nature plants

LETTERS

https://doi.org/10.1038/s41477-018-0231-9

# Light regulates tropical symbiotic nitrogen fixation more strongly than soil nitrogen

Benton N. Taylor* and Duncan N. L. Menge

**Nitrogen limits primary production in almost every biome on Earth[1,2]. Symbiotic nitrogen fixation, conducted by certain angiosperms and their endosymbiotic bacteria, is the largest potential natural source of new nitrogen into the biosphere[3], influencing global primary production, carbon sequestration and element cycling. Because symbiotic nitrogen fixation represents an alternative to soil nitrogen uptake, much of the work on symbiotic nitrogen fixation regulation has focused on soil nitrogen availability[4–8]. However, because symbiotic nitrogen fixation is an energetically expensive process[9], light availability to the plant may also regulate symbiotic nitrogen fixation rates[10,11]. Despite the importance of symbiotic nitrogen fixation to biosphere functioning, the environmental factors that most strongly regulate this process remain unresolved. Here we show that light regulates symbiotic nitrogen fixation more strongly than does soil nitrogen and that light mediates the response of symbiotic nitrogen fixation to soil nitrogen availability. In a shadehouse experiment, low light levels (comparable with forest understories) completely shut down symbiotic nitrogen fixation, whereas soil nitrogen levels that far exceeded plant demand did not fully downregulate symbiotic nitrogen fixation at high light. For in situ forest seedlings, light was a notable predictor of symbiotic nitrogen fixation activity, but soil-extractable nitrogen was not. Light as a primary regulator of symbiotic nitrogen fixation is a departure from decades of focus on soil nitrogen availability. This shift in our understanding of symbiotic nitrogen fixation regulation can resolve a long-standing biogeochemical paradox[12], and it will improve our ability to predict how symbiotic nitrogen fixation will fuel the global forest carbon sink and respond to human alteration of the global nitrogen cycle.**

Symbiotic nitrogen (N)-fixing plants represent one of the largest natural sources of new N into the biosphere. These symbiotic 'N fixers' have the potential to relieve N limitation by converting atmospheric $N_2$ gas into bio-available forms, potentially increasing the terrestrial carbon sink, but whether they do so depends on how much N they fix. Many N fixers can regulate how much N they fix per unit biomass[5–7,13], but which environmental factors govern this regulation remain largely unresolved.

The availability of soil nutrients[4,6,8,13], light[10,11], water[14] and temperature[15] have all been suggested to play a role in regulating symbiotic nitrogen fixation (SNF). Of these, soil resources, especially soil N availability, have received the most attention. Because SNF represents a direct alternative to soil N uptake, the availability of soil N serves as a logical regulator of SNF. However, because SNF is energetically expensive[9], an N fixer's access to light may also determine how much N the plant fixes. Although existing theory suggests that light could strongly regulate SNF[4,10,11] and light has been shown to regulate SNF in some agricultural N fixers (for example, Murphy[16], MacDowall[17] and Lau[18]), little empirical work has directly tested this mechanism in natural systems[19–21].

We paired a shadehouse experiment with natural field sampling under varying conditions of light and soil N availability to ask: does light or soil N availability have a stronger influence on the regulation of SNF? Our shadehouse experiment grew individuals of a common neotropical N fixer, *Pentaclethra macroloba*, under a full-factorial design of three soil N treatments and three light treatments. Although our light treatments (8%, 16% and 40% full sunlight) represented a subset of possible natural conditions a plant could experience[22], the high end of our N treatments (0.51, 20 and 40 g N $m^{-2}$ $yr^{-1}$ added to a sand/soil mix) far exceeded natural N conditions[23]. *P. macroloba* individuals were raised from seed for 6 months in an open-air shadehouse in Costa Rica, after which the plants were harvested to measure biomass growth, allocation to root nodules (the symbiotic structures where SNF occurs) and rates of SNF.

Light limited the biomass growth of *P. macroloba* in the low- and medium-light treatments (Fig. 1a–c, Supplementary Table 1) similar to field observations for this species[24]. Low N levels limited plant growth in the high-light treatments only (Fig. 1c; corrected Akaike Information Criterion [$\Delta AIC_c$] = 0.45 compared with the next best model and $\Delta AIC_c > 164$ compared with our model of N limitation in all light treatments, Supplementary Table 1). Nodule biomass responded most strongly to light availability: Within low-N treatments, N fixers allocated greater than 230-fold more of their belowground biomass to nodules in high versus low light (Fig. 1d–f). The addition of soil N also influenced nodule biomass, but to a much smaller degree. In the high-light treatments, N fixers allocated 7.7-fold more of their belowground biomass to nodules in low- versus high-N fertilization ($\Delta AIC_c = 1.95$ compared with the next best model and $\Delta AIC_c > 109$ compared with our model of N regulation of nodulation).

Using an isotopic soil $^{15}N$ tracer to track the various sources of plant N, we calculated the percentage of each plant's N it derived from SNF ($\%N_{dfa}$). $\%N_{dfa}$ varied across treatments in a similar pattern to nodulation (Fig. 1g–i). In the high-light treatments, plants fixed 34% and 20% of their N in low- and high-N treatments, respectively, whereas plants fixed 0% of their N in the medium- and low-light treatments, regardless of N conditions ($\Delta AIC_c = 1.98$ compared with the next best model and $\Delta AIC_c > 125$ compared with our model of N regulation of $\%N_{dfa}$). We also found that total N fixed by each plant ($\%N_{dfa}$ multiplied by total N in each plant) was much more strongly influenced by light than by soil N (Fig. 2; $\Delta AIC_c = 1.54$ compared with the next best model and $\Delta AIC_c > 123$ compared with our model of N regulation of fixed N, Supplementary Table 1).

Interestingly, light availability at 8% of full sunlight (more than many forest understories[25]) completely inhibited N fixation in

Department of Ecology, Evolution, and Environmental Biology, Columbia University, New York, NY, USA. *e-mail: bentonneiltaylor@gmail.com

