bioRxiv preprint doi: https://doi.org/10.1101/2020.09.29.318550; this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

PISUM SATIVUM HAS NO COMPETITIVE RESPONSES TO NEIGHBOURS: A CASE STUDY IN REPRODUCIBLE ECOLOGY

Mariah L Mobley[1,2], Audrey S Kruse[3], and Gordon G McNickle[1,2,*]

[1]Department of Botany and Plant Pathology, Purdue University, 915 W. State Street, West Lafayette, IN 47907, USA States

[2]Purdue Center for Plant Biology, Purdue University, 915W. State Street, West Lafayette, IN, 47907, USA.

[3]Department of Agronomy, Purdue University, 915 W. State Street, West Lafayette, IN 47907, USA States

* corresponding author: gmcnickle@purdue.edu, phone: 765-494-4645

KEYWORDS: Pisum sativum, tragedy of the commons, Ideal free distribution, plant-plant competition, replication crisis, below ground competition, meta-analysis.

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.29.318550; this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

ABSTRACT:

1. Many fields of science have experienced a replication crisis, where results from experiments with low statistical power published in the literature cannot be replicated. Ecology so far has not been drawn into this crisis, but there is no reason to think that this problem is absent in our field. Here, we originally attempted to replicate findings that showed pea (*Pisum sativum* L.) roots had strong differences in growth in the presence or absence of neighbours. Our original goal was just to develop a simple model system for studying how plant roots respond to competition from neighbours.
2. In an attempt to replicate previous findings, we performed four separate experiments with 480 individual plants, across three years. Each time plants were grown in the full factorial combination of above and belowground competition. In addition, pea has been studied in similar experiments across six additional studies. Thus, we used meta-analysis to combine previous findings with our new findings.
3. We were unable to replicate previous findings, and in all four experiments plants grew the same whether there were neighbours or not. Despite variability in individual studies, meta-analysis revealed that pea has no growth responses to neighbours and grows the same whether there is or is not below ground competition.
4. *Synthesis*: Many other fields have gradually been drawn into a growing replication crisis, that is thought to be the result of low statistical power. Even though this is just one case study where a somewhat controversial result could not be reproduced, there is no reason to think ecology is immune from the replication crisis. We suggest that solutions developed in other fields might pre-emptively ward off similar problems. These include stricter cut-offs for statistical significance, a growing use of large replicated studies, and considering avenues for pre-registration of methods.

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.29.318550; this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

INTRODUCTION:

Many scientific fields have been experiencing a so-called replication crisis (Ioannidis, 2005, Saltelli and Funtowicz, 2017). The crisis is largely thought to stem from a plethora of statistically significant results in studies with low statistical power, that cannot be repeated in follow-up studies and thus represent type I errors (i.e. a failure to accept a true null hypothesis; Moonesinghe et al., 2007, Colquhoun, 2017). Does ecology suffer from a replication crisis? As far as we can tell, there has not been a major event, paper or movement in the field of ecology to spark a replication crisis the way there has been in other fields such as psychology (Pashler and Wagenmakers, 2012), medicine (Ioannidis, 2005) and economics (Camerer et al., 2016). Here, we describe a series of experiments designed to explore how one species, *Pisum sativum* (L.) responds to below ground neighbour competition. Our original goal was not replication, but rather to simply develop this as a model system for plant responses to neighbours. We chose the species and experimental designs based on results in the literature. However, over four separate experiments, we were unable to replicate published findings. Here, we reluctantly present this as a case study in reproducible ecology. However, first, we give a brief overview of the theory and hypotheses associated with plant root responses to neighbours that motivated our experiments.

One of the earliest and better known hypotheses about how roots should grow in the presence of neighbours was that roots should avoid neighbours and segregate in soil (Litav and Harper, 1967, Parrish and Bazzaz, 1976, Schenk et al., 1999). This hypothesis was mostly the result of verbal models that sought to explain high species diversity rather than theory that sought an understanding of how evolution by natural selection might shape a plant's response. Nevertheless, these verbal root segregation models created a two-part testable hypothesis: (i) plants alone simply respond to nutrient and water availability and produce the amount of roots required to adequately supply the shoot with

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.29.318550; this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

resources and; (ii) plants growing with neighbours are predicted to produce fewer roots (i.e. under-proliferate) in an attempt to seek root-free soil, and cede soil resources to competitors. The null hypothesis would be that plants grow the same, whether there are neighbours or not. Importantly, the segregation response requires some as yet unidentified mechanism to sense neighbour roots as different from self-roots (Falik et al., 2003, Gruntman and Novoplansky, 2004). A number of experiments have shown statistical evidence for this under-proliferation behaviour in plants (Schenk et al., 1999, Cahill et al., 2010), that is, these experiments have rejected the null hypothesis that plants have no response to neighbours and accepted the hypothesis that plants avoid neighbours.

Despite the longevity of the verbal root segregation hypothesis, and some empirical support, when evolutionary models were finally developed concerning best neighbour responses they almost all showed that segregation unlikely to be favoured by natural selection. Instead these models predicted the opposite of root segregation: plants should confront neighbours and try to pre-empt their nutrient uptake (Zhang et al., 1999, Gersani et al., 2001, Craine et al., 2005). To understand resource pre-emption, we suggest readers first consider a plant's best response to above ground competition: taller plants shade shorter plants, thereby pre-empting light access (Givnish, 1982, Falster and Westoby, 2003). Indeed, many plants that are shaded can sense being over-topped and will plastically elongate, or over-proliferate, their stems to attempt to avoid this resource pre-emption (Ballare et al., 1987, Murphy and Dudley, 2007). Resource pre-emption of a below ground resource is analogous - though below ground resources are not as easily pre-empted as light - resource pre-emption takes the form of over-proliferation of roots when plants grow among neighbours compared to when plants grow alone. These models created another testable two-part hypothesis: (i) remains the same as above, but; (iii) plants are hypothesized to produce more roots (i.e. over-proliferate) in order to pre-empt resources from a neighbour. Like the segregation hypothesis, the pre-emption hypotheses requires the ability to

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.29.318550; this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

recognize neighbour roots as different from self-roots. Again, a number of studies also show statistical support for the over-proliferation behaviour in plants (Padilla et al., 2013, Smyeka and Herben, 2017), that is, these studies rejected the null hypothesis that plants have no response to neighbours and accepted the alternative that plants over-proliferate roots.

However, while both over- and under-proliferation hypotheses have experimental support, some plant species do not seem to exhibit any response to below ground competition (Litav and Harper, 1967). That is, in these studies the null hypothesis of no response to neighbours was not rejected. A third set of null models have also been developed for such plants on the assumption that such plants may lack the ability to recognize neighbouring competitors (McNickle and Brown, 2014). Here, if plants have no responses to neighbours independent from nutrient depletion then the distribution of roots is predicted to be an ideal free distribution (IFD; sensu Fretwell and Lucas, 1969, Křivan et al., 2008). An IFD is often used to understand how non-territorial animals distribute themselves in relation to their food. Thus, despite the jargon, the IFD has a long history in foraging ecology, and connecting root growth to this long history is informative. The name IFD comes from the hypothesis that the distribution of foragers is ideal in the sense that individuals can 'know' resource distributions, and free in the sense that they are free to distribute themselves anywhere in relation to other individuals (e.g. no allelopathy or another plant behaviour akin to territoriality in animals; Schenk et al., 1999). An IFD type model explains the well-known behaviour of plants to place more roots in nutrient rich areas and fewer roots in nutrient poor areas (E.g. Hutchings and de Kroon, 1994, Hodge, 2004, Hodge, 2006), but predicts that nutrient depletion is the only mechanism of competition. Thus, rather than under- or over-proliferate roots in response to neighbours, an IFD type model creates a third two-part hypothesis: (i) remains the same as above, but; (iv) plants growing with neighbours should produce the same amount of roots as plants alone, providing the nutrient environment is equivalent on a per-plant basis. Again, there is also

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.29.318550; this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

experimental support for the IFD type root foraging model (McNickle and Brown, 2014, McNickle et al., 2020).

The three root competition models described above create three mutually exclusive alternative hypotheses. Importantly, these should be interpreted as mutually exclusive on the per-genotype basis, and not interpreted as mutually exclusive such that one model explains the entire plant kingdom. Indeed, it is entirely possible that different species use different responses, and there some evidence that this is the case (McNickle et al., 2016, Smyeka and Herben, 2017). Second, it also seems to be the case that different genotypes within a species might use different responses, and there is also evidence that this is the case (Murphy and Dudley, 2009, Semchenko et al., 2014). Third, perhaps responses are context dependent and the same genotype uses different responses across different contexts, and there is evidence that this too may be the case (Cahill et al., 2010, McNickle et al., 2016).

Many researchers (including us) seem to have converged on common pea (*Pisum sativum* L.) as a sort of model system for these root responses to neighbours. This is probably due to its ease of growth, and fast life cycle of 50-70 days. However, curiously, for this one species all three outcomes have been reported: One study found under-proliferation ((Chen et al., 2015) but see (McNickle, in press)); one study found over-proliferation ((O'Brien et al., 2005); but see (Laird and Aarssen, 2005, Hess and de Kroon, 2007)), and; (iii) two studies found the IFD no-response behaviour (Meier et al., 2013, Jacob et al., 2017). Given the mutually exclusive nature of the three theories, either pea responds differently across context, or genotype, or there may be seeds of a replication problem in these divergent results.

