity. Tree plantations were also excluded from the analysis in order to better capture natural regeneration patterns. After passing the subset of the FIA dataset through these filters, 207,444 plots remained and were included in the analyses.
Reproducibility	All results obtained with stochastic simulations of community assembly were repeated multiple times, to ensure that the results shown in the main text and supplementary information of the manuscript were robust to variations due to stochasticity.
Randomization	n/a; there were no treatments assigned.
Blinding	n/a; there were no treatments assigned.
Did the study involve field work?	<input type="checkbox"/> Yes <input checked="" type="checkbox"/> No

Reporting for specific materials, systems and methods

Materials & experimental systems

- | n/a | Involvement in the study |
|-------------------------------------|--|
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Unique biological materials |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Antibodies |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Eukaryotic cell lines |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Palaeontology |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Animals and other organisms |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Human research participants |

Methods

- | n/a | Involvement in the study |
|-------------------------------------|---|
| <input checked="" type="checkbox"/> | <input type="checkbox"/> ChIP-seq |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Flow cytometry |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> MRI-based neuroimaging |