PULLING ON MULTIPLE THREADS OF EVIDENCE TO REVEAL TREE SPECIES COEXISTENCE

Gordon G. McNickle[1,2,3,*], Morgan V. Ritzi[1,3], Kliffi M.S. Blackstone[1,2,3], John J. Couture[2,4,5], Taylor Nelson [2,4,5], Brady S. Hardiman[5,6], Madeline S. Montague[5], Douglass F. Jacobs[5]

[1]Department of Botany and Plant Pathology, Purdue University, West Lafayette, IN 47907-2054

[2]Center for Plant Biology, Purdue University, West Lafayette, IN 47907-2054

[3]Department of Biology, Purdue University, West Lafayette, IN 47907-2054

[4]Department of Entomology, Purdue University, West Lafayette, IN 47907-2054

[5]Department of Forestry and Natural Resources, Purdue University, West Lafayette, IN 47907-2054

[6]Environmental and ecological engineering, Purdue University, West Lafayette, IN 47907-2054

* correspondence: gmcnickle@purdue.edu

ABSTRACT

Understanding drivers of species coexistence is a central challenge in ecology [1,2]. Coexistence cannot be observed directly, and while species co-occurrence in time and space is necessary for coexistence, it is not sufficient to prove coexistence [3,4]. Species exclusion from a region is potentially observable, but can take decades to occur, and still might occur stochastically [5,6]. Thus, ecologists generally use theory to identify indirect observations that are indicative of mechanisms driving coexistence or exclusion [7-9]. Various methods have been developed to indirectly infer coexistence, each of which requires different data, and none of which are usually conclusive on their own (e.g. [9-12]). Here, we demonstrate agreement using three different approaches examining coexistence of multiple hardwood species. First, in an experimental planting of three mature tree species we found no relationship between productivity and species diversity, which could be due to a lack of niche differences among species. Second, we used modern coexistence theory to calculate niche and fitness differences for each pair of species, which confirmed the lack of niche differences among species, and showed high fitness differences that could create a neutral distribution of species in nature. Third, we used the United States Department of Agriculture Forest Inventory and Analysis data to examine co-occurrence patterns of our species across thousands of natural forest stands and found that indeed, these three species were distributed randomly throughout the USA. Given that these independent methods agree, we take this as strong evidence about a lack of coexistence.

MAIN BODY

Ecological species coexistence has a well-developed body of theory [7,8,13], but is challenging to study empirically [4,14,15]. Strictly, coexistence occurs when the population growth trajectories of multiple species share a stable equilibrium. Unfortunately, natural systems are never at equilibrium. Furthermore, species may simply find themselves thrust together or apart for entirely stochastic neutral reasons [3,4,7]. Ecologists therefore must rely on theory to identify experimental results that

might imply coexistence [9,10,16]. Generally, ecologists expect that differences among species in many dimensions (e.g. resource use, habitat use, seasonality) control coexistence such that there is some limit to how similar two species can be before one excludes the other. The totality of these dimensions, describing the differences among species, is referred to as the niche of a species [1]. Here, we combine multiple approaches to inferring species coexistence. On their own, each approach has proven ambiguous, but when combined here they all point in the same direction providing compelling evidence that ecologists are gaining ground on a predictive theory of coexistence.

First, one approach to indirectly infer mechanisms of coexistence is the relationship between community productivity and species richness [5,17]. Generally, though not universally, researchers have found that more species rich plant communities are also more productive [10,11]. One explanation for this positive diversity-productivity correlation would be if competition within a species was more intense than competition among species [18]. Under this interpretation, low diversity communities grow poorly due to strong negative intraspecific feedbacks, while species rich mixtures grow better because of diluted negative effects caused by niche differences among species. However, this is not the only explanation. Others have argued that positive diversity productivity relationships merely reflects sampling error with an increasing probability of having large productive species as diversity increases [19].

To examine the relationship between productivity and diversity we planted three species of hardwood trees in 2007: (*Prunus serotina* Ehrh., *Quercus rubra* L. and *Castanea dentata* (Marsh.) Borkh.) in all seven possible combinations of one, two or three species, crossed with two different planting densities (1m, or 2m between trees). The experiment has three replicate blocks, and is located in west-central Indiana (40°26'41.9"N 87°01'46.4"W) within the historic range of these three species (Fig S1) [20]. In the summer and fall of 2018 we took wood cores from the base of three individual trees of each species from each plot and used dendrochronology to estimate the

productivity of these plots between 2007-2017. The tree cores give us the diameter of the base of the tree in each year, and we used this to calculate the cross-sectional area of the tree trunk in each year (i.e. basal area). The area of each annual growth ring is an estimate of productivity and is calculated as the area of the tree trunk in year t minus the area in year $t-1$, also known as the basal area increment (BAI). To estimate leaf production in 2018 and 2019, we used cylindrical litter traps (28.6cm diameter). Finally, to estimate fine root production in 2019 we used an ingrowth core method. In autumn 2018, a cylindrical core (7.5cm diameter, 50cm deep) was removed from the forest floor, fine roots were sorted out and the soil replaced inside of a nylon mesh tube. After one year, the core was retrieved in 2019, roots washed, dried and weighed.

We found that the diversity and productivity of leaves, wood and roots were uncorrelated in these plots across all years (Fig 1; Table S1-3). Given contrasting interpretations of diversity productivity relationships, the lack of a positive relationship is either: (i) an indicator of a lack of niche differences among our species[21] or; (ii) simply evidence that our species grow at similar rates[19].

To further probe the alternative explanations of the relationship between diversity and productivity, we used modern coexistence theory (MCET)[7,13]. MCET focuses on population growth rates, not biomass production, and partitions coexistence into two components: stabilizing mechanisms and fitness equalizing mechanisms [7,13]. Stabilizing mechanisms of coexistence are mechanisms that cause one species to limit itself more than other species meaning that the negative effects of intraspecific interactions are stronger than any effects of interspecific interactions [13]. Thus, this allows us to directly and independently examine one explanation for the lack of correlation between diversity and productivity. Stabilizing mechanisms arise because of niche differences such as adaptation to different environments, or differences in biotic interactions [13]. Second, fitness equalizing mechanisms are mechanisms that minimize average fitness differences among species [13]. If fitness is defined as per capita survival and reproduction, then large fitness differences among

species can cause spikes in births or deaths which can negate stabilizing effects. For the lack of a positive relationship between productivity and species richness to be an indicator of non-coexistence rather than simple species growth rates, then we hypothesized that we should find small niche differences in demography data using MCET.