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.29.318550; this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Here, we report the results of four separate experiments that sought to replicate pea root responses to competitors. In each experiment, peas were grown in the full factorial combination of root and shoot competition, but in completely different contexts (e.g. potting media, nutrient availabilities, times of year). We were unable to replicate previous findings of under- or over-proliferation and found that in all four experiments, pea exhibited the null IFD response. We combined our new results with previously published data in a meta-analysis. When combined the results across the literature overwhelmingly support the null IFD response for pea. We conclude by returning to the idea that this case study might represent seeds of a reproducibility crisis, and describe some approaches used in other fields to find a path forward.

METHODS

Not all studies of pea have used the same genotype. However, one study that reported neighbour avoidance (Gersani et al., 1998), and one study that found over-proliferation (O'Brien et al., 2005) both used the 'Little Marvel' cultivar. Thus, in our attempt to replicate these results, we also used *P. sativum* c.v. Little Marvel in all experiments.

Peas were grown in four different experiments that varied in context (i.e. soil type, nutrient addition and the time of year), but included the same four neighbour treatments (Fig 1). Each time, soil type and nutrient delivery were adjusted in ways that we hypothesized would bring us closer to replicating previous methods, and therefore results. The neighbour treatments were: 1) aboveground neighbour only; 2) below ground neighbour only; 3) no neighbour, and; 4) both above- and below ground neighbour. All experiments were performed in the same greenhouse room, and on the same bench of the Purdue University Lilly greenhouse complex, in West Lafayette, Indiana, USA (40°42'26.0"N, 86°91'88.2"W) but took place over three years. The greenhouse was set to 25°C and supplementary

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.29.318550; this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

lighting from 600 watt high pressure sodium lights was on an 16:8 light:dark schedule for all four experiments. The same pot size was used in all experiments, with pots washed and sterilised between each experiment. In an attempt to minimize hypothesized volume constraints (Hess and de Kroon, 2007, Chen et al., 2015), we used very large 6.2 L pots designed for tree seedlings that were 40 cm deep and 15 cm square at the top (Pot TP616, Stuewe & Sons Inc, Tanget, Oregon, USA). These pots are deeper than pea above ground height, and are extremely large for such a small plant. The planting media, and fertigation varied across the four experiments as described below. Pea is a climbing vine, thus to hold plants upright, 92 cm tall bamboo stakes were added for each plant (i.e. two per pot).

Testing among the hypotheses of under-proliferation, IFD and over-proliferation has been difficult, and plagued with methodological difficulties related to the study of roots hidden in soil (Laird and Aarssen, 2005, Hess and de Kroon, 2007, Mommer et al., 2008, Chen et al., 2015). Unfortunately, these problems do not have easy solutions because of the limitations of geometry and chemistry (McNickle, in press). To impose the competition treatments, we used the barrier method and added dividers either above- or below ground (Figure 1). This has the problem of confounding neighbour addition and rooting soil volume. However, we deemed this to be a true replication of previous work since this was the most common methodological approach in previous studies, and it is one of the only approaches that controls soil nutrients across treatments. We direct readers to McNickle (in press) for guidance on alternative experimental approaches to controls and their interpretation. Above ground, opaque dividers made of white corrugated plastic that were 45 cm tall and 15 cm across with 4 cm flaps to attach to the pot were constructed and placed across the middle of each pot. All pots in our experiments received dividers as a control for their effect on light reduction, and the placement of plants was adjusted such that both plants were on the same side of the divider if interaction above ground was to be permitted, or plants were on opposite sides of the divider if above ground interaction

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.29.318550; this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

was to be blocked (Figure 1). To minimize the effects of shading caused by the above ground dividers, and interaction among pots, replicate blocks were widely spaced on the greenhouse bench (~1m apart), and all pots were turned one-quarter turn to the east every day of the experiment. Below ground, dividers were constructed by cutting the rectangular pots in half, and nesting the two halves together (Figure 1). This created a situation where, as above, a barrier either allowed belowground interaction, or not. In all experiments, pots were arranged in a randomized block design with 15 replicates in all four experiments.

Each pot received two seeds five cm apart, with their location relative to the root or shoot barrier depending on treatment (Figure 1). Here, the alone treatment still included two plants, but there was a barrier between them both above and below ground. Prior to sowing, the soil was saturated with tap water, and freely watered each day until germination. After germination, plants were put on strict watering schedules that differed by experiment because the different potting media had different water holding capacities and are described below. In experiments 1, 3 and 4 (Table 1) fertilizer was dissolved into water and added with water on a pre-defined schedule. In experiment 2, only water was applied. In each case, the fertilizer was water soluble 24-8-16 of N-P-K solution that also contained micronutrients (Miracle-Gro All Purpose Plant Food, The Scotts Miracle-Gro Company, Marysville, Ohio, USA). The concentration and application of fertilizer varied by experiment in attempts to replicate previous findings. Major differences among the four separate experiments are summarised in Table 1, and described below.

Experiment 1, soil:

Experiment one was performed in early spring from February 18, 2016 to April 28, 2016. The planting media was pure potting soil (propagation mix soil, Sungro Company, Agawam, Massachusetts,

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.29.318550; this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

USA). Each 6.2L pot was watered every other week with exactly 1L of water measured and poured into the pots (0.5L per half pot). The large pot size, and in particular the ratio of depth to surface area exposed to air, meant that soil did not completely dry between each watering. Pots were fertilized with a nutrient solution that was 0.25 g/L during weeks 3 and 6. All other experimental details were as above. Though we detail results below, this experiment failed to replicate published results, working from the hypothesis of different responses across different contexts, we adjusted the soil environment.

Experiment 2, soil-gravel:

Plants in experiment 1 grew extremely large and produced many fruits per individual. Thus, we hypothesized that if nutrients were highly available and not limiting to growth, plants may not experience competition because there were more nutrients than either plant could use (Casper and Jackson, 1997). Thus, to reduce nutrient availability and attempt to induce competition for limited resources, the planting media in this experiment was a 1:1 mixture of potting soil and calcined clay gravel. Experiment 2 was performed in early autumn from September 5, 2016 to November 14, 2016. Again, each 6.2L pot volume was watered every two weeks with exactly 1L of water. No fertilizer was applied. Unlike Experiment 1, the smaller plants in experiment 2 appeared to tip away from each other above ground because they failed to grasp the stake with their tendrils. In this experiment, bird netting (1.9 cm mesh, Bird Barricade, DeWitt Company, Sikeston, Missouri, USA) was wrapped loosely around the above ground portion of the experiment to keep plants within the vertical space above the pots. The netting was very fine and has undetectable effects on light levels (data not shown). All other experimental details were as above. Again, though we detail results below, this experiment failed to replicate published results, once again we hypothesized that context was the cause and adjusted the soil environment.

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.29.318550; this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Experiment 3, gravel:

Gersani et al. (1998), O'Brien et al. (2005) and Chen et al. (2015) all used nutrient free potting media in their pea experiments, and applied nutrients exclusively during watering. Thus, we hypothesized that there might be something in the context of this nutrient delivery mechanism that led to their detection of neighbour responses. The planting media in experiment 3 was therefore pure calcined clay gravel (Turface Athletics MVP, PROFILE Products LLC, Buffalo Grove, Illinois, USA). Experiment three was performed in winter from December 6, 2017 to February 14, 2017. Plants were loosely tied to the stakes with stretch tie tape. The gravel did not hold water well, and so these plants were watered once each week, again by measuring exactly 1L of water per pot, however in this experiment plants were fertigated every week because the gravel did not hold water well. The fertilizer concentration was 0.5 g/L, and supplied during weeks 3, 5, 7, and 9 of growth. All other experimental details were as above. Though we detail results below, this experiment also failed to replicate published results, again we hypothesized that context was the cause and adjusted the soil environment one last time.

Experiment 4, vermiculite:

While Gersani et al. (1998) and O'Brien et al. (2005) indeed used nutrient free potting media and nutrients supplied only in aqueous solution, they used vermiculite in both experiments, not gravel. Thus, our final hypothesis was that there was something unique about the context of growing in vermiculite compared to gravel and/or soil that might have led to their results. Thus, the planting media in experiment 4 was pure vermiculite (Coarse Vermiculite, Perlite Vermiculite Packaging Industries, Inc., North Bloomfield, Ohio, USA), which contained no nutrients. Experiment four was performed in early autumn from September 6, 2018 to November 15. Plants were loosely tied to the stakes with stretch tie tape. Again, all nutrients were supplied by liquid fertilizer using the same water-fertigation schedule as

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.29.318550; this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

in experiment 3. The other experimental details were as above. After failing to repeat published results four times, we attempted no further experiments.

Harvest

In all four experiments, after 10 weeks of growth the peas were harvested. The leaves, shoots, pods, and roots were collected separately. Roots were washed on a 2mm sieve. As expected, intermingled roots of neighbouring plants could not be separated. Tissues were dried at 60°C to constant mass, and then weighed.