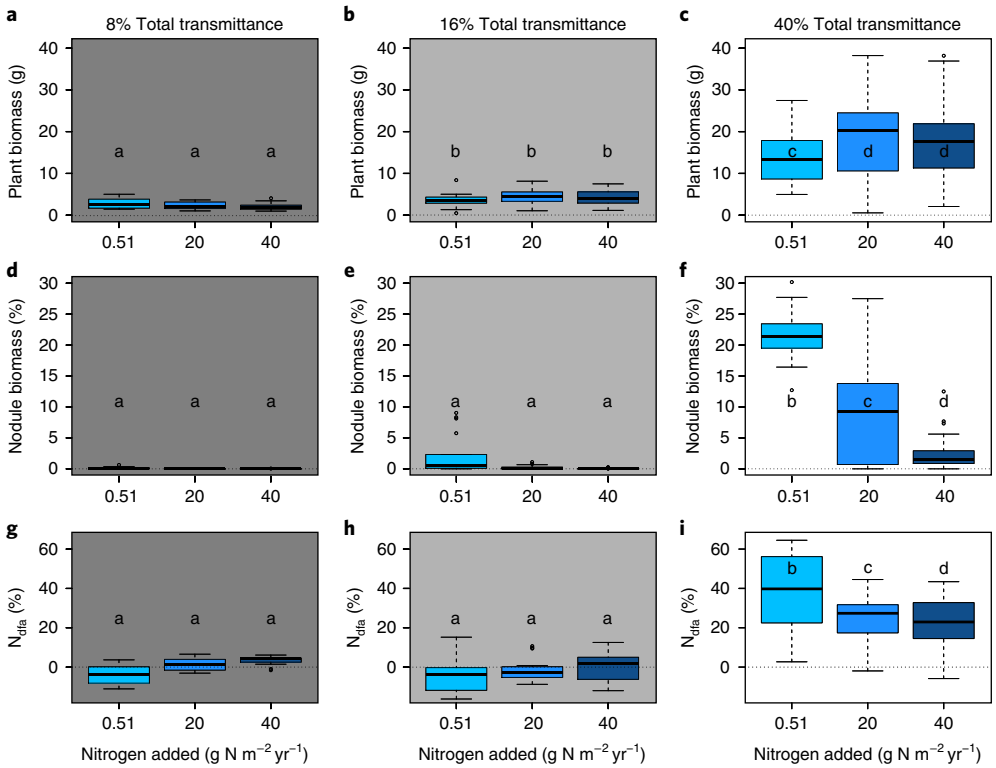


**Fig. 1 | Light is a stronger driver than soil N for plant biomass and N fixation in shadehouse-grown plants. a–i**, Plant biomass (**a–c**), allocation to nodules (% of belowground biomass) (**d–f**) and $\%N_{dfa}$ (**g–i**) all varied more strongly across light treatments than across N treatments. Light treatments are shown as the left, middle and right columns. N treatments are represented by the three boxes within each panel and are shown in units of g N m$^{-2}$ yr$^{-1}$ added to each pot. Low-, medium- and high-N treatments correspond approximately to soil N concentrations of 14, 46 and 78 mg N kg soil$^{-1}$, respectively, using conversions described in the Methods. Boxes represent the lower and upper quartiles with the median shown as the black central line. Whiskers indicate either the most extreme value or 1.5 times the interquartile range from the box. Sample sizes (number of individual plants) for each treatment across all three rows were (from left to right): 18, 17, 16, 17, 20, 19, 20, 19 and 19 plants (Supplementary Table 3). Mixing model end-member variation was incorporated into the statistics for **g–i** but is not shown here. Within each row, bars with different letters are statistically different as determined by our best-fit maximum likelihood model (Supplementary Table 1). **a–i**, $\Delta AIC_c$ values for our best-fit models compared with our next best model were 0.45 for plant biomass (**a–c**), 1.95 for allocation to nodules (**d–f**) and 1.98 for $\%N_{dfa}$ (**g–i**) (Supplementary Table 1).

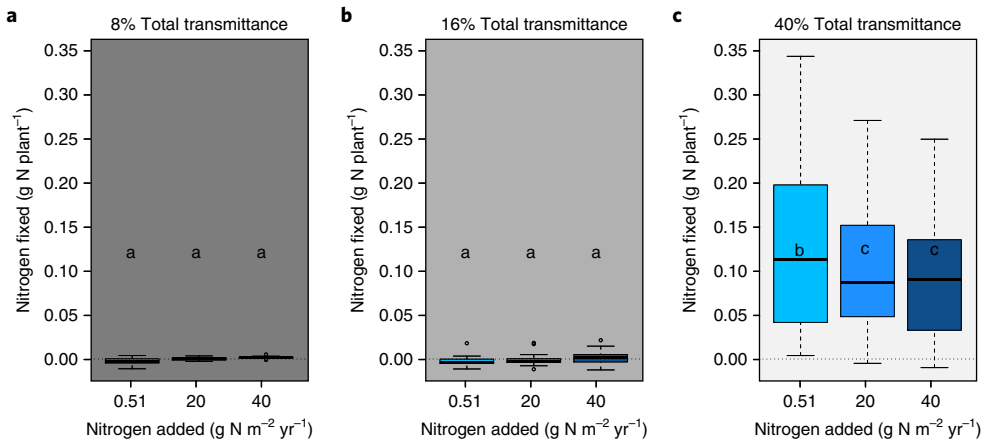


**Fig. 2 | Light drives total fixed N in plants more strongly than soil N. a–c**, The total amount of N fixed per plant for shadehouse plants grown in low- (**a**), medium- (**b**) and high-light (**c**) treatments. Low-, medium- and high-N treatments are represented by the three boxes in each panel, which correspond to soil N concentrations of 14, 46 and 78 mg N kg soil$^{-1}$ as described for Fig. 1. Treatment arrangements are as in Fig. 1. Boxes represent the lower and upper quartiles with the median shown as the black central line. Whiskers indicate either the most extreme value or 1.5 times the interquartile range from the box. Sample sizes (number of individual plants) for each treatment were (from left to right): 18, 17, 16, 17, 20, 19, 20, 19 and 19 plants (Supplementary Table 3). Bars with different letters are statistically different as determined by our best-fit maximum likelihood model (Supplementary Table 1). $\Delta AIC_c$ for our best-fit model compared with our next best model was 1.54 (Supplementary Table 3).

nearly all our replicates, as did 16% sunlight (although we did find some nodules in the lowest N treatment). However, soil N availability that demonstrably relieved N limitation (medium- and high-N treatments) did not fully inhibit SNF at 40% full sunlight. The response of fixation to light suggests that below some threshold of light availability, *P. macroloba* downregulates SNF completely because it is not N limited; that is, reduced light to fuel photosynthesis in our low- and medium-light treatments creates low N demand for these plants, which leads them to completely downregulate SNF. By contrast, the incomplete downregulation of SNF in response to N fertilization at high light demonstrates that, given sufficient light, *P. macroloba* continues to engage in SNF even when it is not N limited.

Environmental conditions, such as light and soil N availability, largely determine plant size (Fig. 1a–c), and plant size could influence nodulation independent of the direct effects of the environmental conditions. For example, plants growing in low-light environments may not have nodulated simply because small plant size inhibited nodulation rather than because light availability had a direct effect on SNF. To test for this possibility and assess the effects of light and soil N independent of plant size in our shadehouse experiment, we used a subset of our plants that were grown in different light and soil N treatments but attained similar final plant sizes (4–10 g plant biomass). With this subset of plants, an analysis of variance (ANOVA) model testing for main effects of light, soil N and plant size on allocation to nodules showed that light availability significantly increased nodulation ($P > 0.0001$), soil N availability significantly decreased nodulation ($P = 0.00127$), but that plant biomass had no effect on nodulation ($P = 0.362$) (Supplementary Fig. 1). These results demonstrate that light and soil N availability, independent of plant size, primarily determined nodulation in our plants.

Because plants reared in shadehouse conditions are only a limited reflection of nature, we sought to verify the patterns in our experimental results with forest seedlings in situ. We measured the percentage of belowground biomass allocated to nodules (to control for variation in plant size) on 100 *P. macroloba* seedlings growing across gradients of light and soil N conditions in rainforest understories. We found that nodulation was significantly positively correlated with light availability ($\Delta AIC_c > 1.63$; Fig. 3a) but was not correlated with soil N ($\Delta AIC_c > 1.62$; Fig. 3b). Our models indicated that the positive effect of light on nodulation was driven by an increase in the probability that a plant nodulated with increasing light availability, rather than an increase in nodule mass within plants that nodulated. Due to differences in other potential SNF regulators (for example, non-N soil resources, plant competition or herbivory) between our shadehouse and field studies, direct quantitative comparisons of our light and soil N gradients between these two studies are difficult. Still, these in situ data lend support to our shadehouse results that light is a stronger driver of SNF than soil N availability.