We estimated the slope of per-capita population changes from 2007-2019 as a function of planting density for all three species either in monoculture (α_{ii}), or in each species pair (α_{ij}; Fig S2, Table S4). Niche differences (ND) were calculated as:

$$ND = 1 - \sqrt{\left|\frac{\alpha_{ij}\alpha_{ji}}{\alpha_{ii}\alpha_{jj}}\right|}, \qquad \text{Eqn. 1}$$

and fitness differences (FD) were estimated using the competitive response ratio component of fitness differences, calculated as:

$$FD = \sqrt{\left|\frac{\alpha_{ij}\alpha_{ii}}{\alpha_{ji}\alpha_{jj}}\right|}. \qquad \text{Eqn. 2}$$

These formulations for ND and FD are derived elsewhere [9,22]. ND can be negative if species are extremely similar such that $\alpha_{ij}\alpha_{ji} > \alpha_{ii}\alpha_{jj}$. In the case of negative frequency dependence, regions of coexistence and exclusion can be demarcated by:

$$FD = \frac{1}{1-ND}, \qquad \text{Eqn. 3}$$

such that exclusion occurs above and to the right of this curve, and coexistence occurs below and to the left of the curve [7]. In the case of positive frequency dependence [23], exclusion and priority effects are demarcated by:

$$FD = 1 - ND, \qquad \text{Eqn. 4}$$

where exclusion occurs above and to the right of this curve, and priority effects below and to two the left.

Population growth was generally lower in mixture than in monoculture (Fig S2). MCET calculations reveal that the niche differences among all combinations of the three species in our study were too small to permit species coexistence (Fig 2). Thus, MCET provides evidence that the

lack of a diversity-productivity relationship was due to niche similarity. However, the estimated confidence intervals for niche and fitness differences of *Q. rubra* and *P. serotina* partially overlap with the coexistence region (Fig 2). Furthermore, most species pairs had high fitness differences that would predict stochasticity in co-occurrence of these species across their native ranges. Therefore, we hypothesized that if these species cannot coexist, then they should not be found together in natural forest stands more frequently than expected due to chance. Alternatively, since each species pair also had high fitness differences, these species might occur randomly in nature due to neutral stochastic fitness differences.

To probe these alternative hypotheses about co-occurrence, we used the publicly available USDA Forest Inventory and Analysis (FIA) data. The FIA data are an inventory of trees from thousands of permanent plots across the USA, including data for hundreds of thousands of trees allowing us to examine co-occurrence across the entire native range of our three focal species. *P. serotina*, *Q. rubra* and *C. dentata* all have overlapping ranges in the USA (Fig 3A-C). We used the checkerboard index of co-occurrence [3,24,25] (C-score) to quantify and statistically compare the observed co-occurrence pattern to a null expectation that the three species are randomly distributed throughout their range. We found that the observed negative co-occurrence pattern was not significantly different from random (Fig 3D). This outcome means that, on average, these species do not positively co-occur throughout their range as might be expected if coexistence was common, but nor are they negatively associated as might be expected if exclusion was common. We conclude that the seemingly random association of these three species across their entire range is consistent with large fitness differences (Fig 2) that predict neutral stochastic processes are likely to control the co-occurrence of these three hardwood species.

Here, by pulling on multiple threads of evidence, none of which was conclusive on its own, we found strong evidence that three species of hardwood trees cannot coexist. First, the lack of a

positive relationship between productivity and species richness suggested a lack of niche differences among species (Fig 1), but on its own could be explained by something other than coexistence [19]. Second, MCET independently confirmed the small niche differences. However, MCET also showed large fitness differences that suggest neutral processes may control the distribution of these species. Indeed, we found that the three study species randomly co-occur other across their entire range (Fig 3). Each of these methods has been used separately to try and infer potential mechanisms of coexistence, but on their own are not always conclusive proof of coexistence. Predicting species coexistence has long been a goal of ecological theory, and the confluence of multiple threads of evidence all pointing in the same direction demonstrates that ecologists have made significant progress in uniting theory and observation surrounding species coexistence. Indeed, it was not a given that these three methods would hang together to tell a cohesive story.

We suggest that going forward, combining multiple threads of evidence related to biomass production of monocultures and mixtures, demographic trajectories combined with modern coexistence theory to reveal fitness and niche differences as well as patterns of co-occurrence at large spatial scales can be a powerful set of tools that will better reveal patterns of ecological species coexistence than any one tool on its own. The first two require manipulative experiments that vary both species diversity as well as population density, while the last requires observational data from natural systems. Data projects such as the USDA FIA data, or the US' National Science Foundation National Earth Observatory Network (NSF NEON), or the Smithsonian's ForestGEO project make such large scale patterns of co-occurrence increasingly easy to examine.

Understanding species coexistence, particularly for hardwood trees such as our study species, is not simply a problem for basic science. The world's timber production increasingly relies on planted stands of trees. Planted forests grew from 5% to 7% of global forest cover between 2001 and 2010 and supply as much as 33% of global timber supply [26-28]. Species diversity in forests is generally valued by society, considered to have conservation value for rare species, and can foster important wildlife resources for society [29]. Thus, a greater understanding of how to successfully

combine tree species into high diversity and high productivity timber stands also has significant applied science value too. Excitingly, understanding the mechanisms that drive tree species coexistence turns out to be one of those rare situations where basic and applied science, as well as economic, conservation and societal values are almost uniquely aligned.

METHODS:

Experimental plot design

As described in the main text, we established an experimental planting of trees in west-central Indiana (40°26'41.9"N, 87°01'46.4"W) in the spring of 2007. The main soil type is Rockfield silt loam, and the site is moderately well drained. Mean annual temperature is 10.4 °C and mean annual precipitation is 970mm[20]. Three species of hardwood trees were planted: *Prunus serotina* Ehrh. (black cherry), *Quercus rubra* L. (northern red oak) and *Castanea dentata* (Marsh.) Borkh. (American chestnut) in all seven possible combinations of one, two or three species. These combinations were: (1) *P. serotina* monoculture, (2) *Q. rubra* monoculture, (3) *C. dentata* monoculture, (4) *P. serotina* and *Q. rubra* together, (5) *Q. rubra* and *C. dentata* together, (6) *P serotina* and *C. dentata* together, and (7) all three species together (Fig S1). In addition, we included three density treatments of 1m, 2m or 3m between trees which correspond to 10,000 stems per hectare, 2,500 stems per hectare, or 1,111 stems per hectare respectively. The experiment was arranged into a randomized split plot design where each replicate block contained three densities by seven combinations of species for a total of 21 plots per block. With three replicate blocks this was a total of 63 plots (Fig S1).

Each individual plot always contained 56 trees arranged in a 7 tree by 8 tree grid (Fig S1). Thus, plot size depended on the density treatment. Therefore, plots planted with trees at a 1m spacing were 7 by 8 metres (54m^2), plots planted with trees at 2m spacing were 14 by 16 metres (224 m^2), and plots planted with trees at 3m spacing were 21 by 24 metres (504 m^2). In total, the 63

plots take up approximately 2.4 hectares. In the mixed species plots, trees were planted sequentially across the row with seven trees in a torus wrap which results in a checkerboard arrangement of individual trees within a plot such that individual of the same species were never planted next to each other (E.g. Fig S1B). The 26 perimeter trees around each plot were considered a buffer row to account for potential edge effects and were not measured, leaving 30 focal trees in the middle of each plot (Fig S1B). Thus, across the entire experiment there were 1,648 buffer trees and 1,800 focal trees for a total of 3,438 trees. Every one of the 1,800 focal trees were individually tagged and numbered.