Analysis

Since roots of interacting neighbours typically cannot not be separated, there are two approaches to dealing with this. One is to divide the total mass of plants with neighbours by 2 and compare this to the mass of one plant grown alone. However, this does not control for size asymmetry among the two interacting plants (Laird and Aarssen, 2005). The other approach, is to sum the biomass of two plants grown alone and compare this to the mass of the two neighbour plants across all treatments, thereby controlling for size asymmetry whether or not the plants actually interacted (McNickle and Brown, 2014). We took this second approach. Data from each experiment was analysed using GLMM with treatment as a fixed effect, and block as a random effect using lme4 (Bates, 2007) in the R statistical environment. Plants grew very differently in different soil contexts, but we were not interested in comparing soil contexts. Rather we viewed each experiment was an independent attempt at replicating a previously published result. Therefore, we analysed each experiment individually. For each experiment we examined leaf, stem, root, fruit and shoot (i.e. leaf + stem) biomass of both plants summed together as separate response variables. All biomass data was $\log(x + c)$ transformed where $c = 0.01$.

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.29.318550; this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Meta-analysis

Once it became clear that we were unable to replicate published results of under- or over-proliferation, we also sought to combine our four experiments with existing results in the literature. From the known studies in the literature, we extracted the mean root production with and without neighbours, and the mean pod production with and without neighbours, as well as the standard deviations. Some studies imposed a second treatment, and these means are recorded separately resulting in multiple data points for the following studies: (i) O'Brien et al. (2005) crossed the neighbour addition treatment with low and high nutrient addition; (ii) Chen et al. (2015) manipulated volume (6 levels) and nutrient concentration (3 levels with neighbour controls) simultaneously and; (iii) McNickle et al. (2020) had plants with and without mycorrhizae. The supplementary data details the different treatments. Though, Gersani et al. (1998) used pea, they used a split root approach. In the experiment one pot had neighbours, while the other had neighbour-free soil. Given that all other studies compared neighbours that shared soil, to plants alone that did not share soil, we did not know how to compare the results of Gersani et al. (1998) to the others, and therefore do not include it in the meta-analysis.

When raw data were available (Chen et al., 2015, McNickle unpubl., McNickle et al., 2020), means and standard deviations were calculated from the data for alone and with neighbours, within each additional treatment if present. When data were not available (O'Brien et al., 2005, Meier et al., 2013, Jacob et al., 2017), means and standard deviations were extracted from the figures using imageJ by calibrating the length measurement tool to the scale of the y-axis. Two studies only collected root data and did not have pod data, and these two studies also did not report any estimate of variation around the reported means (Meier et al., 2013, Jacob et al., 2017). In addition, we had access to an old pilot experiment from our own research that also used the ‘Little Marvel’ cultivar but was not followed

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.29.318550; this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

up on (McNickle unpubl.; Supplementary information, Table S1, Figure S1). We include those data here as well.

To account for potential methodological differences among studies, we used a log response ratio ($lnRR$) as our test statistic (Hedges et al., 1999) calculated according to:

$$lnRR = \ln\left(\frac{\bar{X}_n}{\bar{X}_a}\right), \qquad eqn\ 1$$

where $\bar{X}_n$ is the mean response variable in the presence of a neighbour, and $\bar{X}_a$ is the mean response variable when grown alone. By constructing the ratio this way, it will be negative in the case of under-proliferation, positive in the case of over-proliferation, and zero in the case of an IFD response. The standard deviation of lnRR is given by:

$$SD_{RR} = \sqrt{\frac{{SD_n}^2}{n{\bar{X}_n}^2} + \frac{{SD_a}^2}{n{\bar{X}_a}^2}}, \qquad Eqn\ 2$$

where n is the sample size in the study, SD_n is the standard deviation of the neighbour response, and SD_a is the standard deviation of the alone response (Lajeunesse, 2015). We calculated $lnRR$ for individual root biomass, and for lifetime seed yield.

In addition to the debate about root responses to neighbours, there is debate about methodological controls. Specifically, how and whether to control nutrients in neighbour treatments. One study (Chen et al., 2015) in the meta-analysis has been interpreted two different ways to draw very different conclusions. Chen et al (2015) argued that controlling pot volume was more important than nutrients in the study of plant competition, and made planned comparisons that confounded soil nutrient concentration with neighbour addition while controlling for pot volume. Analysis of these six planned comparisons concluded that plants were under-proliferating roots. McNickle et al (In press) argued that nutrients were important in the study of below ground plant competition and used the

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.29.318550; this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

same data to make three planned comparisons that confounded pot volume with neighbours but controlled soil nutrients. This analysis concluded plants were ignoring neighbours and exhibited the IFD behaviour. To deal with the fact that these two interpretations are alternative theoretical research programmes (sensu Lakatos, 1978) which cannot both be true, we performed the meta-analysis data both ways as if there were two distinct literatures and leave the interpretation to readers depending on their world view (sensu Brown, 2001).

RESULTS

Experiments 1-4

Though plants grew differently in each of our four experiments, plants produced the same pod mass (Fig 2 a-d), and root mass (Fig 2 e-h) regardless of whether there was a neighbour above or below ground in every experiment (Table 2). Thus, we were unable to replicate previous findings of either neighbour avoidance, or root over-proliferation despite four different experiments that included 480 individual plants across 3 years. We conclude that the results of all four experiments support the IFD hypothesis. Because of previously expressed concerns about pot volume effects (Hess and de Kroon, 2007, Chen et al., 2015), we also note that since all plants with root barriers were in pots of half the volume of all plants without root barriers, these results mean that potting volume also had no effect on plant growth in any of our experiments.

Most root based theories described here make no specific hypotheses about leaf and stem responses, though an implicit assumption is probably that there are no shoot differences among treatments. We analysed leaf and stem tissue pools separately (Fig 3). Only two significant differences arose at the $\alpha = 0.05$ significance level (Table 2). First, in experiment 2, plants that experienced root competition only, produced significantly more stem biomass than plants that experienced shoot

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.29.318550; this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

competition only (Fig 3B). Second, in experiment 3, plants that experienced no competition produced significantly more leaf biomass than plants that experienced root competition only (Fig 3G). Though these differences were significant in the statistical sense, they were not biologically large differences. In addition, when leaf and stem mass are summed as the more commonly analysed "shoot" biomass, the significance disappeared (Table 2). Given that: (i) most root theories make no specific shoot hypotheses; (ii) the fact that the biological differences were slight (Fig 3B, G), and; (iii) this is a paper partially about type I errors, we do not make very much of these differences. Indeed, Table 2 has exactly 20 independent statistical tests. With $\alpha = 0.05$, on average we would expect one spurious significant result, and two spurious results would not be out of the question. A simple Bonferroni correction reduces the significance cut-off to $\alpha = 0.0025$, and erases all statistical significance.

Meta-analysis

The meta-analysis included 14 or 17 estimates of root and fruit log response ratios from 224 or 299 replicate alone and with neighbour pairs depending on how Chen et al. (2015) is included. When all available results were combined the consensus among studies was that pea has no response to neighbours, and that root and pod growth supports the null IFD model (Fig 4). Interestingly, this was the case under either interpretation of the Chen et al (2015) data.

DISCUSSION

Many scientific fields have been experiencing a reproducibility crisis (Ioannidis, 2005, Saltelli and Funtowicz, 2017). Our understanding is that this reproducibility crisis began in psychology in part when a study reported evidence for psychic perception in humans (Bem, 2011, Pashler and Wagenmakers, 2012). Not surprisingly, this result was quickly criticised (Wagenmakers et al., 2011), and the result was not repeated in a subsequent study (Galak et al., 2012). The mounting concerns about reproducability

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.29.318550; this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

led some in psychology to wonder how often any results could be repeated, and the crisis culminated in a large collaborative effort to try and reproduce 100 published results. This effort found that only 36% of these 100 previously published results were reproducible (Aarts et al., 2015). The cause was largely blamed on low statistical power, and a culture that values novelty over replication (Hunter, 2001, Moonesinghe et al., 2007, Stanley et al., 2018). As far as we can tell, ecology has not been fully drawn into the replication crisis like other fields such as psychology (Pashler and Wagenmakers, 2012), medicine (Ioannidis, 2005) and economics (Camerer et al., 2016). However, others have noted that true replications are relatively rare in ecology (Belovsky et al., 2004) and statistical power to detect all but the largest effect sizes can be low in ecological experiments (Steidl et al., 1997). It is likely that there is some level of non-reproducibility in ecology.

In addition, to the replication crisis, there is also the problem of p-hacking, which can enlarge the number of type-I errors in the published literature. This p-hacking occurs when a just barely insignificant result at the $\alpha = 0.05$ level is nudged across the significance line through slight adjustments in the statistical approach (Bruns and Ioannidis, 2016). Without p-hacking, and with statistical significance set arbitrarily at $\alpha = 0.05$, we would expect one in twenty published results in the literature which have not been replicated to be type-I errors where a true null hypothesis was erroneously rejected. However, in the presence of p-hacking, the number of published un-replicated type-I errors can be significantly higher. Indeed, as in many fields there is evidence of p-hacking in the broader field of biology as a whole (Head et al., 2015). The presence of p-hacking was also thought to be part of the reason that only 35% of the 100 psychology studies in the literature were reproducible(Aarts et al., 2015).