Three prior studies assessing the effects of light on nodulation in natural settings found complete or near-complete downregulation of allocation to nodules in low-light conditions, just as we did[19–21]. Although these prior studies did not directly measure SNF rates, our current understanding of SNF suggests that complete downregulation of nodulation fully inhibits SNF activity. These studies, along with work on agricultural N fixers (for example, Murphy[16], MacDowall[17] and Lau[18]) and the results presented here, cover a broad spectrum of woody and herbaceous N-fixing taxa from multiple legume subfamilies occurring across a range of biomes. All of these studies show strong effects of light on SNF regulation, which suggests that our findings are probably widely applicable, but additional studies on other N fixers in other locations are needed to fully verify these results. Evidence for how strongly soil N regulates SNF is more mixed. Several studies found incomplete downregulation of SNF in response to N fertilizer[8,19,20,26], similar to our high-light treatment, whereas others report that N additions completely downregulated SNF[5,8,19,20], similar to our medium- and low-light treatments. Our data suggest that one possible explanation for the discrepancy between these studies may be differences in light availability under different experimental conditions. The ability of light to mediate SNF responses to soil N (and potentially other factors such as phosphorus) may help us develop our emerging understanding of why and how some species match SNF closely with N demand, whereas others are either over-regulators or under-regulators[5].

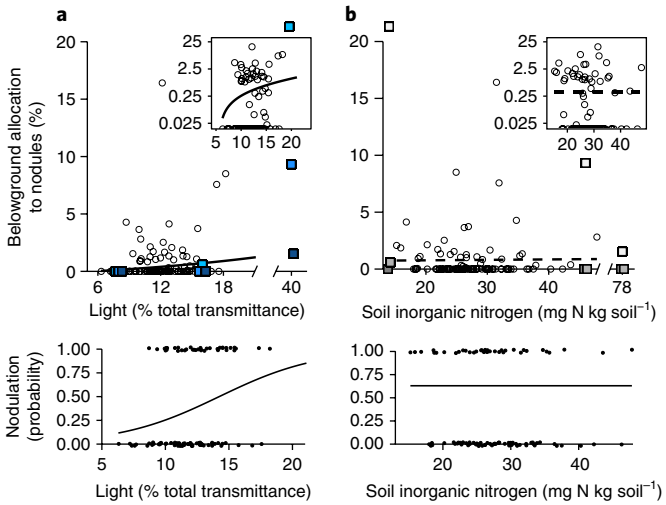


**Fig. 3 | In situ field nodulation varies with light, but not soil N. a**, Nodule biomass (linear-scaled in the main plot and log-scaled in inset) and the probability of a plant nodulating increased significantly with seedling light availability in 100 seedlings growing in the rainforest understory. **b**, However, nodule biomass and the probability of a plant nodulating did not vary considerably with soil inorganic N ($NO_3^- + NH_4^+$) availability in field-sampled seedlings. Even though log-scaled axes do not have a 'zero' value, individual plants with 0% allocation to nodules are displayed as points on the *x* axis of the log-scaled insets. Coloured squares in each panel represent median nodulation from the nine treatments of our shadehouse study (Fig. 1d–f) for comparison. **a**, Squares of light, medium and dark blue correspond to low, medium and high N additions in the shadehouse for each shadehouse light group. **b**, Squares of dark, medium and light grey correspond to low, medium, and high light for each N shadehouse N treatment.

Our results can also help refine SNF regulation theory. Current theory[5] envisions that different plant species vary in how they regulate SNF in response to limitation by N and another resource (for example, light[10], soil phosphorus[5,15] or an undefined density-dependent resource[27,28]). Perfectly facultative N fixers decrease from relatively high SNF under conditions of N limitation to zero SNF under conditions of limitation by the other resource. Obligate N fixers maintain similar SNF rates per unit biomass regardless of which resource limits them and incomplete downregulators are in between[5]. Our data demonstrate that species' SNF strategies, rather than simply being traits intrinsic to the taxa, vary as functions of light availability. *P. macroloba* is perfectly facultative at low light, but an incomplete downregulator at high light (Figs. 1g–i, 2 and 4a). Incorporating light regulation of SNF strategies may inform aspects of SNF theory ranging from the ability of individual plants to regulate SNF[5] to the effects that SNF regulation has on global patterns of N cycling[15,28].

In particular, SNF has for almost a decade been the focus of an apparent paradox in the biogeochemistry literature—that of tropical forest N richness. Many tropical forests export large

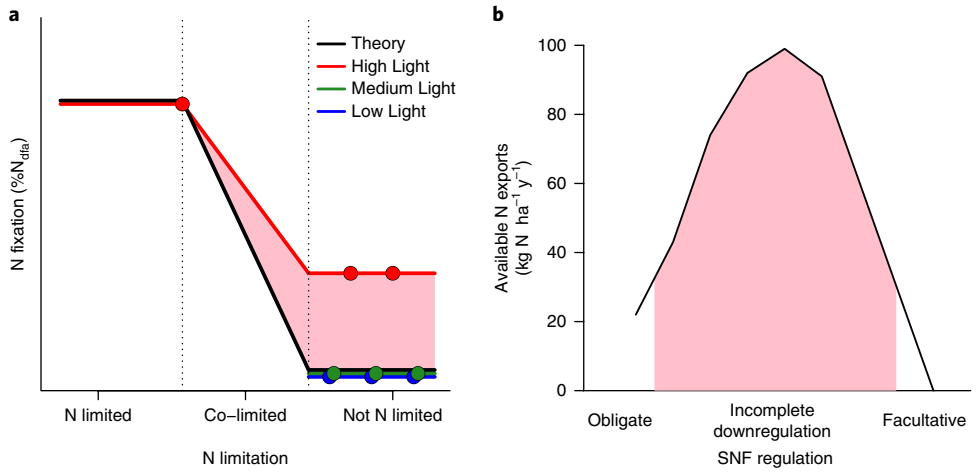


**Fig. 4 | Our results in the context of SNF theory. a**, Theory (black) predicts that facultative SNF will respond to limitation by N and another resource in equal but opposite directions. By contrast, our data suggest that SNF responds to N limitation differently depending on light availability. When plants in our experiment were not N limited, they continued to fix N in high light (red), but not in low (blue) or medium light (green), possibly because they lacked the carbon resources to engage in SNF. Points represent results from our nine shadehouse treatments plotted along an axis of N limitation (environmental N supply relative to plant N demand, as inferred by biomass responses to elevated resource levels). The pink shaded region shows conditions where incomplete downregulation of SNF causes our plants to 'over fix' theoretical predictions. **b**, Incomplete downregulation of SNF can lead to large ecosystem N exports (adapted with permission from Menge et al.[5]). In the model simulations of Menge et al.[5], SNF regulation that is perfectly facultative in response to N limitation does not create ecosystem N exports. Obligate SNF (no regulation) leads to substantial ecosystem N exports, but not as large as incomplete downregulation because obligate N fixers are outcompeted in N-saturated environments. Incomplete downregulation (represented by the corresponding pink areas in **a** and **b**, which correspond to our results from high light) can lead to the largest ecosystem N exports because N fixers can persist in N-saturated conditions but continue to engage in some SNF even when N is not limiting.