In 2007, trees were planted as bare-root seedlings. Seedlings were obtained from Cascade Forest Nursery in Cascade, Iowa, USA. The *Q. rubra* and *P. serotina* nursery seed source was from stands local to the nursery. Pure *C. dentata* seeds were collected from a forest stand near Galesville, Wisconsin, USA [20]. Prior to planting, 2% glyphosate was applied to eliminate herbaceous vegetation. After planting, a 10-foot-high fence was erected around the perimeter of the entire experiment to prevent damage from deer. The fence was removed in 2012. In addition, a mixture of pendimethalin and glyphosate was applied in the rows between trees in 2008, 2009 and 2010 with backpack and side mounted band sprayer as needed to reduce competition from herbaceous weeds as the seedlings established. In 2011, the rows were mowed. Starting in 2012 no herbicide or mowing was used on the rows as the trees were taller than herbaceous weeds.

Experimental plot data collection

There had been significant mortality between 2007 and 2017 in the southern most 3m spaced plots (Fig S1C), and thus the 3m plots were dropped from most productivity analyses due to lack of replication. Leaf litter was collected in buckets staked to the ground which were 28.6cm diameter, 36 cm deep, and had eight half inch holes drilled into the bottom for drainage. In 2018 and 2019, leaf litter collection began in September, and continued bi-weekly until December. Upon

return to the lab, leaves were sorted by species, dried and weighed. Plot size differed by tree density, and so we applied sampling effort proportional to the plot size. This resulted in a grid of leaf litter traps spaced 6 metres apart across the site with one litter trap in the center of the 1m spaced plots, four in the center of the 2m spaced plots and nine in the center of the 3m plots. We had a total of 258 sampling stations, as the four of the dead southernmost 3m plots were not sampled. Since sampling occurred every two weeks across three months, this generated 1,548 samples which had to be sorted to species. To ensure the most efficient data collection, using the 2019 data we determined that this could be reduced to 1 leaf litter sampling station in the 1m spaced plots, 2 in the 2m spaced plots and 3 in the 3m spaced plots with no change in estimated mean or variance. Sampling stations to be removed were determined randomly. Thus, in 2019, only 114 leaf litter stations were sampled across three months in the autumn. Total leaf production in each year was analyzed by summing across the six sampling dates, and all species within each plot, divided by the area of the litter traps (0.064 m^2). This calculation provided an estimate of leaf production per square metre per year.. The data analysis included the fully factorial combination of year by species combination by density in a linear model (Table S1).

Wood production between 2007 and 2017 was estimated using dendrochronology on cores taken during the 2018 growing season. Of the 1800 total focal trees in the experiment, three of each species in each plot were randomly selected to be sampled for a total of 24 cores per species and 72 cores in total. Due to loss of replicates in the 3m spaced treatments, only the 1m and 2m spaced plots were sampled. Cores were air dried in the lab, mounted on wood blocks and sanded flat with sequentially finer sand-paper ending with 800 grit. Cores were imaged at a resolution of 1200 dots per inch. Ring widths on digitized images were measured in CooRecorder (v 9.00, Cybis Elektronik & Data AB, Saltsjöbaden, Sweden). Because we knew all trees were planted in 2007, it was straightforward to get the diameter of each tree in each year. Assuming the cross-sectional area of the trunk was a circle, we estimated basal area in each year. The area of a ring grown in year t was

therefore estimated as the basal area at the end of year t, minus the basal area at the end of year $t-1$ creating a basal area increment in each year. The data analysis included the fully factorial combination of year by species combination by density in a linear model (Table S2).

Fine root production in 2019 was estimated using root in-growth cores. In this method, a 7.5cm diameter and 50cm deep cylindrical soil core was extracted from the ground in the autumn of 2018. The living roots were sorted out of the soil and then the root-free soil was replaced inside a size 7 mesh cylinder. The in-growth cores were paired with litter trap stations, and so there was one in each 1m plot, two in each 2m spaced plot, and three in each 3m spaced plot for a total of 114 in-growth cores. After one year, in the autumn of 2019, we retrieved the in-growth cores. These were sorted into the top 10cm, the middle 20 cm, and the bottom 20cm, dried and weighed. Analyses summed all three depths within a plot and divided by the total surface area of the soil at the top of the core (0.0044 m^2) to get an estimate of fine root production per year per square metre.

To estimate population demography, we recorded the survival of all 1,800 individual focal trees in 2018. The 3m spaced plots were included in the demography data. We also walked the grid of the entire plot to survey for any recruitment, as all trees were reproductive by 2018. The population growth rate was then regressed against population density using a binomial general linearized model and species identity by species combination by density in the full factorial combination (Fig S2, Table S4). The slope of each individual species line was then used as the estimate of the competition coefficient (α_{ii} or α_{ij}) for modern coexistence theory calculations as described in the main text. Error bars were also estimated as described in the supplementary information.

United States Department of Agriculture Forest Inventory and Analysis Data

In 1999, the Farm Bill (PL 105-185) directed the Forest Service to begin annual surveys of the national inventory of tree resources. These USDA FIA data are publicly available and free for anyone to download. They represent surveys from 124,563 permanent plots across the entire USA allowing us to examine the co-occurrence of our three study species across their entire native range [30]. An FIA plot is made up of four sub-plots with a radius of 24 feet (7.3 m) arranged in a triangle, with one subplot at the center, and the other subplots at the corners of an equilateral triangle centred on the first subplot. The identity and size of all trees with a diameter of ≥5 inches (12.7 cm) at a height of 4.5 feet (1.37m) in each subplot are then recorded. Only the presence or absence of each species at the plot level was used in our analyses to understand patterns of co-occurrence of our study species across their entire native range.

Checkerboard score analysis

The checkerboard score is named for the appearance of a perfect negative or positive co-occurrence matrix, which takes on the appearance of a checkerboard based on the diagonal arrangement of 0s and 1s. Here we detail how to calculate the C-score. First the number of checkerboard units for each pair of species is given by:

$$C_{ij} = (r_i - X_{ij})(r_j - X_{ij}), \qquad \text{Eqn 4}$$

where r_i is the number of times species i occurs without species j, r_j is the number of times species j occurs without species i and X_{ij} is the number of sites where species i and j co-occur [3]. Typically, C_{ij} is normalized by the number of possible species pairs (P) and the number of possible site pairs (N), which using basic combinatorics are given by

$$P = S(S-1)/2 \text{ and } N = n(n-1)/2, \qquad \text{Eqns 5}$$

where S is the number of species, and n is the total number of replicate plots [3,16]. Across all sites, the normalized C-score is given by:

$$C_n = \sum_{j=0}^{S} \sum_{i<j} \frac{C_{ij}}{PN}. \qquad \text{Eqn 6}$$

To estimate the null expectation, we used a fixed-equiprobable randomization where row sums (i.e. species occurrences) are held constant, but the column sums (species occurrence within sites) are allowed to vary with equiprobability [31]. A fixed-fixed randomization is more conservative and less prone to type-I errors, but due to the rarity of *C. dentata*, it cannot be applied to the frequent two species communities containing just *P. serotina* and *Q. rubra.* This is because retaining the row and column sums of a two species matrix rather trivially always returns the original occurrence matrix when there are only two species. Null community matrices were created using the permatswap function in the vegan library in R [32,33]. For each observed matrix, 1,000 null matrices were sampled with 30,000 swaps between each sample and with a burn-in of 30,000 swaps prior to the first sample. We used the quasiswap method for generating null matrices, which does not produce sequential null matrices but instead generates a matrix at each time step that is fully independent of previous matrices [34]. The standardized effect size of the C_n-score can be calculated based on the observed C_n-score (C_n) and the mean (μ) and standard deviation (σ) of the expected C_n-scores produced from the 1000 null matrices by calculating a z statistic as [25,35]:

$$z = \frac{C_n - \mu}{\sigma}. \qquad \text{Eqn 7}$$

The z statistic is significant at the two-tailed p=0.05 level for $-1.96<z>1.96$.

ACKNOWLEDGEMENTS

This work was funded by the American Chestnut Foundation, a Purdue AgSeed award, and a Hardwood Tree Improvement and Regeneration Centre grant to GGM, DFJ, JJC and BSH, and Hatch awards 1010722 to GGM and ND011490 to JJC, and McIntire Stennis project IND011535 to DFJ. The authors declare no conflicts of interest.

AUTHOR CONTRIBUTIONS

MVR, KMSB and GGM performed dendrochronology analysis. MSM, DFJ and GGM performed root ingrowth core analyses. KMSB, TN, JJC and GGM performed leaf analyses. All authors contributed to tree tagging and survival data collection. GGM wrote the manuscript, performed MCET and C-score calculations and all authors contributed to revisions and discussion of approaches.

DATA AVAILABILITY

The USDA FIA data are publically available from the USDA. Our site specific productivity and demography data are publically available on DRYAD at: https://doi.org/10.5061/dryad.280gb5mqc. All code required to reproduce analyses, figures and tables are publically available on Zenodo at: https://doi.org/10.5281/zenodo.5151988

335 **REFERENCES**

336 1 Hutchinson, G. E. Homage to santa-rosalia or why are there so many kinds of animals. *Am.*
337 *Nat.* **93**, 145-159 (1959).
338 2 Silvertown, J. Plant coexistence and the niche. *Trends Ecol. Evol.* **19**, 605-611, doi:DOI
339 10.1016/j.tree.2004.09.003 (2004).
340 3 Stone, L. & Roberts, A. The checkerboard score and species distributions. *Oecologia* **85**, 74-
341 79, doi:10.1007/bf00317345 (1990).
342 4 Blanchet, F. G., Cazelles, K. & Gravel, D. Co-occurrence is not evidence of ecological
343 interactions. *Ecol. Lett.* **23**, 1050-1063, doi:10.1111/ele.13525 (2020).
344 5 Tilman, D. *et al.* Diversity and productivity in a long-term grassland experiment. *Science* **294**,
345 843-845, doi:10.1126/science.1060391 (2001).
346 6 Crawley, M. J. *et al.* Determinants of species richness in the park grass experiment. *The*
347 *American Naturalist* **165**, 179-192, doi:10.1086/427270 (2005).
348 7 Adler, P. B., HilleRisLambers, J. & Levine, J. M. A niche for neutrality. *Ecol. Lett.* **10**, 95-104,
349 doi:10.1111/j.1461-0248.2006.00996.x (2007).
350 8 Hart, S. P., Freckleton, R. P. & Levine, J. M. How to quantify competitive ability. *J. Ecol.* **106**,
351 1902-1909, doi:doi:10.1111/1365-2745.12954 (2018).
352 9 Kraft, N. J. B., Godoy, O. & Levine, J. M. Plant functional traits and the multidimensional
353 nature of species coexistence. *Proceedings of the National Academy of Sciences* **112**, 797-
354 802, doi:10.1073/pnas.1413650112 (2015).
355 10 Liang, J. J. *et al.* Positive biodiversity-productivity relationship predominant in global forests.
356 *Science* **354**, 12, doi:10.1126/science.aaf8957 (2016).
357 11 Grace, J. B. *et al.* Integrative modelling reveals mechanisms linking productivity and plant
358 species richness. *Nature* **529**, 390-393, doi:10.1038/nature16524 (2016).
359 12 McNickle, G. G. *et al.* Checkerboard score–area relationships reveal spatial scales of plant
360 community structure. *Oikos* **127**, 415-426, doi:10.1111/oik.04620 (2018).
361 13 Chesson, P. Mechanisms of maintenance of species diversity. *Annu. Rev. Ecol. Syst.* **31**, 343-+
362 (2000).
363 14 Wright, S. J. Plant diversity in tropical forests: A review of mechanisms of species
364 coexistence. *Oecologia* **130**, 1-14, doi:10.1007/s004420100809 (2002).
365 15 Avolio, M. L. *et al.* A comprehensive approach to analyzing community dynamics using rank
366 abundance curves. *Ecosphere* **10**, e02881, doi:https://doi.org/10.1002/ecs2.2881 (2019).
367 16 Ulrich, W. & Gotelli, N. J. Pattern detection in null model analysis. *Oikos* **122**, 2-18,
368 doi:10.1111/j.1600-0706.2012.20325.x (2013).
369 17 Grime, J. P. Control of species density in herbaceous vegation. *J. Environ. Manage.* **1**, 151-
370 167 (1973).
371 18 Lotka, A. J. The growth of mixed populations: Two species competing for a common food
372 supply. *Journal of the Washington Academy of Sciences* **22**, 461-469 (1932).
373 19 Cheng, Y. X., Zhang, C. Y., Zhao, X. H. & von Gadow, K. Biomass-dominant species shape the
374 productivity-diversity relationship in two temperate forests. *Ann. For. Sci.* **75**, 9,
375 doi:10.1007/s13595-018-0780-0 (2018).
376 20 Gauthier, M. M., Zellers, K. E., Lof, M. & Jacobs, D. F. Inter- and intra-specific
377 competitiveness of plantation-grown american chestnut (castanea dentata). *For. Ecol.*
378 *Manage.* **291**, 289-299, doi:10.1016/j.foreco.2012.11.014 (2013).
379 21 Cardinale, B. J. *et al.* Impacts of plant diversity on biomass production increase through time
380 because of species complementarity. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 18123-18128,
381 doi:10.1073/pnas.0709069104 (2007).
382 22 Narwani, A., Alexandrou, M. A., Oakley, T. H., Carroll, I. T. & Cardinale, B. J. Experimental
383 evidence that evolutionary relatedness does not affect the ecological mechanisms of

coexistence in freshwater green algae. *Ecol. Lett.* **16**, 1373-1381, doi:10.1111/ele.12182 (2013).