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.29.318550; this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Here, we used a debated problem of how plants respond to neighbouring roots framed as a case study in reproducible ecology. Different research groups had developed three mutually exclusive hypotheses, two of which had been supported by the growth behaviour of a single species and a single genotype in the literature (Gersani et al., 1998, O'Brien et al., 2005). Our original goal was simply to develop a model system for exploring pea responses to neighbours across contexts. We began with an assumption that the published results were 'true', and then worked from the hypothesis that the difference in results were caused by different contexts. As such, we tried a number of different soil types and nutrient availabilities, all seeking to replicate previous methods. Our failure to replicate published results (Figure 2, 3) combined with meta-analysis of all available studies (Fig 4) show that the best current evidence is that pea exhibits the null IFD response, and that previous findings may have simply been type I errors made at the $\alpha = 0.05$ level. This is true whether one views nutrients as important (McNickle, in press) or unimportant (Chen et al., 2015) to plant competition below ground (Fig 4).

We stress that our conclusion that common pea exhibits an IFD response should be confined to pea, and possibly even to the 'Little Marvel' cultivar. This conclusion should not be extended to the entire plant kingdom. The hypothesis that different species and different genotypes may have any of the three theoretical root responses (under- or over-proliferation, or IFD) is not falsified through the study of one species *P. sativum* and one cultivar Little Marvel. Indeed, surprisingly few species have been assayed for their response to neighbouring plants, with (strangely) almost half of assayed species coming from the fabaceae family (Smyeka and Herben, 2017). In addition, though different responses across different contexts do not seem to have been the cause for variation in pea responses in the literature, this does not refute the hypothesis that other species may have different responses among different contexts (Cahill et al., 2010, McNickle et al., 2016). Much more work that includes replication

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.29.318550; this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

both within and among species, and within and among genotypes is needed to draw general conclusions across the hundreds of thousands of species in the plant kingdom (Joppa et al., 2011).

The studies included in our meta-analysis used on average of ~13 replicates but included as few as 4 (Jacob et al., 2017) replicates, and as many as 25 (Chen et al., 2015). Thus, each study varied significantly in its statistical power (where statistical power is defined as 1 minus the probability of making a type II error; accepting a false null). None of the published studies (including our experiments 1 – 4) had the statistical power to detect small or medium effect sizes with a type I error rate of one in twenty, and a type II error rate of one in five (Fig 5A). Indeed, even with 224 replicates across all studies in the meta-analysis (Fig 4), this just barely would allow for the statistical power to detect small effect sizes (Fig 5A). Thus, as ultimately suspected in the failure to replicate previously published results in psychology, we conclude the non-reproducibility of the pea under- or over-proliferation behaviour was due to low statistical power.

This is just one case study in ecology, but there are other examples. One in an adjacent field was the idea that there should be a trade-off in a plant's ability to precisely place roots into nutrient rich zones of soil, and the ability of a plant to explore large spatial scales of soil volume (Campbell et al., 1991). This scale-precision trade-off hypothesis emerged from a study that with ten species but only five replicates per species, and shaped root foraging research for decades (Hodge, 2004, Hodge, 2006). Ultimately, meta-analysis revealed that there was not really any evidence for this scale-precision trade-off, and that the first paper was likely just a low statistical power type-I error (Kembel and Cahill, 2005, de Kroon and Mommer, 2006, Kembel et al., 2008) but see (Grime, 2007). It is not our goal to review every such case of a failure to replicate previously published results across the ecology literature, but these two involving plant roots are unlikely to be the only such examples across the field.

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.29.318550; this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Given that a few results were enough to spark a replication crises in other fields (Pashler and Wagenmakers, 2012, Saltelli and Funtowicz, 2017), it is worth considering how other fields have approached a solution. One of the simplest solutions is to make data publically available in data repositories, and ecologists seem to have largely embraced this approach already judging from the requirements of most journals in our field. In addition, some fields approach the problem with very high power to detect small effect sizes. For example, sample sizes in medicine are often in the thousands. Here, human lives literally depend on the ability to detect even small effect sizes. However, this may not be a reasonable approach in ecology. It is one thing to administer a medical intervention to thousands of people who then go off and live their lives during the experiment, and another to care for thousands of plants in a greenhouse, or survey millions of hectares of forest.

As another example, fields also differ in the cut-off they use for statistical significance. Most scientific fields, including ecology, use a statistical significance cut-off of $\alpha = 0.05$, or a type I error rate of one in twenty. In the language of confidence intervals, where a standard deviation is denoted by the symbol sigma, this statistical cut-off is also sometimes called two-sigma. However, in physics a five-sigma level of significance, which is $\alpha = 0.0000003$, or a type I error rate of one in 3.5 million is required for a discovery (E.g. the detection of gravity waves: Abbott et al., 2016). Alternatively, three sigma, which is $\alpha = 0.0027$, or an error rate of one in about 370, is taken as weak evidence that a phenomena might exist (Lyons, 2013). Importantly, either two, three or five sigma has no theoretical basis, it is just a convenient cut-off based on a field's willingness to accept error type I error (Colquhoun, 2017). Perhaps five-sigma is unrealistic for ecology (Fig 5c), but when we look through our own past published results, three-sigma seems frequently achievable, and perhaps an error rate of one in 370 is preferable to an error rate of one in twenty. For context, in the study of pea roots, a sample sizes of

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.29.318550; this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

n=30 to n=60 would be able to detect both large and medium effect sizes at sigma three (Fig 5B). Furthermore, in the context of experiments 1-4, the somewhat ambiguous leaf and stem effect sizes would no longer be "statistically significant" under three-sigma.

Another solution is a culture of science that values replication studies, and ecologists seem increasingly open to such replication. Indeed, replication of studies allows for meta-analysis, which has been an increasingly powerful tool in ecology to resolve large questions and problems (E.g. Cardinale et al., 2007, Kembel et al., 2008, Vellend et al., 2013). In addition, large scale experiments that are replicated by independent groups across space and time are becoming increasingly common in ecology. Examples of these include the plethora of Free Air Carbon dioxide Enrichment studies (Norby and Zak, 2011), networks of large atmospheric eddy covariance towers (Luyssaert et al., 2007), the National Science Foundation of the USA's National Ecological Observatory network (NEON; https://www.neonscience.org/). Ecologists are also forming increasingly large networks of studies using common methods such as nutrient net (https://nutnet.org/), drought net (https://drought-net.colostate.edu/), the global biodiversity initiative (https://www.gfbinitiative.org/) or the Smithsonian ForestGeo network (https://forestgeo.si.edu/).

Finally, pre-registration has been one solution increasingly adopted in other fields (Van 't Veer and Giner-Sorolla, 2016). Pre-registration requires researchers to detail and publish their hypotheses, methods and analyses before the experiments are performed. Pre-registration can potentially protect against p-hacking, and post-hoc data exploration. We are unaware of any wide-spread attempts at pre-registration in ecology.

Conclusion

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.29.318550; this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

A small number of plant ecologists have been arguing about plant root responses to neighbours for several decades, but only sporadically performing experiments to advance the debate. Three mutually exclusive theories have been devised: (i) roots should avoid neighbours and under-proliferate roots in response to competition; (ii) roots should confront neighbours and over-proliferate in response to competition, and; (iii) roots cannot recognize neighbours and do not respond independent from their responses to nutrients forming an IFD of roots in soil. Published studies in the literature showed support for all three theories with the common pea. We were unable to replicate previous findings of over-proliferation, or under-proliferation in four separate experiments. Combined with meta-analysis, we conclude that pea exhibits the IFD response. There is a growing recognition that many fields in science have a reproducibility problem. This study of one problem in plant ecology may not represent a replication crisis in ecology. Yet, it highlights the risk of making conclusions which are not replicated, and independently verifiable. Following success in other fields of science, we recommend ecologists consider a more stringent statistical significance cut-off, that ecologists be more sceptical of results that have not been independently replicated, and that the field explores pre-registration of experimental design.

ACKNOWLEDGEMENTS

This work was supported by the USDA National Institute of Food and Agriculture Hatch project 1010722. The authors declare no conflicts of interest.

AUTHOR CONTRIBUTIONS

GGM and MM designed the experiments. MM performed experiments 1,2 and 4 and statistical analyses. AK performed experiment 3, and statistical analyses. MM drafted the initial manuscript, and all authors contributed to revisions.

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.29.318550; this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

REFERENCES

Aarts, A. A., Anderson, J. E., Anderson, C. J., Attridge, P. R., Attwood, A., Axt, J., Babel, M., Bahnik, S., Baranski, E., Barnett-Cowan, M., Bartmess, E., Beer, J., Bell, R., Bentley, H., Beyan, L., Binion, G., Borsboom, D., Bosch, A., Bosco, F. A., Bowman, S. D., Brandt, M. J., Braswell, E., Brohmer, H., Brown, B. T., Brown, K., Bruning, J., Calhoun-Sauls, A., Callahan, S. P., Chagnon, E., Chandler, J., Chartier, C. R., Cheung, F., Christopherson, C. D., Cillessen, L., Clay, R., Cleary, H., Cloud, M. D., Cohn, M., Cohoon, J., Columbus, S., Cordes, A., Costantini, G., Alvarez, L. D. C., Cremata, E., Crusius, J., DeCoster, J., DeGaetano, M. A., Della Penna, N., den Bezemer, B., Deserno, M. K., Devitt, O., Dewitte, L., Dobolyi, D. G., Dodson, G. T., Donnellan, M. B., Donohue, R., Dore, R. A., Dorrough, A., Dreber, A., Dugas, M., Dunn, E. W., Easey, K., Eboigbe, S., Eggleston, C., Embley, J., Epskamp, S., Errington, T. M., Estel, V., Farach, F. J., Feather, J., Fedor, A., Fernandez-Castilla, B., Fiedler, S., Field, J. G., Fitneva, S. A., Flagan, T., Forest, A. L., Forsell, E., Foster, J. D., Frank, M. C., Frazier, R. S., Fuchs, H., Gable, P., Galak, J., Galliani, E. M., Gampa, A., Garcia, S., Gazarian, D., Gilbert, E., Giner-Sorolla, R., Glockner, A., Goellner, L., Goh, J. X., Goldberg, R., Goodbourn, P. T., Gordon-McKeon, S., Gorges, B., Gorges, J., Goss, J., Graham, J., et al. (2015) Estimating the reproducibility of psychological science. *Science,* **349,** 8.