amounts of bio-available N, which budgets suggest come largely from overactive SNF[29]. Yet intuition, theory and some experimental results[8,19,20] have suggested that N fixers in tropical forests downregulate SNF in a perfectly facultative manner (black line, Fig. 4), which should minimize exports of bio-available N[5,29]. Large N exports from tropical forests are only paradoxical if SNF shuts off completely when N limitation has been overcome[5]. Thus, 'over-fixation' driven by high light would resolve the paradox of over-abundant N in tropical forests. In our shadehouse experiment, plants that were demonstrably N saturated continued to fix N at ~80% the rate of N-limited plants when given ample light availability (Fig. 2c). Even small amounts of SNF that continue after N limitation has been relieved (pink areas in Fig. 4) can lead to large bio-available N export[5] (Fig. 4b), and our results suggest that SNF may be substantial in N-saturated environments when there is sufficient light. Given that tropical forests receive ample solar energy[30], light-driven SNF rates well exceeding plant demand could be common, which would lead to large exports of bio-available N[5] and would, therefore, resolve the paradox. This paradox is just one important example of how a stronger consideration of light as the primary regulator of SNF can improve our understanding of N inputs into the biosphere.

Here we provide clear evidence that light can be a strong and absolute (has the capacity to completely inhibit) driver of SNF, and can mediate the responses of SNF to soil N. This suggests that SNF research should shift from past decades' focus on soil N to looking at other factors, such as light, as the dominant regulators of N inputs into the biosphere. The taxonomic diversity and geographic extent of symbiotic N fixers imply that a variety of environmental factors may play a role in regulating SNF, but our results suggest that the strongest regulator in some ecosystems is not soil N. Given the magnitude of the responses seen in this study, we suggest that regulation by light be a primary consideration as we continue to improve our understanding of the role that SNF plays in global N and C cycling.

## Methods

**Study site.** We conducted both the open shadehouse experiment and the forest seedling sampling at La Selva Biological Station (10° 25′ 53.14″ N, 84° 0′ 10.51″ W) in the premontane wet forests of Heredia province in Costa Rica. This site experiences an average daytime temperature of 25 °C, which is relatively constant throughout the year. Annual precipitation at La Selva is about 4,500 mm yr$^{-1}$ with a pronounced dry season occurring from January through April and a second, less pronounced dry season in September and October. Soils at this site are primarily ultisols derived from weathered basalt[31].

**Study species.** To test for the effects of light and soil N on SNF, we used the native N fixer *P. macroloba* (Willd.) Kuntze. *P. macroloba* is common in lowland forests throughout its range from Nicaragua to the Amazon basin[32] and it is the most abundant canopy species at this site[33]. Adults of *P. macroloba* typically reach 30–35 m in height and up to 130 cm in diameter, and produce dry, dehiscent fruits that ballistically disperse three to eight seeds, each weighing about 8 g. Seed germination occurs within 8–10 days[34]. *P. macroloba* is often considered a shade-tolerant species because it can recruit under mature-forest canopies, but it also thrives in high-light riparian environments. This species can tolerate ranges of light availability from 1% to 100% full sunlight, but its growth is sensitive to changing light environments[22]. Prior studies have also shown that levels of N fixation in *P. macroloba* vary in response to soil N availability[26].

**Shadehouse experiment.** *Experimental design.* We measured the effects of light and soil N availability on *P. macroloba* by comparing the growth of plants in 7-litre pots exposed to varying levels of each environmental variable between July 2015 and January 2016. Seedlings were grown from seed in soil inoculated with *P. macroloba*'s N-fixing rhizobia endosymbionts using 4 ml of slurry containing distilled water and locally collected active *P. macroloba* root nodules. High-, medium- and low-light treatments were created using varying thicknesses of shade cloth. The light treatments corresponded to 40%, 16%, and 8% full irradiance, respectively, which we determined by comparing paired sensor readings for photosynthetically active radiation where one sensor was placed above plants in each light treatment and the other sensor was placed in an open field exposed to full sunlight. Within each of these light treatments were three N fertilization treatments. Pots were filled with a mixture of forest soil and locally sourced sand in a 1/1 ratio. Pots were then fertilized using an ammonium-nitrate ($NH_4$-$NO_3$) solution applied once at the time of seed sowing and once at the mid-point of the experiment's duration in amounts equivalent to 0.51, 20 and 40 g N m$^{-2}$ yr$^{-1}$ for low-, medium- and high-N treatments, respectively. Included in the fertilizer was a 98% isotopically enriched $^{15}$N tracer used to determine N fixation (see later). We converted these N additions to concentrations of total inorganic N in each pot

using the mean soil N concentration of forest soil from our field sampling (see later) for the forest soil fraction of our potting mix and assuming 7,000 $cm^3$ of soil with a 1.4 $g\,cm^{-3}$ bulk density[35]. N treatments were arranged randomly within each light treatment. Concurrent with N fertilizations (at the start and mid-point of the growth period), all plants received 2 ml of N-free Hoagland's solution containing a mixture of inorganic phosphorus and micronutrients. This represents a fully replicated factorial design with nine environmental treatments, each containing 20 plants at the start of the experiment.

*Seedling growth and data collection.* Plants were grown from locally collected seeds planted in late June 2015. Initial plant (seed) size was determined by weighing seeds before planting. After planting, all visibly dead or inviable seeds were replaced after 2 weeks. Plants were grown for a period of 6 months during which they were exposed to ambient fluctuations in temperature and relative humidity, but were provided ample water via individual watering pans placed beneath each pot. About 94.4% of seeds germinated and survived through the experiment's duration. Sample sizes of surviving plants in each treatment, which were used in all statistical tests for our shadehouse experiment, are available in Supplementary Table 3.

After the 6-month growing period, plants were harvested by removing them from the potting soil and rinsing excess soil from root surfaces with distilled water. Each plant was then dissected into root, nodule, stem and leaf fractions. Each tissue component was dried at 60 °C to constant mass (about 3 days) and massed.

*Calculating $\%N_{dfa}$.* We calculated the $\%N_{dfa}$ for each individual using a three-end-member mixing model: one end member for the isotopic signature of soil-derived N, one for the isotopic signature of atmospherically fixed N and one for the fraction of plant N derived from the seed. Our calculation accounted for variation around each end member. Because we added a highly isotopically enriched $^{15}N$ tracer to the growing medium, we present $^{15}N$ data as atom percent, representing the percentage of N atoms in a sample that are $^{15}N$ rather than $^{14}N$ (as opposed to the ‰ notation that is common for natural abundance levels of isotopes).