23 Grainger, T. N., Letten, A. D., Gilbert, B. & Fukami, T. Applying modern coexistence theory to priority effects. *Proceedings of the National Academy of Sciences* **116**, 6205-6210, doi:10.1073/pnas.1803122116 (2019).

24 Diamond, J. in *Ecology and evolution of communities* (eds ML. Cody & JM. Diamond) 342-444 (Belknap Press, 1975).

25 Ulrich, W. & Gotelli, N. J. Null model analysis of species associations using abundance data. *Ecology* **91**, 3384-3397, doi:10.1890/09-2157.1 (2010).

26 FAO. Global forest resources assessment 2010, main report. 85 (Food and Agriculture Organization of the United Nations, Rome, 2010).

27 FAO. Global forest resources assessment 2001, main report., 479 (Food and Agriculture Organization of the United Nations, Rome, 2001).

28 Jurgensen, C., W., K. & Lebedys, A. Assessment of industrial roundwood production from planted forests. (Department of Agriculture, Forest Service, Forest Products Laboratory, Rome, Italy, 2014).

29 Faith, D. P. Conservation evaluation and phylogenetic diversity. *Biological Conservation* **61**, 1-10, doi:10.1016/0006-3207(92)91201-3 (1992).

30 Gray, A. N., Brandeis, T. J., Shaw, J. D., McWilliams, W. H. & Miles, P. D. Forest inventory and analysis database of the united states of america (fia). *Biodiversity and Ecology.* **4**, 225-231, doi:https://doi.org/10.7809/b-e.00079 (2012).

31 Gotelli, N. J. Null model analysis of species co-occurrence patterns. *Ecology* **81**, 2606-2621, doi:10.1890/0012-9658(2000)081[2606:NMAOSC]2.0.CO;2 (2000).

32 Hardy, O. J. Testing the spatial phylogenetic structure of local communities: Statistical performances of different null models and test statistics on a locally neutral community. *J. Ecol.* **96**, 914-926, doi:10.1111/j.1365-2745.2008.01421.x (2008).

33 Vegan: Community ecology package v.2.4-1 (2016).

34 Miklós, I. & Podani, J. Randomization of presence–absence matrices: Comments and new algorithms. *Ecology* **85**, 86-92, doi:10.1890/03-0101 (2004).

35 Ulrich, W. & Gotelli, N. J. Disentangling community patterns of nestedness and species co-occurrence. *Oikos* **116**, 2053-2061, doi:10.1111/j.2007.0030-1299.16173.x (2007).

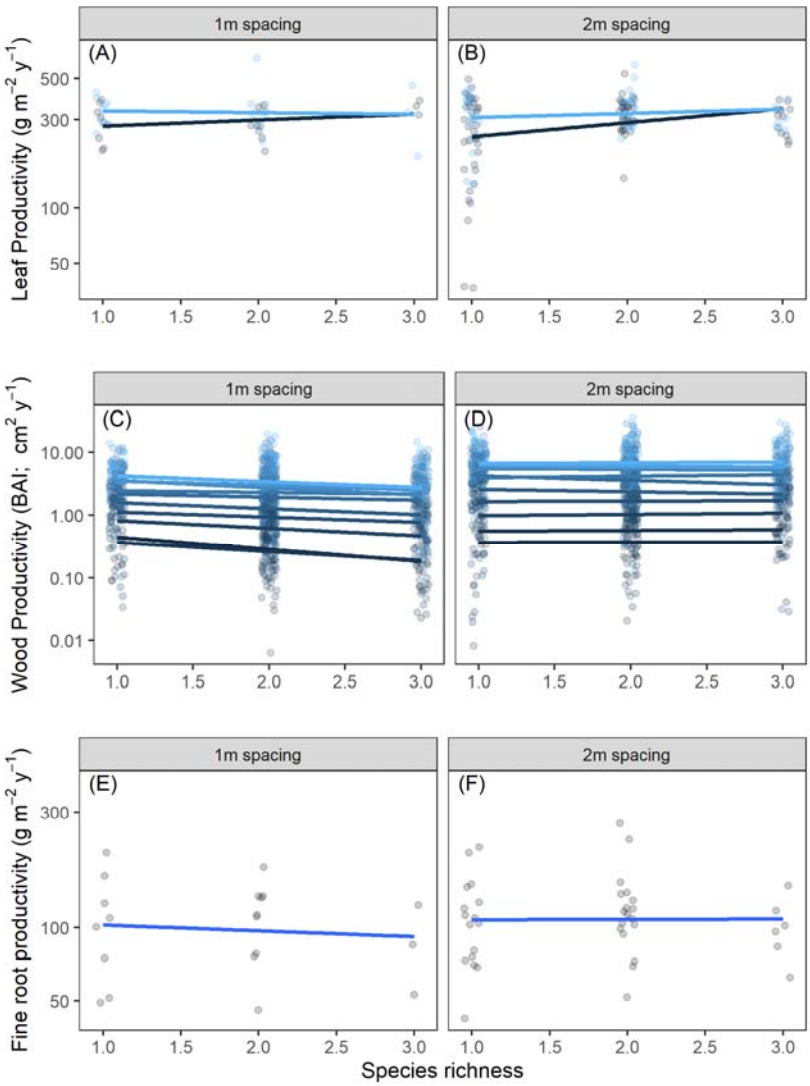
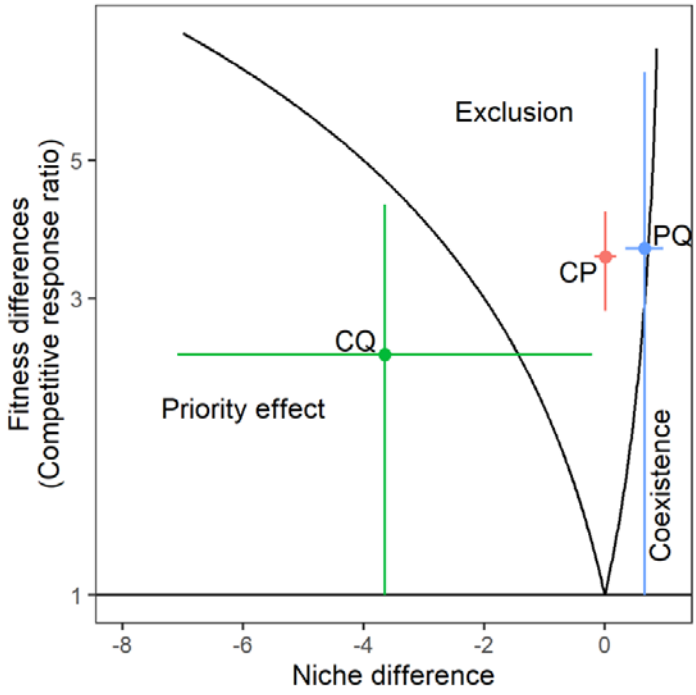


Fig 1: The relationship between species richness, and wood productivity estimated from leaf litter traps from 2018-2019 (**A,B**); dendrochronology as basal area increments from 2007-2017 (BAI; **C,D**), and fine root productivity in 2019 (**E,F**). The y-axis is on a $\log_{10}$ scale in all panels. Trees were spaced either 1m or 2m apart (1 tree/m^2, or 0.5 trees/m^2 respectively). Points are drawn with a jitter around richness, and the y-axis is plotted on a $\log_{10}$ scale. Lines come from a linear regression (Tables S1-3). None of the estimates for the individual slopes of the diversity productivity relationship were different from zero.