Abbott, B. P., Abbott, R., Abbott, T. D., Abernathy, M. R., Acernese, F., Ackley, K., Adams, C., Adams, T., Addesso, P., Adhikari, R. X., Adya, V. B., Affeldt, C., Agathos, M., Agatsuma, K., Aggarwal, N., Aguiar, O. D., Aiello, L., Ain, A., Ajith, P., Allen, B., Allocca, A., Altin, P. A., Anderson, S. B., Anderson, W. G., Arai, K., Arain, M. A., Araya, M. C., Arceneaux, C. C., Areeda, J. S., Arnaud, N., Arun, K. G., Ascenzi, S., Ashton, G., Ast, M., Aston, S. M., Astone, P., Aufmuth, P., Aulbert, C., Babak, S., Bacon, P., Bader, M. K. M., Baker, P. T., Baldaccini, F., Ballardin, G., Ballmer, S. W., Barayoga, J. C., Barclay, S. E., Barish, B. C., Barker, D., Barone, F., Barr, B., Barsotti, L., Barsuglia, M., Barta, D., Bartlett, J., Barton, M. A., Bartos, I., Bassiri, R., Basti, A., Batch, J. C., Baune, C., Bavigadda, V., Bazzan, M., Behnke, B., Bejger, M., Belczynski, C., Bell, A. S., Bell, C. J., Berger, B. K., Bergman, J., Bergmann, G., Berry, C. P. L., Bersanetti, D., Bertolini, A., Betzwieser, J., Bhagwat, S., Bhandare, R., Bilenko, I. A., Billingsley, G., Birch, J., Birney, R., Birnholtz, O., Biscans, S., Bisht, A., Bitossi, M., Biwer, C., Bizouard, M. A., Blackburn, J. K., Blair, C. D., Blair, D. G., Blair, R. M., Bloemen, S., Bock, O., Bodiya, T. P., Boer, M., Bogaert, G., Bogan, C., Bohe, A., Bojtos, P., Bond, C., et al. (2016) Observation of gravitational waves from a binary black hole merger. *Physical Review Letters,* **116,** 16.

Ballare, C. L., Sanchez, R. A., Scopel, A. L., Casal, J. J. & Ghersa, C. M. (1987) Early detection of neighbor plants by phytochrome perception of spectral changes in reflected sunlight. *Plant Cell and Environment,* **10,** 551-557.

Bates, D. M. (2007) Linear mixed model implementation in lme4.

Belovsky, G. E., Botkin, D. B., Crowl, T. A., Cummins, K. W., Franklin, J. F., Hunter, M. L., Joern, A., Lindenmayer, D. B., MacMahon, J. A., Margules, C. R. & Scott, J. M. (2004) Ten suggestions to strengthen the science of ecology. *Bioscience,* **54,** 345-351.

Bem, D. J. (2011) Feeling the future: Experimental evidence for anomalous retroactive influences on cognition and affect. *Journal of Personality and Social Psychology,* **100,** 407-425.

Brown, J. S. (2001) Ngongas and ecology: On having a worldview. *Oikos,* **94,** 6-16.

Bruns, S. B. & Ioannidis, J. P. A. (2016) P-curve and p-hacking in observational research. *PLOS ONE,* **11,** e0149144.

Cahill, J. F., Jr., McNickle, G. G., Haag, J. J., Lamb, E. G., Nyanumba, S. M. & St. Clair, C. C. (2010) Plants integrate information about nutrients and neighbors. *Science,* **328,** 1657-.

Camerer, C. F., Dreber, A., Forsell, E., Ho, T.-H., Huber, J., Johannesson, M., Kirchler, M., Almenberg, J., Altmejd, A., Chan, T., Heikensten, E., Holzmeister, F., Imai, T., Isaksson, S., Nave, G., Pfeiffer, T.,

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.29.318550; this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Razen, M. & Wu, H. (2016) Evaluating replicability of laboratory experiments in economics. *Science,* **351,** 1433-1436.

Campbell, B. D., Grime, J. P. & Mackey, J. M. L. (1991) A trade-off between scale and precision in resource foraging. *Oecologia,* **87,** 532-538.

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

Casper, B. B. & Jackson, R. B. (1997) Plant competition underground. *Annual Review of Ecology and Systematics,* **28,** 545-570.

Chen, B. J. W., During, H. J., Vermeulen, P. J., de Kroon, H., Poorter, H. & Anten, N. P. R. (2015) Corrections for rooting volume and plant size reveal negative effects of neighbour presence on root allocation in pea. *Functional Ecology,* **29,** 1383-1391.

Colquhoun, D. (2017) The reproducibility of research and the misinterpretation of *p*-values. *Royal Society Open Science,* **4,** 171085.

Craine, J. M., Fargione, J. & Sugita, S. (2005) Supply pre-emption, not concentration reduction, is the mechanism of competition for nutrients. *New Phytologist,* **166,** 933-940.

de Kroon, H. & Mommer, L. (2006) Root foraging theory put to the test. *Trends in Ecology & Evolution,* **21,** 113-116.

Falik, O., Reides, P., Gersani, M. & Novoplansky, A. (2003) Self/non-self discrimination in roots. *Journal of Ecology,* **91,** 525-531.

Falster, D. S. & Westoby, M. (2003) Plant height and evolutionary games. *Trends in Ecology & Evolution,* **18,** 337-343.

Fretwell, S. D. & Lucas, H. L. (1969) On territorial behavior and other factors influencing habitat distribution in birds. *Acta Biotheoretica,* **19,** 16-36.

Galak, J., LeBoeuf, R. A., Nelson, L. D. & Simmons, J. P. (2012) Correcting the past: Failures to replicate psi. *Journal of Personality and Social Psychology,* **103,** 933-948.

Gersani, M., Abramsky, Z. & Falik, O. (1998) Density-dependent habitat selection in plants. *Evolutionary Ecology,* **12,** 223-234.

Gersani, M., Brown, J. S., O'Brien, E. E., Maina, G. M. & Abramsky, Z. (2001) Tragedy of the commons as a result of root competition. *Journal of Ecology,* **89,** 660-669.

Givnish, T. J. (1982) On the adaptive significance of leaf height in forest herbs. *American Naturalist,* **120,** 353-381.

Grime, J. P. (2007) The scale-precision trade-off in spacial resource foraging by plants: Restoring perspective. *Annals of Botany,* **99,** 1017-1021.

Gruntman, M. & Novoplansky, A. (2004) Physiologically mediated self/non-self discrimination in roots. *Proceedings of the National Academy of Sciences of the United States of America,* **101,** 3863-3867.

Head, M. L., Holman, L., Lanfear, R., Kahn, A. T. & Jennions, M. D. (2015) The extent and consequences of p-hacking in science. *PLOS Biology,* **13,** e1002106.

Hedges, L. V., Gurevitch, J. & Curtis, P. S. (1999) The meta-analysis of response ratios in experimental ecology. *Ecology,* **80,** 1150-1156.

Hess, L. & de Kroon, H. (2007) Effects of rooting volume and nutrient availability as an alternative explanation for root self/non-self discrimination. *Journal of Ecology,* **95,** 241-251.

Hodge, A. (2004) The plastic plant: Root responses to heterogeneous supplies of nutrients. *New Phytologist,* **162,** 9-24.

Hodge, A. (2006) Plastic plants and patchy soils. *Journal of Experimental Botany,* **57,** 401-411.

Hunter, J. E. (2001) The desperate need for replications. *Journal of Consumer Research,* **28,** 149-158.

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.29.318550; this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Hutchings, M. J. & de Kroon, H. (1994) Foraging in plants - the role of morphological plasticity in resource acquisition. *Advances in Ecological Research, Vol 25,* **25,** 159-238.

Ioannidis, J. P. A. (2005) Why most published research findings are false. *PLOS Medicine,* **2,** e124.

Jacob, C. E., Tozzi, E. & Willenborg, C. J. (2017) Neighbour presence, not identity, influences root and shoot allocation in pea. *PLOS ONE,* **12,** e0173758.

Joppa, L. N., Roberts, D. L. & Pimm, S. L. (2011) How many species of flowering plants are there? *Proceedings of the Royal Society B: Biological Sciences,* **278,** 554-559.

Kembel, S. W. & Cahill, J. F. (2005) Plant phenotypic plasticity belowground: A phylogenetic perspective on root foraging trade-offs. *American Naturalist,* **166,** 216-230.

Kembel, S. W., de Kroon, H., Cahill, J. F. & Mommer, L. (2008) Improving the scale and precision of hypotheses to explain root foraging ability. *Annals of Botany,* **101,** 1295-1301.