We calculated the isotopic signature of the soil N end member by estimating N that naturally mineralized in the soil used in our potting mixture, any non-isotopically enriched N fertilizer added to the pot, and the isotopically enriched N fertilizer added. Although the soil N end member is often estimated using a reference plant[36], some of the known problems with the reference plant method[36] made it unviable in our study. Specifically, substantial differences in the root distributions and particularly seed size between our study plants and reference plants, along with relatively well-known amounts and isotopic signatures of N additions to each pot, meant that estimating the isotopic signature of the soil N end member directly gave a better measure than using a reference plant. Mineralized and asymbiotically fixed N was estimated for the amount of forest-derived soil added to each pot using the mean and s.d. of N mineralization rates from Brookshire et al.[29] for neotropical forests and asymbiotic N fixation rates from Reed et al.[37] to randomly generate a series of 10,000 normally distributed N mineralization and asymbiotic N fixation values for each pot. Sensitivity analyses using ±50% of these literature values showed no greater than a 2.5 $\%N_{dfa}$ change in any treatment, suggesting our results were insensitive to variation in these values. Each pot then received a total of 0.51 $g\,N\,m^{-2}\,yr^{-1}$ of 98% $^{15}N$ tracer. The amount of non-isotopically enriched fertilizer varied by N treatment: 0, 19.49 and 39.49 $g\,N\,m^{-2}\,yr^{-1}$ for low, medium and high N, respectively. The isotopic signature of soil N ($\%^{15}N_{soil}$) from each pot was then calculated as follows:

$$\%^{15}N_{soil} = ((0.003663 \times f_{min}) + (0.98 \times f_{iso}) + (0.003663 \times f_{fert})) \times 100$$

where $f_{min}$, $f_{iso}$ and $f_{fert}$ are the fraction of soil N coming from mineralization, isotopically enriched fertilizer and non-isotopically enriched fertilizer. This resulted in a distribution of 10,000 values of $\%^{15}N_{soil}$ for each pot to account for uncertainty in soil N mineralization rates.

We used an isotopic signature of 0.3663 atom percent for the atmospherically fixed N end member. Variation for this end member is likely to be very small on the scale of enriched isotopes[5]. Therefore, for each pot we randomly generated 10,000 $\%^{15}N_{fixation}$ values normally distributed with a mean of 0.3663 and an s.d. of 0.01 (equivalent to 27.5‰ in $\delta^{15}N$ notation, which is much larger than the actual variation in $\delta^{15}N$ of atmospherically fixed N). Our results were insensitive to variation in s.d. values around this end member ranging from 0.005 to 0.05 (13.8‰ to 137.2‰ in $\delta^{15}N$ notation).

*Fraction of plant N derived from seed.* The large seeds of *P. macroloba* (mean = 6.15 g for plants used in this study) mean that a substantial amount of N within our study plants came from seed reserves, especially for the plants grown in low- and medium-light conditions where final total plant biomass was often similar to the original seed biomass (explained in detail later). To estimate the proportion of a plant's N derived from the seed (Supplementary Table 2), we used a two-step process. To calculate the original seed dry mass, we used a conversion factor for wet mass to dry mass of seeds, which we derived from a linear model of wet and dry seed masses of 36 *P. macroloba* seeds collected along with the seeds used for our study plants. However, the entire mass of a seed is not incorporated into a plant, so we then subtracted out the amount of seed mass that was not used by the growing seedling. To do this, we collected, dried and massed the seed material for each plant remaining in the pot at the end of the experiment.

We then subtracted this unused fraction of the seed mass from the total dry seed mass and multiplied this potentially used seed mass by the N concentration of *P. macroloba* seeds (3.359% obtained by elemental analysis of 36 seeds collected at the same time as seeds used in our study) to estimate the amount of seed N that was potentially used by each plant. We divided this seed-derived N by the total N contained in each plant to calculate the fraction of seed-derived N in the plant ($f_{seed}$). For several of our smallest plants, this fraction of potentially seed-derived N exceeded 1, which is impossible (a plant cannot get >100% of its N from seed). The minimum $\%^{15}N$ value for any plant in our study was 0.383% (well more than the natural abundance range; equivalent to $\delta^{15}N = 46$‰), confirming that even our smallest plants did not derive all of their N from seed. For plants with an estimated $f_{seed} > 1$, we assumed that the majority of plant N was, in fact, derived from the seed, and thus we assigned these plants a value of 0.9 for $f_{seed}$. Sensitivity analyses showed that our results were qualitatively unaffected by varying this assigned value for $f_{seed}$ between .83 and .999. We used 0.3663 atm % for the isotopic signature of seed-derived N ($\%^{15}N_{seed}$) because the seeds were not isotopically labelled.

*$\%N_{dfa}$ calculation.* After generating distributions for each of the three end members ($\%^{15}N_{soil}$, $\%^{15}N_{fixation}$ and $f_{seed}$), we incorporated these end members into a mixing model to calculate the $\%N_{dfa}$ as follows:

$$\%N_{dfa} = \left(\frac{\%^{15}N_{samp} - (f_{seed} \times \%^{15}N_{seed}) - ((1 - f_{seed}) \times \%^{15}N_{soil})}{(\%^{15}N_{fixation} - \%^{15}N_{soil})}\right) \times 100 \quad (1)$$

where $\%^{15}N_{samp}$ is the isotopic signature of the *P. macroloba* seedling sample, $f_{seed}$ is the fraction of plant N derived from seed, $\%^{15}N_{seed}$ is the isotopic signature of N derived from seed, $\%^{15}N_{soil}$ is the isotopic signature of the N derived from the soil and $\%^{15}N_{fixation}$ is the isotopic signature of N derived from fixation. This calculation differs slightly from the $\%N_{dfa}$ equation used in Menge et al.[5] in that they defined $\%N_{dfa}$ as the percentage of a plant's newly acquired N (postgermination), whereas here we define $\%N_{dfa}$ as the percentage of the plant's total N (including N derived from the seed). Because $f_{seed}$, $\%^{15}N_{soil}$ and $\%^{15}N_{fixation}$ were distributions of 10,000 values each, this calculation produced 10,000 values of $\%N_{dfa}$ for each plant. Mean and 95% confidence intervals were calculated from this distribution of $\%N_{dfa}$ for each plant to provide our estimate of N derived from fixation and the uncertainty around this estimate.

*Calculating total N fixed per plant.* To determine the total amount of N in each plant that was derived from N fixation, we multiplied the total amount of N in each plant (g N) by the $\%N_{dfa}$ for each plant (described earlier). N content of each plant was calculated separately for each tissue pool. We multiplied the dry mass of each tissue pool (leaves, stems, roots and nodules) by the percentage of N for that tissue pool reported for *P. macroloba* in Russell and Raich[38] to obtain total N content (g N) for each tissue pool for each plant. N content was then summed across tissue pools to calculate the total N content (g N) for the plant.

*Statistical analyses for shadehouse experiment.* Although our shadehouse experiment was structured as an ANOVA design, it is not possible to incorporate mixing model end-member variation into a standard ANOVA. We therefore used maximum likelihood models and information theory-based model comparison to achieve the same end as an ANOVA—testing for differences between treatment means—while staying true to the error structure of our data. Specifically, to incorporate error propagation from the three-end-member mixing model calculation of $\%N_{dfa}$ into our maximum likelihood models, we had the maximum likelihood model estimate the mean and s.d. for $\%N_{dfa}$ in each treatment directly from the $\%^{15}N$ values for each end member of the mixing model[5]. This resulted in a more robust assessment of the uncertainty around $\%N_{dfa}$ estimates and a more conservative assignment of treatment differences than a standard ANOVA test that cannot incorporate end-member variation[39]. Furthermore, model comparison allows us to test only the treatment differences that are biologically reasonable or of particular interest to our study (rather than all pairwise comparisons), vastly decreasing the number of possible model tests and avoiding the need for corrected *post hoc* comparisons.