425

426 **Fig 2**: Comparison of niche and fitness differences calculated using population growth rates and

427 MCET for each pair of species: *P. serotina* vs *Q. rubra* (PQ); *C. dentata* vs *P. serotina* (CP) and *C.*

428 *dentata* vs *Q. rubra* (CQ). Since our propagation of error calculations could not account for

429 covariance this makes them larger, thus, error bars are 80% confidence intervals. The black lines

430 demarcate the regions of coexistence, priority effects and exclusion as labeled.

431

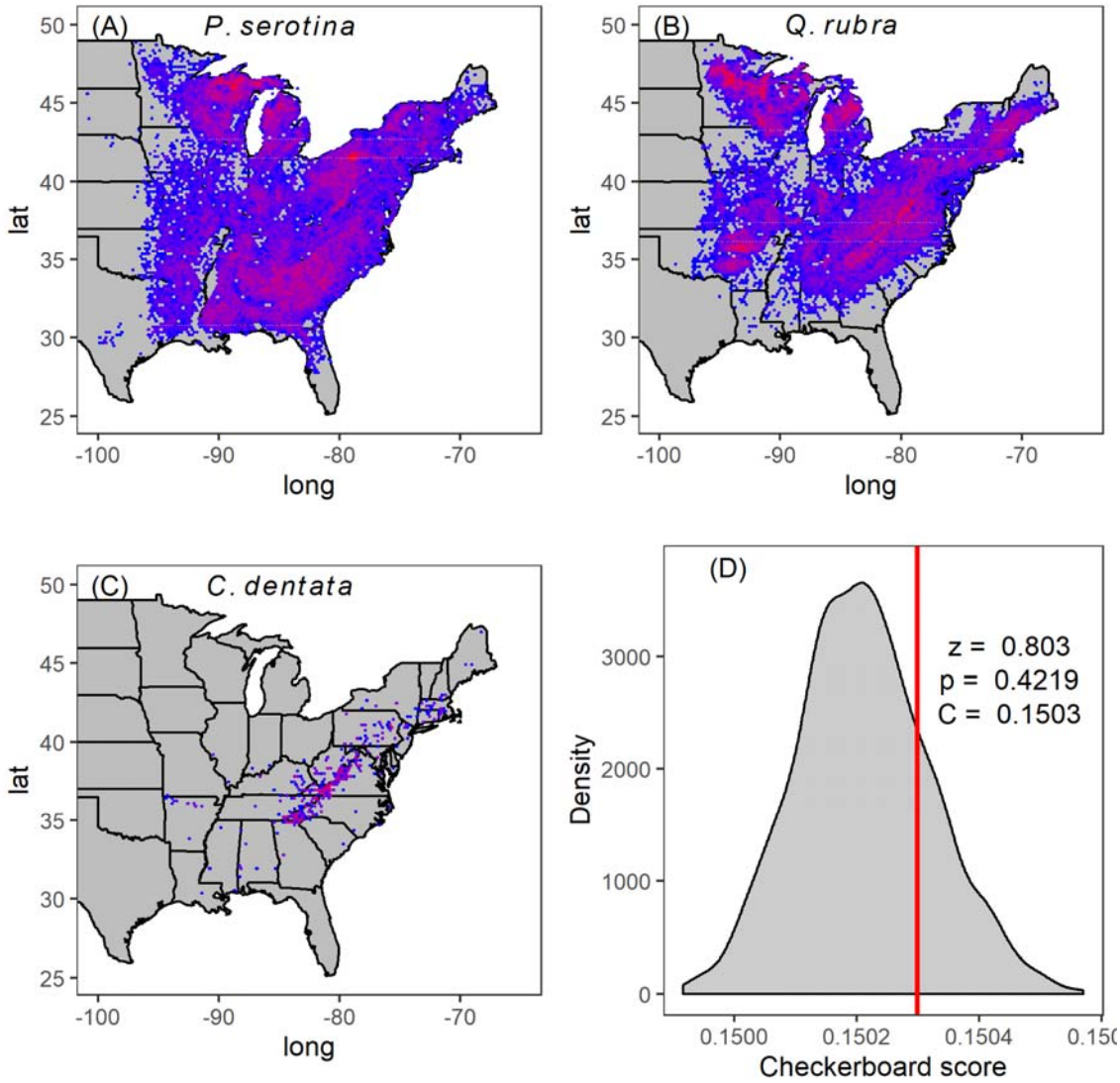


Fig 3: Distribution and population density (count) in natural forest stands of *P. serotina* **(A)** *Q. rubra* **(B)** and *C. dentata* **(C)** across the FIA plots in the USA. Red represents higher density, while blue represents low density. We compared observed C-score (red vertical line) with the null expectation of random distribution (grey histogram) across all FIA plots **(D)**. The C score is high (relative to null) when species negatively co-occur, and low (relative to null) when they positively co-occur. The lack of difference compared to null (p=0.4219) suggests these three species are randomly distributed in nature with respect to each other on average.

SUPPLEMENTARY METHODS:

Error propagation in modern coexistence theory

Since the components of modern coexistence theory are typically predicted from regression coefficients of per-capita population growth rates (i.e. fitness) versus population density [8] that have an associated variance-covariance matrix it is possible to use the known propagation of error methods to create confidence intervals around the estimates of niche and fitness differences (e.g. [23]). Here, we outline the approach we took to estimating error for niche and fitness differences. This approach assumes no covariance between terms, which actually leads to larger errors than if we assumed covariance, and therefore can be considered a conservative approach to estimating error.

First, to obtain the competition coefficients (α) as the slopes of the relationship between per capita survival and population density, we must sum individual m parts of the regression model shown in Table S2 to obtain the species and treatment specific slopes (β_i) according to:

$$f_\alpha = \sum_{i}^{m} \beta_i \qquad Eqn\ S1$$

Error propagates in the sum of m estimates according to:

$$\sigma_\alpha^2 = \sum_{i=1}^{m} \sigma_i^2. \qquad Eqn\ S2$$

Once the competition coefficients are estimated, the niche and fitness differences are combined into ratios via multiplication and division using Eqns 1 and 2 in the main text. For the purposes of this derivation, we can write this as:

$$f_r = \prod_{i=1}^{l} \alpha_i, \qquad Eqn\ S3$$

where, quotient of A by B is simply written as the product of $A \times \frac{1}{B}$. Error propagates through multiplication or division, for l elements of value α_i to create variance for the internal ratios of competition coefficients according to:

$$\sigma_r^2 = \sqrt{\sum_{i=1}^{l} \left(\frac{\sigma_{\alpha,i}^2}{\alpha_i}\right)^2}. \qquad Eqn\ S4$$

Finally, niche and fitness differences both involve taking the square root of the competition coefficient ratios (f_r). The final variance after taking the square root becomes

$$\sigma_{MCET}^2 = \frac{\sigma_r^2}{2f_r}. \qquad Eqn\ S5$$

With the final variance for either niche or fitness differences in hand, the standard deviation and standard error can be used to obtain a confidence interval for both niche and fitness differences.