Křivan, V., Cressman, R. & Schneider, C. (2008) The ideal free distribution: A review and synthesis of the game-theoretic perspective. *Theoretical Population Biology,* **73,** 403-425.

Laird, R. A. & Aarssen, L. W. (2005) Size inequality and the tragedy of the commons phenomenon in plant competition. *Plant Ecology,* **179,** 127-131.

Lajeunesse, M. J. (2015) Bias and correction for the log response ratio in ecological meta-analysis. *Ecology,* **96,** 2056-2063.

Lakatos, I. (1978) *The methodology of scientific research programmes, philisophical papers volume 1.* Cambridge University Press, New Jersey, New York, USA.

Litav, M. & Harper, J. L. (1967) A method for studying spatial relationships between root systems of 2 neighbouring plants. *Plant and Soil,* **26,** 389-392.

Luyssaert, S., Inglima, I., Jung, M., Richardson, A. D., Reichstein, M., Papale, D., Piao, S. L., Schulzes, E. D., Wingate, L., Matteucci, G., Aragao, L., Aubinet, M., Beers, C., Bernhofer, C., Black, K. G., Bonal, D., Bonnefond, J. M., Chambers, J., Ciais, P., Cook, B., Davis, K. J., Dolman, A. J., Gielen, B., Goulden, M., Grace, J., Granier, A., Grelle, A., Griffis, T., Grunwald, T., Guidolotti, G., Hanson, P. J., Harding, R., Hollinger, D. Y., Hutyra, L. R., Kolar, P., Kruijt, B., Kutsch, W., Lagergren, F., Laurila, T., Law, B. E., Le Maire, G., Lindroth, A., Loustau, D., Malhi, Y., Mateus, J., Migliavacca, M., Misson, L., Montagnani, L., Moncrieff, J., Moors, E., Munger, J. W., Nikinmaa, E., Ollinger, S. V., Pita, G., Rebmann, C., Roupsard, O., Saigusa, N., Sanz, M. J., Seufert, G., Sierra, C., Smith, M. L., Tang, J., Valentini, R., Vesala, T. & Janssens, I. A. (2007) Co2 balance of boreal, temperate, and tropical forests derived from a global database. *Global Change Biology,* **13,** 2509-2537.

Lyons, L. (2013) Discovering the significance of 5 sigma. *arXiv***,** 1310.1284.

McNickle, G. G. (in press) Interpreting plant root responses to nutrients, neighbours and pot volume depends on researchers' assumptions. *Functional Ecology*.

McNickle, G. G. & Brown, J. S. (2014) An ideal free distribution explains the root production of plants that do not engage in a tragedy of the commons game. *Journal of Ecology,* **102,** 963–971.

McNickle, G. G., Deyholos, M. K. & Cahill, J. F. (2016) Nutrient foraging behaviour of four co-occurring perennial grassland plant species alone does not predict behaviour with neighbours. *Functional Ecology,* **30,** 420-430.

McNickle, G. G., Guinel, F. C., Sniderhan, A. E., Wallace, C. A., McManus, A. S., Fafard, M. M. & Baltzer, J. L. (2020) The benefits of mycorrhizae are frequency-dependent: A case study with a non-mycorrhizal mutant of pisum sativum. *bioRxiv***,** 2020.09.08.288084.

Meier, I. C., Angert, A., Falik, O., Shelef, O. & Rachmilevitch, S. (2013) Increased root oxygen uptake in pea plants responding to non-self neighbors. *Planta,* **238,** 577-586.

Mommer, L., Wagemaker, C. A. M., De Kroon, H. & Ouborg, N. J. (2008) Unravelling below-ground plant distributions: A real-time polymerase chain reaction method for quantifying species proportions in mixed root samples. *Molecular Ecology Resources,* **8,** 947-953.

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.29.318550; this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Moonesinghe, R., Khoury, M. J. & Janssens, A. C. J. W. (2007) Most published research findings are false—but a little replication goes a long way. *PLOS Medicine,* **4,** e28.

Murphy, G. P. & Dudley, S. A. (2007) Above- and below-ground competition cues elicit independent responses. *Journal of Ecology,* **95,** 261-272.

Murphy, G. P. & Dudley, S. A. (2009) Kin recognition: Competition and cooperation in impatiens (balsaminaceae). *American Journal of Botany,* **96,** 1990-1996.

Norby, R. J. & Zak, D. R. (2011) Ecological lessons from free-air co2 enrichment (face) experiments. *Annual Review of Ecology, Evolution, and Systematics,* **42,** 181-203.

O'Brien, E. E., Gersani, M. & Brown, J. S. (2005) Root proliferation and seed yield in response to spatial heterogeneity of below-ground competition. *New Phytologist,* **168,** 401-412.

Padilla, F. M., Mommer, L., de Caluwe, H., Smit-Tiekstra, A. E., Wagemaker, C. A. M., Ouborg, N. J. & de Kroon, H. (2013) Early root overproduction not triggered by nutrients decisive for competitive success belowground. *Plos One,* **8,** e55805.

Parrish, J. A. D. & Bazzaz, F. A. (1976) Underground niche separation in successional plants. *Ecology,* **57,** 1281-1288.

Pashler, H. & Wagenmakers, E. J. (2012) Editors' introduction to the special section on replicability in psychological science:A crisis of confidence? *Perspectives on Psychological Science,* **7,** 528-530.

Saltelli, A. & Funtowicz, S. (2017) What is science's crisis really about? *Futures,* **91,** 5-11.

Schenk, H. J., Callaway, R. M. & Mahall, B. E. (1999) Spatial root segregation: Are plants territorial? *Advances in Ecological Research, Vol 28,* **28,** 145-180.

Semchenko, M., Saar, S. & Lepik, A. (2014) Plant root exudates mediate neighbour recognition and trigger complex behavioural changes. *New Phytologist,* **204,** 631-637.

Smyeka, J. & Herben, T. (2017) Phylogenetic patterns of tragedy of commons in intraspecific root competition. *Plant and Soil,* **417,** 87-97.

Stanley, T. D., Carter, E. C. & Doucouliagos, H. (2018) What meta-analyses reveal about the replicability of psychological research. *Psychological Bulletin,* **144,** 1325-1346.

Steidl, R. J., Hayes, J. P. & Schauber, E. (1997) Statistical power analysis in wildlife research. *Journal of Wildlife Management,* **61,** 270-279.

Van 't Veer, A. E. & Giner-Sorolla, R. (2016) Pre-registration in social psychology-a discussion and suggested template. *Journal of Experimental Social Psychology,* **67,** 2-12.

Vellend, M., Baeten, L., Myers-Smith, I. H., Elmendorf, S. C., Beauséjour, R., Brown, C. D., De Frenne, P., Verheyen, K. & Wipf, S. (2013) Global meta-analysis reveals no net change in local-scale plant biodiversity over time. *Proceedings of the National Academy of Sciences,* **110,** 19456-19459.

Wagenmakers, E.-J., Wetzels, R., Borsboom, D. & van der Maas, H. L. J. (2011) Why psychologists must change the way they analyze their data: The case of psi: Comment on bem (2011). *Journal of Personality and Social Psychology,* **100,** 426-432.

Zhang, D. Y., Sun, G. J. & Jiang, X. H. (1999) Donald's ideotype and growth redundancy: A game theoretical analysis. *Field Crops Research,* **61,** 179-187.

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.29.318550; this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

TABLE 1: Summary of methodological differences among experiments 1-4. Throughout most of the paper we refer to the experiments by the media type.

	Experiment 1	**Experiment 2**	**Experiment 3**	**Experiment 4**
Timing	Spring 2016	Autumn 2016	Winter 2017	Autumn 2018
Media type	100% Soil	50% soil 50% gravel	100% gravel	100% Vermiculite
Plant support	Stakes	Stakes and Bird netting	Stakes and ties	Stakes and ties
Watering Schedule	Every other week	Every other week	Weekly	Weekly
Nutrient Application	¼ g/L nutrients twice	None	½ g/L nutrients four times	½ g/L nutrients four times

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.29.318550; this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

TABLE 2: ANOVA tables for the GLMMs on the biomass production of different plant tissues across neighbour addition treatments in experiments 1-4. Statistical significance at the $\alpha = 0.05$ level is marked with * and bold face font. Denominator degrees of freedom (Den. df) were estimated using Satterthwaite's method.