The effects of light and soil N on each response variable were assessed using a series of at least six maximum likelihood models: (1) a null model fitting a single mean to all treatments; (2) a model fitting different means for each light treatment; (3) a model fitting different means for each soil N treatment; (4) a high-light model that fit a single mean for the low- and medium-light treatments and a different mean for high-light treatments; (5) a variation-in-high-light model that fit a single mean to low- and medium-light treatments and an individual mean for each high-light treatment; and (6) an individual-treatment model that fit means for each of the nine experimental treatments. Additional models were tested on the basis of variation between treatments and specific scientific questions

for each response variable (Supplementary Table 1). When multiple treatment means for a response variable were not significantly different from zero in our individual-treatment model (such as N fixed per plant in our low- and medium-light treatments), no further models were used to assess differences between these means. We determined the best model for our data by comparing $\Delta AIC_c$ values[40] and interpreted one model as being considerably better than another model if $\Delta AIC_c > 2$. We report $\Delta AIC_c$ for our best model compared with the next best model and also compared with our model fitting different means for each N treatment (to compare our results with the expected results if soil N were the primary driver of SNF).

**Field sampling.** During the summer of 2017 we sampled 100 *P. macroloba* seedlings growing in natural forest understories exposed to varying levels of light and soil N availability. We sampled 20 seedlings from five sites each ranging in stand age from 20 years since abandonment to old-growth forest: each site was adjacent to plots used in the Bosques long-term forest dynamics project[41], where we have studied the dynamics of adult *P. macroloba*[42,43]. Seedlings 30–200 cm in height were selected to obtain variation in light availability (it was not possible to assess soil N availability before sampling).

A hemispherical photograph was taken directly above the tallest leaf of each seedling to assess light availability. Each photograph was analysed using Gap Light Analyser software (Cary Institute) for the % total light transmittance. Following photography, each seedling was extracted from the soil taking care to ensure that roots and nodules were not dislocated from the seedling during soil extraction. In cases where we thought it possible that some roots or nodules were dislocated from the plant, this was noted, but no differences were found between analyses conducted with and without these potentially broken plants in the dataset. To measure soil N availability, we sampled ~5 g of soil directly from the rooting zone of each seedling. Soil samples were extracted in 2 M KCl and analysed for nitrate and ammonium on a Smartchem 170 discrete analyser (Westco Scientific Instruments) at Columbia University. Following harvesting, we cleaned the root system of each seedling and removed all nodule material using forceps. Root and nodule material were dried and massed separately for each plant to calculate the % of belowground biomass that each plant allocated to nodule biomass.

*Statistical analyses for field sampling.* Because field nodulation data typically contain many zeros and values > 0 are often log-normally distributed, we analysed nodulation in field-sampled seedlings using models for zero-inflated log-normal data adapted from Tian and Wu[44], which were then evaluated using a maximum likelihood framework. This method predicts nodulation using a dual-process model where the probability of encountering a zero (nodule presence versus absence) and the mean of non-zero data (nodule mass when present) are modelled simultaneously. This takes the form of:

$$G(x,\mu,\sigma,\delta) = \begin{cases} \delta & \text{if } x = 0 \\ \delta + (1-\delta)\ F(x,\mu,\sigma) & \text{if } x > 0 \end{cases} \quad (2)$$

where $x$ is the log-normal response variable (% belowground allocation to nodules), $\mu$ and $\sigma$ are the mean and s.d. of the response variable in log space, and $\delta$ is the probability of encountering a zero value. $F(x,\mu,\sigma)$ is the log-normal cumulative distribution function of non-zero values. $\mu$ and $\delta$ could either be single values (for our null models) or vary in response to an independent variable such as light or soil N availability. This allowed us to calculate the predicted population geometric mean, $M$, as:

$$M = (1-\delta) \times e^{(\mu)} \quad (3)$$

where $\mu$ is the mean of the non-zero data in log space.

For each predictor variable—percentage total light transmittance and soil N availability (ammonium + nitrate concentrations)—separate models were created in which nodulation did not vary with the predictor variable (null model), only the probability of nodulation varied with the predictor variable, only the mean value of non-zero data varied with the predictor variable and where both the probability of nodulation and the mean of non-zero values varied according to the predictor variable. We tested these four model types for each predictor variable using the bbmle package for maximum likelihood tests in R statistical software[45,46]. We assessed differences between our models using the $\Delta AIC_c$[40] and present the difference in $\Delta AIC_c$ between our best and next best-fit models.

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

**Data availability.** Data for shadehouse and field-sampled seedlings used for these analyses are publicly available via the Dryad Digital Repository at https://doi.org/10.5061/dryad.p9f5160. All code used for data preparation and analyses is available via Github at https://github.com/bentonneiltaylor/Taylor-Menge-2018-N-Fixation-Code.

Received: 8 February 2018; Accepted: 23 July 2018;
Published online: 20 August 2018