TABLE S1: Regression model for leaf data shown in Fig 1A,B. The model was lm(Leaf.production~ density * diversity * year) in R. The overall regression was not significant ($F_{7,160}$ = 1.66, p=0.1217), and critically, none of the individual slopes for the diversity productivity relationship were significantly different from zero.

Coefficients	Estimate	SE	t	p
Intercept	260.6	51.2	5.1	<0.001
Diversity	21.3	27.7	0.77	0.443
Density(2m)	-10.5	57.3	-0.18	0.855
Year(2019)	76.2	72.4	1.0	0.295
Diversity*Density(2m)	6.0	30.9	0.19	0.847
Diversity*Year(2019)	-19.8	39.1	-0.51	0.614
Density(2m)*Year(2019)	-13.0	84.9	-0.15	0.88
Diversity*Density(2m)*Year(2019)	2.71	45.9	0.06	0.953

TABLE S2: Regression model for wood productivity shown in Fig 1C,D. The model was lm(BAI ~ density * diversity * year) in R. The regression was significant ($F_{43,1925}$ = 24.1, p<0.0001), though the data contained a large amount of unexplained error (R^2 = 0.33). None of the individual slopes for the diversity productivity relationship were significantly different from zero.

Coefficients	Estimate	SE	t	p
Intercept	0.58	2.05	0.284	0.7761
Density(2m)	0.39	2.50	0.154	0.8773
Year(2008)	0.08	2.33	0.034	0.9728
Year(2009)	0.93	2.29	0.405	0.6859
Year(2010)	1.10	2.28	0.484	0.6283
Year(2011)	1.81	2.27	0.799	0.4246
Year(2012)	2.57	2.26	1.134	0.2567
Year(2013)	2.60	2.26	1.147	0.2516
Year(2014)	4.05	2.26	1.788	0.074
Year(2015)	4.29	2.26	1.892	0.0586
Year(2016)	5.24	2.26	2.314	0.0208
Year(2017)	4.79	2.26	2.114	0.0346
Diversity	-0.08	0.94	-0.087	0.9307
Density(2m):Year(2008)	0.28	2.96	0.094	0.925
Density(2m):Year(2009)	0.25	2.88	0.086	0.9312
Density(2m):Year(2010)	0.66	2.87	0.229	0.8193
Density(2m):Year(2011)	1.52	2.86	0.533	0.5943
Density(2m):Year(2012)	2.51	2.85	0.88	0.3787
Density(2m):Year(2013)	1.68	2.85	0.592	0.5541
Density(2m):Year(2014)	1.27	2.85	0.447	0.6546
Density(2m):Year(2015)	2.61	2.85	0.918	0.3585
Density(2m):Year(2016)	1.46	2.85	0.514	0.6076
Density(2m):Year(2017)	2.14	2.85	0.75	0.4531
Diversity:Density(2m)	-0.08	1.16	-0.066	0.9472
Diversity:Year(2008)	-0.01	1.07	-0.009	0.9931
Diversity:Year(2009)	-0.17	1.05	-0.162	0.8712
Diversity:Year(2010)	-0.06	1.05	-0.057	0.9542
Diversity:Year(2011)	-0.25	1.04	-0.239	0.8108
Diversity:Year(2012)	-0.15	1.04	-0.143	0.8866
Diversity:Year(2013)	0.03	1.04	0.026	0.9791
Diversity:Year(2014)	-0.46	1.04	-0.443	0.6576
Diversity:Year(2015)	-0.40	1.04	-0.38	0.7037
Diversity:Year(2016)	-0.69	1.04	-0.659	0.51
Diversity:Year(2017)	-0.26	1.04	-0.25	0.8025
Diversity:Density(2m):Year(2008)	0.02	1.38	0.011	0.9913
Diversity:Density(2m):Year(2009)	0.07	1.34	0.052	0.9587
Diversity:Density(2m):Year(2010)	0.12	1.34	0.09	0.9284
Diversity:Density(2m):Year(2011)	-0.04	1.33	-0.034	0.9727
Diversity:Density(2m):Year(2012)	0.39	1.33	-0.295	0.7681
Diversity:Density(2m):Year(2013)	0.25	1.33	0.186	0.8523
Diversity:Density(2m):Year(2014)	0.81	1.33	0.614	0.5396
Diversity:Density(2m):Year(2015)	0.26	1.33	0.197	0.8438
Diversity:Density(2m):Year(2016)	0.80	1.33	0.606	0.5446
Diversity:Density(2m):Year(2017)	0.82	1.33	0.619	0.5358

TABLE S3: Regression model for root data shown in Fig 1E,F. The model was lm(Root.production~ density * diversity) in R. The overall regression was not significant ($F_{3,56}$ = 0.27, p=0.8476), and critically, none of the individual slopes for the diversity productivity relationship were significantly different from zero.

Coefficients	Estimate	SE	t	p
Intercept	121.5	28.5	4.3	<0.0001
Diversity	-8.3	15.1	-0.55	0.585
Density(2m)	-2.8	34.9	-0.08	0.936
Diversity*Density(2m)	6.7	18.5	0.364	0.717

TABLE S4: Regression coefficients for the logistic curves fit in Fig S2 and used to calculate competitive ability parameters for niche and fitness differences. The model was glm(survival ~ density * species + treatment*species) in R using a binomial distribution to achieve a logistic regression. Statistically significant coefficients are shown in boldface font. Modern coexistence theory calculations involve ratios of competition coefficients. Thus, to avoid dividing by zero, we dealt with non-significant regression coefficients by incorporating error into our modern coexistence theory calculations. This allows non-significant parameters to have their appropriately high error.