Tissue	Experiment	Num. df	Den. df	F	p
Root	**1. Soil**	3	40.96	1.04	0.3852
	2. Soil-gravel	3	42.00	2.65	0.0611
	3. gravel	3	55.00	1.03	0.388
	4. Vermiculite	3	33.04	1.52	0.2276
Stem	**1. Soil**	3	41.46	1.57	0.2103
	2. Soil-gravel	3	49.00	3.06	**0.0368***
	3. gravel	3	56.00	0.84	0.4761
	4. Vermiculite	3	34.64	1.19	0.3268
Leaf	**1. Soil**	3	41.01	0.87	0.4642
	2. Soil-gravel	3	49.00	2.43	0.0762
	3. gravel	3	56.00	3.4	**0.0237***
	4. Vermiculite	3	34.98	1.09	0.3666
Fruit	**1. Soil**	3	41.46	2.02	0.1259
	2. Soil-gravel	3	38.6	0.71	0.5497
	3. gravel	3	42.00	2.79	0.0523
	4. Vermiculite	3	35.1	1.66	0.1939
Shoot = (Leaf + Stem)	**1. Soil**	3	41.24	1.25	0.3039
	2. Soil-gravel	3	49.00	2.68	0.0562
	3. gravel	3	56.00	2.47	0.0715
	4. Vermiculite	3	34.91	1.11	0.3561

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.29.318550; this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

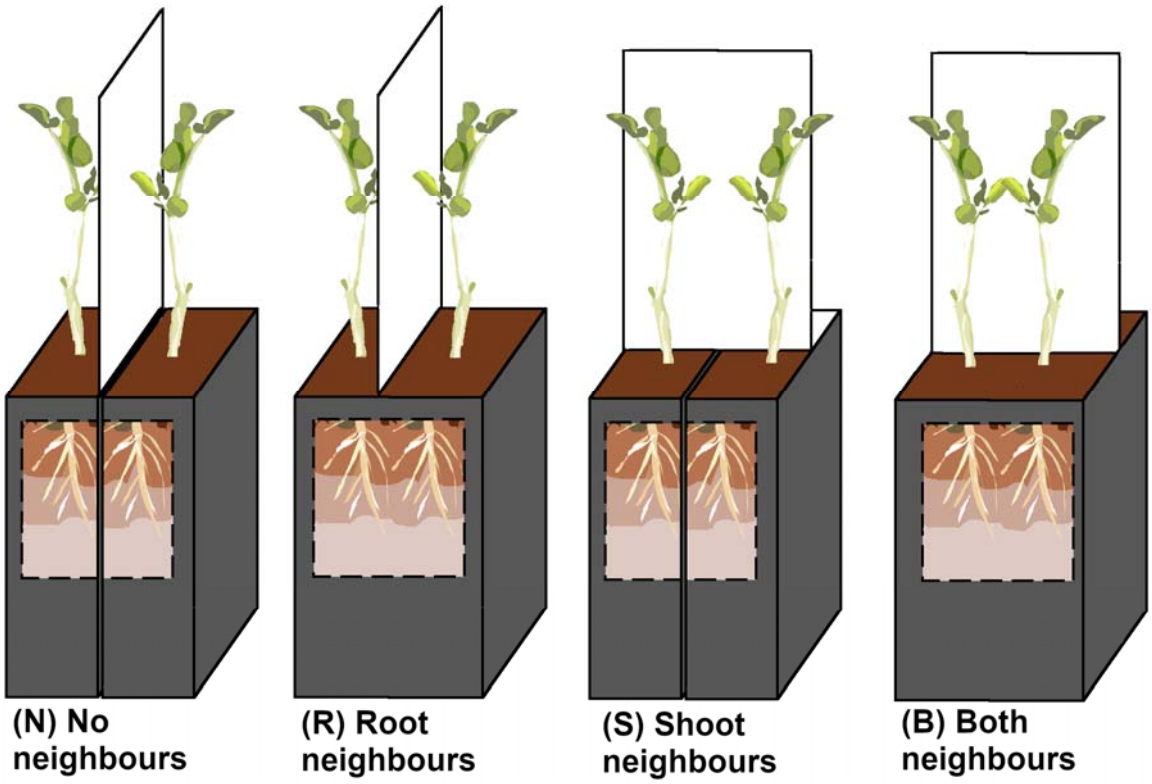

FIGURE 1: Schematic representation of the experimental design used to separate root and shoot competition. Dividers were placed between plants either above or belowground to create four treatments that included: (N) no interaction with neighbours; (R) only root interactions with neighbours; (S) only shoot interactions with neighbours, or; (B) both root and shoot interactions with neighbours. Image is not to scale, see methods for dimensions of pots and dividers.

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.29.318550; this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

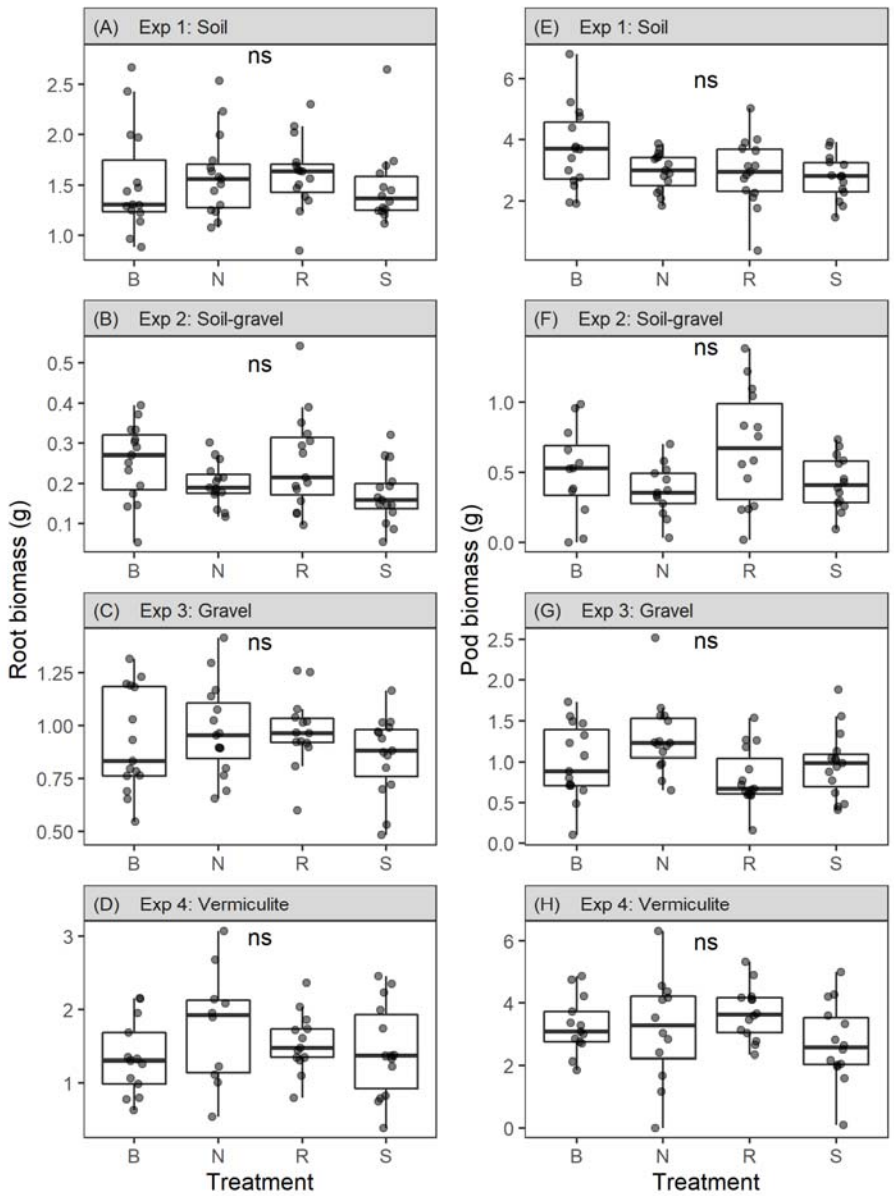

FIGURE 2: Results of root (A-D) and pod biomass (E-H) production across the four experiments as a test of different theories of root production in response to neighbours above and below ground. Treatment codes refer to both root and shoot interactions (B), no interactions (N), root interaction only (R), or shoot interactions only (S). The code ns indicates lack of statistical significance in GLMMs which supports the IFD model, and rejects the root under- or over-proliferation hypotheses (Table 2).

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.29.318550; this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