## References

1. LeBauer, D. & Treseder, K. Nitrogen limitation of net primary productivity in terrestrial ecosystems is globally distributed. *Ecology* **89**, 371–379 (2008).
2. Elser, J. J. et al. Global analysis of nitrogen and phosphorus limitation of primary producers in freshwater, marine and terrestrial ecosystems. *Ecol. Lett.* **10**, 1135–1142 (2007).
3. Chapin, F. III, Matson, P. & Vitousek, P. *Principles of Terrestrial Ecosystem Ecology* (Springer, New York, NY, 2011).
4. Vitousek, P. & Howarth, R. Nitrogen limitation on land and in the sea: how can it occur? *Biogeochemistry* **13**, 87–115 (1991).
5. Menge, D. N. L., Wolf, A. A. & Funk, J. L. Diversity of nitrogen fixation strategies in Mediterranean legumes. *Nat. Plants* **1**, 15064 (2015).
6. Barron, A. R., Purves, D. W. & Hedin, L. O. Facultative nitrogen fixation by canopy legumes in a lowland tropical forest. *Oecologia* **165**, 511–520 (2011).
7. Batterman, S. A. et al. Key role of symbiotic dinitrogen fixation in tropical forest secondary succession. *Nature* **502**, 224–227 (2013).
8. Batterman, S. A., Wurzburger, N. & Hedin, L. O. Nitrogen and phosphorus interact to control tropical symbiotic $N_2$ fixation: a test in *Inga punctata*. *J. Ecol.* **101**, 1400–1408 (2013).
9. Gutschick, V. Evolved strategies in nitrogen acquisition by plants. *Am. Nat.* **118**, 607–637 (1981).
10. Rastetter, E., Vitousek, P. & Field, C. Resource optimization and symbiotic nitrogen fixation. *Ecosystems* **4**, 369–388 (2001).
11. Vitousek, P. & Field, C. Ecosystem constraints to symbiotic nitrogen fixers: a simple model and its implications. *Biogeochemistry* **46**, 179–202 (1999).
12. Hedin, L. O., Brookshire, E. N. J., Menge, D. N. L. & Barron, A. R. The nitrogen paradox in tropical forest ecosystems. *Annu. Rev. Ecol. Evol. Syst.* **40**, 613–635 (2009).
13. Crews, T. E. Phosphorus regulation of nitrogen fixation in a traditional Mexican agroecosystem. *Biogeochemistry* **21**, 141–166 (1993).
14. Wurzburger, N. & Miniat, C. F. Drought enhances symbiotic dinitrogen fixation and competitive ability of a temperate forest tree. *Oecologia* **174**, 1117–1126 (2014).
15. Houlton, B. Z., Wang, Y.-P., Vitousek, P. M. & Field, C. B. A unifying framework for dinitrogen fixation in the terrestrial biosphere. *Nature* **454**, 327–330 (2008).
16. Murphy, P. M. Effect of light and atmospheric carbon dioxide concentration on nitrogen fixation by herbage legumes. *Plant Soil* **95**, 399–409 (1986).
17. MacDowall, F. D. H. Effects of light intensity and $CO_2$ concentration on the kinetics of 1st month growth and nitrogen fixation of alfalfa. *Can. J. Bot.* **61**, 731–740 (1982).
18. Lau, J. A. et al. Direct and interactive effects of light and nutrients on the legume-rhizobia mutualism. *Acta Oecol.* **39**, 80–86 (2012).
19. McHargue, L. A. *Factors Affecting the Nodulation and Growth of Tropical Woody Legume Seedlings* (Florida International University, Miami, FL, 1999).
20. Myster, R. W. Light and nutrient effects on growth and allocation of *Inga vera* (Leguminosae), a successional tree of Puerto Rico. *Can. J. For. Res.* **36**, 1121–1128 (2006).
21. Sprent, J. I. Growth and nitrogen fixation in *Lupinus arboreus* as affected by shading and water supply. *New Phytol.* **72**, 1005–1022 (1973).
22. Oberbauer, S. F., & Strain, B. R. Effects of light regime on the growth and physiology of *Pentaclethra macroloba* (Mimosaceae) in Costa Rica. *J. Trop. Ecol.* **1**, 303–320 (1985).
23. Vitousek, P. M. & Matson, P. A. Nitrogen transformations in a range of tropical forest soils. *Soil Biol. Biochem.* **20**, 361–367 (1988).
24. Chou, C. B., Hedin, L. O. & Pacala, S. W. Functional groups, species and light interact with nutrient limitation during tropical rainforest sapling bottleneck. *J. Ecol.* **106**, 157–167 (2018).
25. Chazdon, R. L. & Pearcy, R. W. The importance of sunflecks for forest understory plants. *Bioscience* **41**, 760–766 (1991).
26. Pons, T. L., Perreijn, K., van Kessel, C. & Werger, M. J. A. Symbiotic nitrogen fixation in a tropical rainforest: $^{15}N$ natural abundance measurements supported by experimental isotopic enrichment. *New Phytol.* **173**, 154–167 (2007).
27. Menge, D. N. L., Levin, S. A. & Hedin, L. O. Facultative versus obligate nitrogen fixation strategies and their ecosystem consequences. *Am. Nat.* **174**, 465–477 (2009).
28. Menge, D. N. L. et al. Why are nitrogen-fixing trees rare at higher compared to lower latitudes? *Ecology* **98**, 3127–3140 (2017).
29. Brookshire, E. N. J., Gerber, S., Menge, D. N. L. & Hedin, L. O. Large losses of inorganic nitrogen from tropical rainforests suggest a lack of nitrogen limitation. *Ecol. Lett.* **15**, 9–16 (2012).
30. Liang, S. et al. *Global LAnd Surface Satellite (GLASS) Products* (Springer, New York, NY, 2014).
31. Sollins, P., Sancho, M. F., Mata, C. R. & Sanford, R. L. in *La Selva: Ecology and Natural History of a Neotropical Rain Forest* (ed McDade, L.) Ch. 4 (University of Chicago Press: Chicago, IL, 1994).

32. Flores, E. M. in *Tropical Tree Seed Manual: Agricultural Handbook* (Academia Nacional de Ciencias de Costa Rica, San Jose, 2002).
33. Lieberman, D. & Lieberman, M. Forest tree growth and dynamics at La Selva, Costa Rica (1969–1982). *J. Trop. Ecol.* **3**, 347–358 (1987).
34. Joker, D. & Salazar, R. *Pentaclethra macroloba*. *Seed Leafl.* **35**, 1–3 (2000).
35. Brady, N. C. & Weil, R. R. *The Nature and Properties of Soils* (Prentice Hall, Upper Saddle River, NJ, 2002).
36. Chalk, P. M. Estimation of $N_2$ fixation by isotope dilution: an appraisal of techniques involving $^{15}N$ enrichment and their application. *Soil Biol. Biochem.* **17**, 389–410 (1985).
37. Reed, S. C., Cleveland, C. C. & Townsend, A. R. Functional ecology of free-living nitrogen fixation: a contemporary perspective. *Annu. Rev. Ecol. Evol. Syst.* **42**, 489–512 (2011).
38. Russell, A. E. & Raich, J. W. Rapidly growing tropical trees mobilize remarkable amounts of nitrogen, in ways that differ surprisingly among species. *Proc. Natl Acad. Sci. USA* **109**, 10398–10402 (2012).
39. Phillips, D. & Gregg, J. Uncertainty in source partitioning using stable isotopes. *Oecologia* **127**, 171–179 (2001).
40. Anderson, D. R. *Model Based Inference in the Life Sciences: A Primer on Evidence* (Springer, New York, NY, 2008).
41. Chazdon, R. L. et al. Rates of change in tree communities of secondary Neotropical forests following major disturbances. *Philos. T. Roy. Soc. B* **362**, 273–289 (2007).
42. Menge, D. N. L. & Chazdon, R. L. Higher survival drives the success of nitrogen-fixing trees through succession in Costa Rican rainforests. *New Phytol.* **209**, 965–977 (2016).
43. Taylor,B. N., Chazdon, R. L., Bachelot, B. & Menge, D. N. L. Nitrogen-fixing trees inhibit growth of regenerating Costa Rican rainforests. *Proc. Natl Acad. Sci. USA* **114**, 8817–8822 (2017).
44. Tian, L. & Wu, J. Confidence intervals for the mean of lognormal data with excess zeros. *Biometrical J.* **48**, 149–156 (2006).
45. Bolker, B. M. & R Development Core Team. bbmle: tools for general maximum likelihood estimation. https://cran.r-project.org/package=bbmle (2017).
46. R Core Team. R: A language and environment for statistical computing. http://www.r-project.org/ (2017).


## Acknowledgements

The authors thank E. Salicetti, M. Wilcots, R. Li, S. Taylor, B. Scott, E. Utset, B. Matarrita, D. Madrigal and O. Vargas for help conducting the experiment, and collecting and processing samples. This work was supported by the Garden Club of America's Award in Tropical Botany, Columbia University's Earth Institute and the Institute for Latin American Studies.