Coefficients	Estimate	SE	t	p
Intercept	**2.4**	**1.08**	**2.13**	**0.0268**
Density	-0.65	1.44	-0.450	0.6528
Species(*P. serotina*)	**2.18**	**0.88**	**2.48**	**0.0133**
Species(*Q. rubra*)	**2.56**	**0.97**	**2.64**	**0.0082**
Treatment(C)	0.42	1.24	0.34	0.7343
Treatment(Q)	-1.53	1.42	-1.08	0.2797
Treatment(PQ)	**-2.44**	**0.97**	**-2.50**	**0.0124**
Treatment(CP)	-1.19	0.95	-1.25	0.2098
Treatment(CQ)	-1.69	1.16	-1.46	0.1452
Density:Species(*P. serotina*)	**-3.67**	**1.25**	**-2.94**	**0.0033**
Density:Species(*Q. rubra*)	**-4.41**	**1.47**	**-3.00**	**0.0027**
Density:treatment(C)	-0.06	1.65	-0.04	0.9696
Density:treatment(Q)	**4.79**	**1.95**	**2.46**	**0.0141**
Density:treatment(PQ)	**4.65**	**1.31**	**3.53**	**0.0004**
Density:treatment(CP)	1.87	1.13	1.65	0.0999
Density:treatment(CQ)	**3.70**	**1.61**	**2.30**	**0.0218**

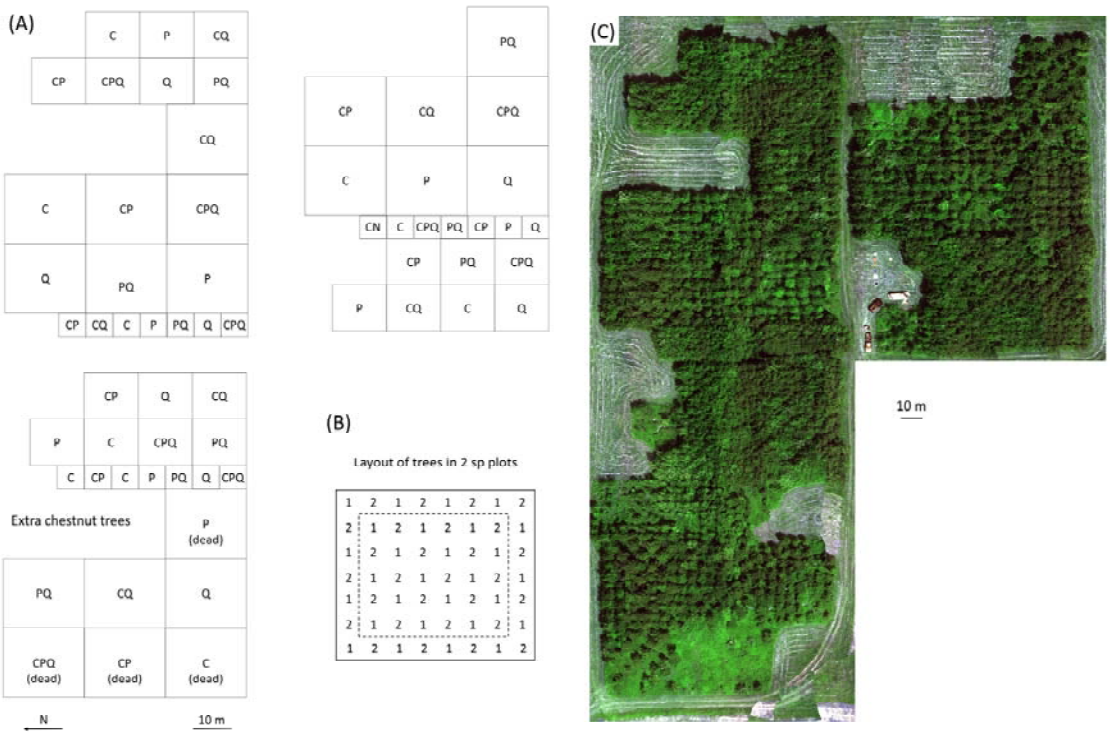


FIG S1: Plot layout. **(A)** The experiment has three replicate blocks that include all possible combinations of *C. dentata* (C, American chestnut), *P. serotina* (P, black cherry) or *Q. rubra* (Q, northern red oak), at three different planting densities (1m, 2m or 3m between trees; 1 tree/m^2, 0.5 trees/m^2, or 0.33 trees/m^2 respectively) in a full factorial split plot design. Unfortunately, many trees in the southern most 3m spacing plots died due to an unusual flooding event. These plots are not included in productivity data due to inadequate sample size for analysis, however these plots are included in demography data. Surplus chestnut trees were planted in an open space at the southern end of the experiment because of their conservation value. **(B)** Each plot contains 56 trees that alternate between species (indicated by 1 and 2). There is a buffer row of trees that are not measured around each plot (dashed line) leaving 30 focal trees in the center of each plot, and 1260 individual focal trees in 1m and 2m spaced plots. (C) An aerial image of our plots showing the basic layout. The dead 3m spaced plots, and the extra chestnut trees that are not part of the experiment, are especially evident in this image. Field vehicles are also visible in the center of the image. Full methodological details about the planting and experimental design can be found in: [20].

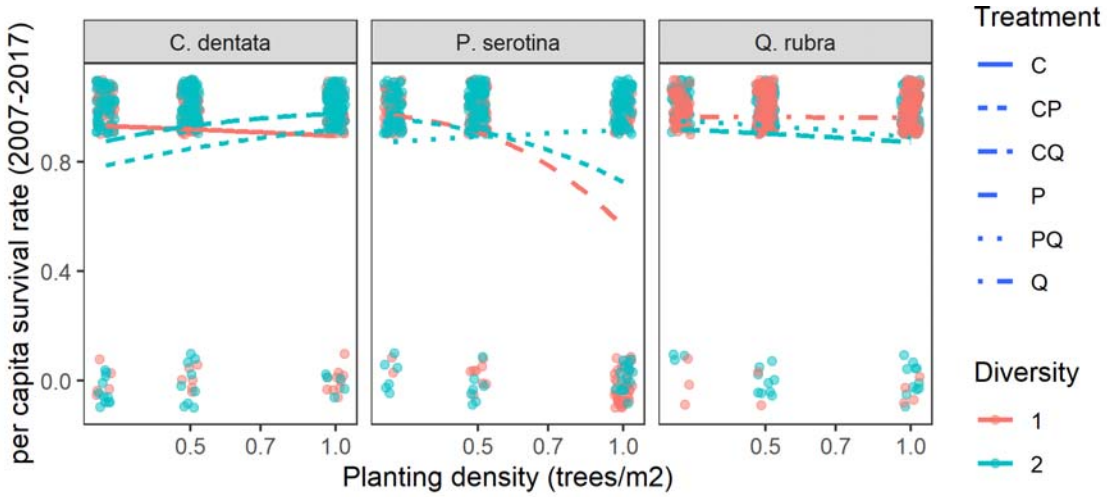


Fig S2: Per capita survival rates (our estimate of fitness) across a ten-year interval for each monoculture, or two species mixture for all three species: *C. dentata* (C, American chestnut), *P. serotina* (P, black cherry) or *Q. rubra* (Q, northern red oak). The slope of each line is shown above and was our estimate of competition coefficients (α) used to calculate stabilizing and equalizing mechanisms of coexistence. Lines represent logistic generalized linear model fits across all data (regression coefficients given in Table S2). All y data are either 0 or 1, and points are plotted with a jitter on x and y to aide in visualisation of the density of points at each position.