## Author contributions

B.N.T. designed and implemented the study, analysed data and wrote the first draft. D.N.L.M. designed the study, analysed data and revised 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/s41477-018-0231-9.

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

**Correspondence and requests for materials** should be addressed to B.N.T.

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

nature research

Corresponding author(s): Benton N. Taylor

# 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<br>*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<br>*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<br>*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 open source software "Gap Light Analyzer" from the Cary Institute was used to collect light data from hemispherical photographs. This software is referenced in the methods section of the manuscript. |
| Data analysis | All statistical analyses were conducted using the open source R statistical software. Code is available on Github with a link available in the data availability statement of the manuscript. |

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

Data for shadehouse and field-sampled seedlings used for these analyses are publicly available via the Dryad Digital Repository at https://doi.org/10.5061/dryad.p9f5160. All code used for data preparation and analyses is available via Github at https://github.com/bentonneiltaylor/Taylor-Menge-2018-N-Fixation-Code.

# 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

# Ecological, evolutionary & environmental sciences study design

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

| | |
|---|---|
| Study description | Our manuscript describes two complementary studies. In our shadehouse study, we grew 180 P. macroloba seedlings under 9 different environmental treatments. Our design included 3 light treatments (8%, 16%, and 40% full sunlight) by 3 nitrogen addition treatments (0.51, 20, and 40 gN/m2/yr added) established in a full-factorial design with 20 plants grown in each of the 9 environmental treatments. After a 6-month growing period, we harvested plants and used isotopic analyses to determine plant growth and symbiotic nitrogen fixation for plants grown in each treatment. For our field-sampling study, we harvested 100 P. macroloba saplings (30cm - 200 cm in height) from 5 different forest understory sites (20 plants from each site). For each plant we assessed inorganic N availability using KCl extractions directly from the rooting zone of the plant, and we assessed light availability using hemispherical photographs taken directly above each plant prior to harvesting. We measure symbiotic N fixation activity on each field-collected seedling by measuring the percent of the plant's root tissue that was allocated to nodule biomass. |
| Research sample | Samples were individuals of the nitrogen-fixing tree Pentaclethra macroloba located at La Selva Biological Station in Northeast Costa Rica. This sample was chosen because P. macroloba is a common nitrogen fixing tree throughout much of Central and South America and is the most common tree at La Selva. While this sample population most closely describes others in the species P. macroloba, we have provided text in the manuscript demonstrating why the results of our study are likely widely generalizable to tropical nitrogenfixing trees. |
| Sampling strategy | Prior to establishing our shadehouse experiment, we ran power analyses allowing us to determine adequate sample sizes for a range of effect sizes and data variance. We chose starting sample sizes that were ~33% higher (to account for potential plant mortality during the experiment) than the results of our power analyses for our best estimates of effect size and variance. Mortality was well under 33% for all experimental treatments allowing us to maintain sufficient robust sample sizes. For our field sampling, we ran similar power analyses based on effect sizes and variance from our shadehouse data and determined final sample size to best satisfy these power analyses while minimizing destructive sampling. |
| Data collection | For our shadehouse experiment, plants were harvested following the experimental growth period by removing each plant from it's pot and washing the soil from the root system by hand. The sand/soil mix used in each pot allowed for easy removal of the root system from the soil without damaging any roots or nodules. Following plant harvest, all plants were separated into tissue pools (leaves, stems, roots, etc.) and dried to constant mass. Tissue was then massed (to obtain plant biomass and allocation to nodules) and data were recorded directly into an excel spreadsheet. Isotopic analyses on ground leaf tissue were done by an independent lab, which records the data directly from the analytical machine to a csv file, which was sent to BNT. For field sampling, root and nodule tissue for each plant was harvested and dried in the same fashion as the shadehouse plants. Dried tissue was massed and data were recorded directly into an Excel spreadsheet. Following soil KCl extraction, data for soil nitrate and ammonium were downloaded directly from the software for the Smartchem 170 discrete analyzer that we used to run those analyses. Hemispherical photos were taken by BNT to obtain data on light transmittance above each plant. Each photo was analyzed using Gap Light Analyzer software, and data were recorded directly to a .csv file. The lead author, BNT, was involved in all data collection. For collection of shadehouse data, a research assistant, Richard Li, aided in collecting data, but all data were recorded by BNT. For our field study, a research assistant, Ben Scott, aided in data collection, but all data were recorded by BNT. data were recorded by BNT. |
| Timing and spatial scale | Our shadehouse experiment was conducted from late June 2015 - January 2016. All plants were harvested within two weeks of each other in January 2016. Our field work was conducted over a period of 3 weeks in June 2017. All data were taken either at La Selva Biological Station or in surrounding forests within 20 km of La Selva Biological Station. |
| Data exclusions | No data were excluded from either the shadehouse experiment or field sampling sections of this study. |
| Reproducibility | The long duration of the experimental period (the plant growth period itself was 6 months) and the distance between our university |

| | |
|---|---|
| Reproducibility | and the experimental site precluded temporal replication of this experiment. Thus, our robust sample size for a tropical shadehouse study is the primary verification that the effects we see are reproducible. We have also provided sufficiently detailed methods in the main text and supplementary information to allow others to reproduce this study, which we look forward to. |
| Randomization | An equal number of plants were allocated to each of our 9 environmental treatments (3 light x 3 N treatments) in our shadehouse experiment. Plants were randomly assigned to one of the 3 N treatments within each light treatment ensuring that N treatments were arranged randomly within each light treatment. We also tested for effects of position within each light treatment on our response variables and found none. |
| Blinding | During post-harvest data collection (e.g. massing, KCl extraction analysis and isotopic analysis), all investigators were blind to treatment assignment. |

Did the study involve field work? ☒ Yes ☐ No

## Field work, collection and transport

| | |
|---|---|
| Field conditions | All field data were collected under similar conditions. Daily temperatures were around 25 degrees celcius. Weather conditions were always cloudy but without rain, as those are the conditions for the highest accuracy when taking hemispherical photos. |
| Location | All samples were taken at La Selva Biological Station (10° 25′ 53.14″ N, 84° 0′ 10.51″ W). Soil samples were taken at an approximate depth of 10 cm directly in the rooting zone of the freshly harvested P. macroloba saplings. |
| Access and import/export | La Selva Biological Station is a research forest run by the Organization for Tropical Studies. All necessary sampling, research, and export permits were obtained from the Costa Rican Ministerio de Ambiente y Energia (MINAE). |
| Disturbance | Only small holes were disturbed for the excavation of each individual sapling. Once the sapling had been harvested, the soil and litter layer was replaced. |

# Reporting for specific materials, systems and methods

### Materials & experimental systems

| n/a | Involved in the study |
|---|---|
| ☒ | ☐ Unique biological materials |
| ☒ | ☐ Antibodies |
| ☒ | ☐ Eukaryotic cell lines |
| ☒ | ☐ Palaeontology |
| ☒ | ☐ Animals and other organisms |
| ☒ | ☐ Human research participants |

### Methods

| n/a | Involved in the study |
|---|---|
| ☒ | ☐ ChIP-seq |
| ☒ | ☐ Flow cytometry |
| ☒ | ☐ MRI-based neuroimaging |