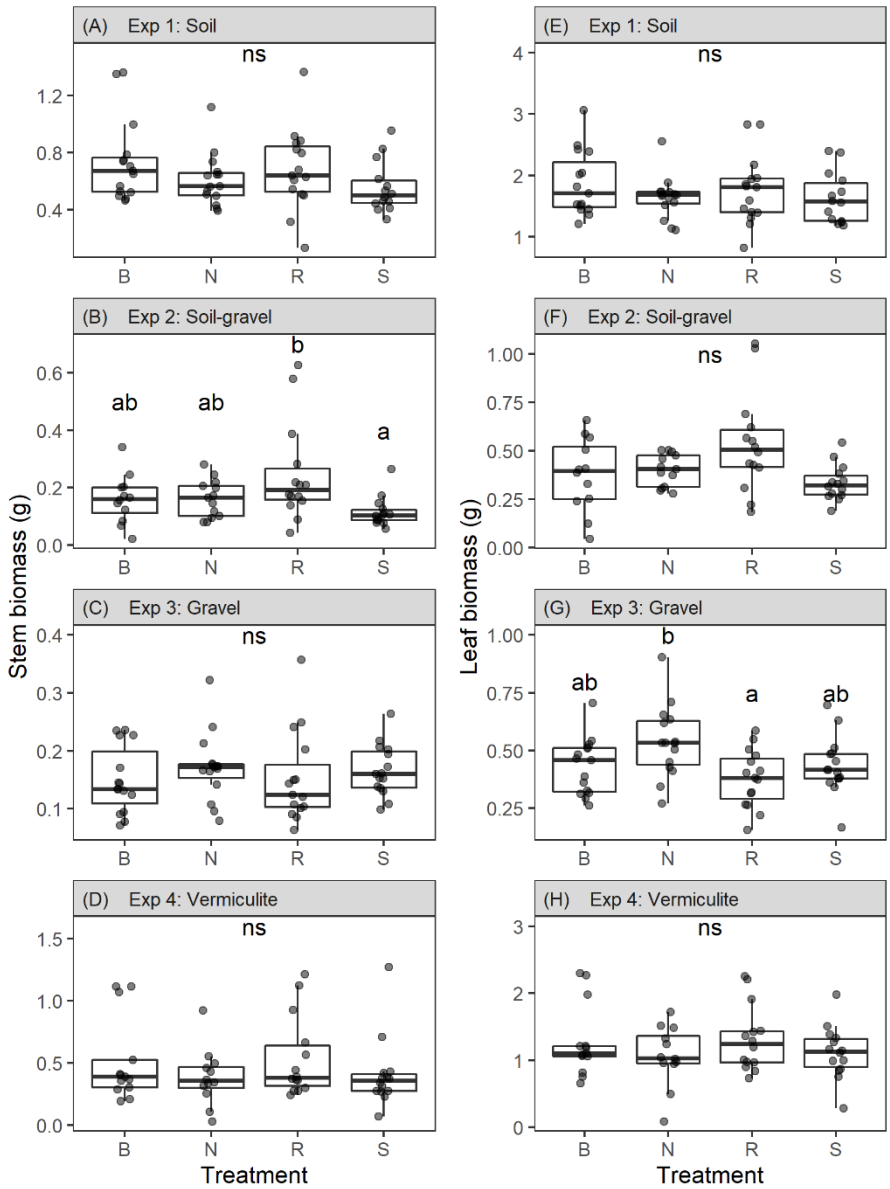

FIGURE 3: Results of stem and leaf production in response to neighbours either above or below ground. Most root based theories make no specific predictions about above ground growth. Treatment codes refer to both root and shoot interactions (B), no interactions (N), root interaction only (R), or shoot interactions only (S). The label *ns* indicates lack of statistical significance at the $\alpha = 0.05$ level in GLMMs, while lower case letters indicate post-hoc comparisons of differences among treatments at this $\alpha = 0.05$ level (Table 2).

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.29.318550; this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

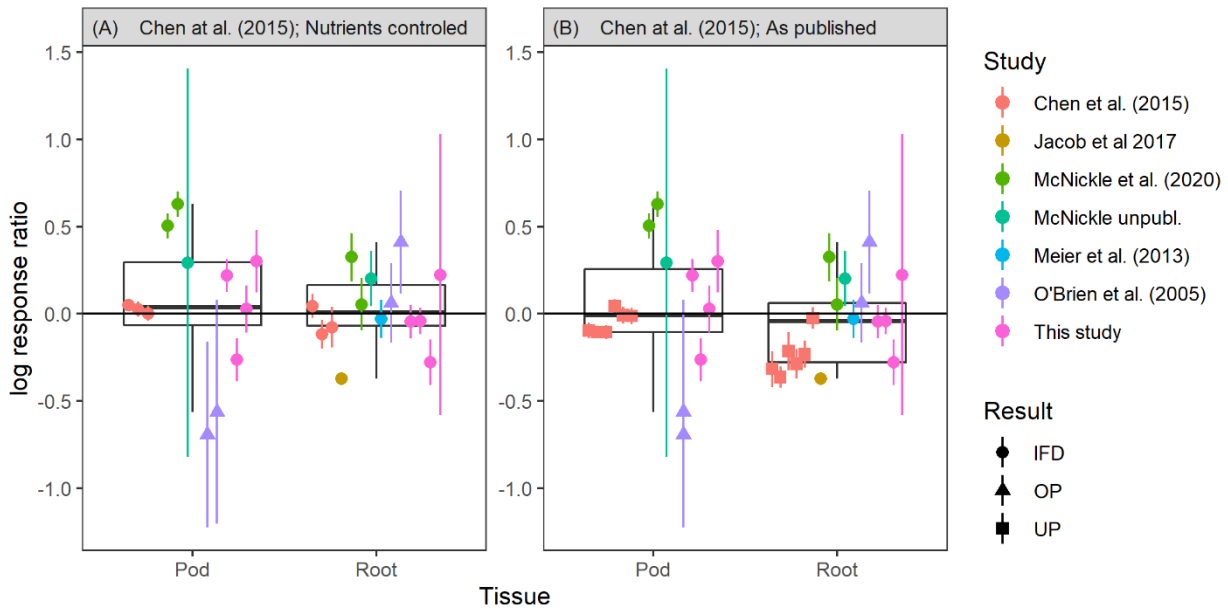

FIGURE 4: Summary of all available studies of pea responses to the presence of below ground neighbours. (A) results plotted with Chen et al. (2015) data analysed with nutrients controlled rather than pot volume. (B) results plotted with Chen et al. (2015) data presented as published, with pot volume controlled rather than nutrients. The log response ratio is the natural log of the biomass with neighbours to alone. Positive values for roots indicate support for the over-proliferation (OP) hypothesis, negative values support the under-proliferation (UP) hypothesis, and values equal to zero support the IFD hypothesis. The shape of each point represents the original conclusion (Result) of the published study, colour represents the reference of the original study, and error bars are 1 standard deviation of the log response ratio (which was not available for all studies).

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.29.318550; this version posted October 1, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

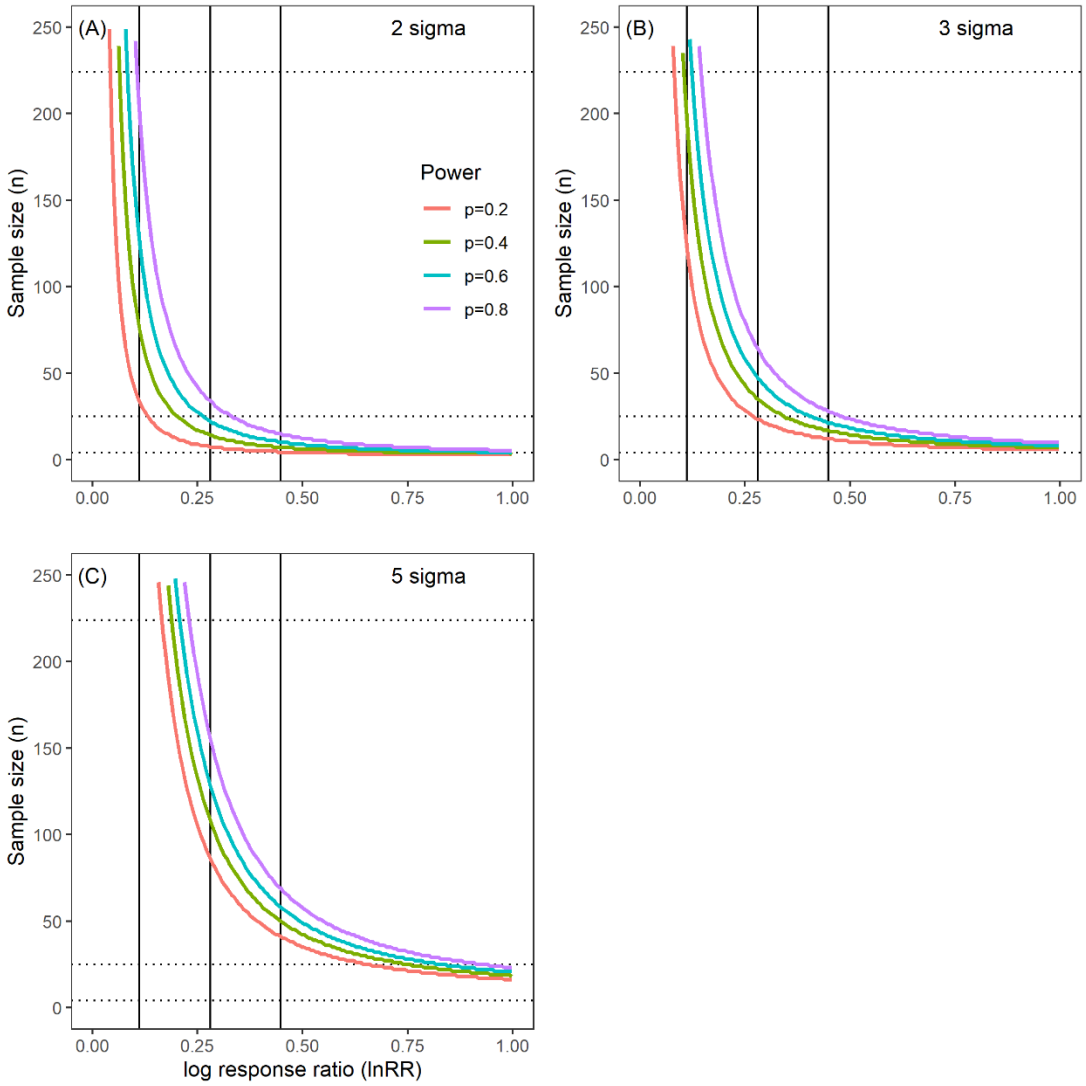

FIGURE 5: Power analysis for statistical significance using a different willingness to make a type I error: (A) 2 sigma, i.e. a 1 in 20 type I error rate, (B) 3 sigma, i.e. a 1 in 370 type I error rate and (C) five sigma, i.e. a 1 in 3.5million type I error rate. Coloured lines represent different levels of statistical power, that is, the ability to correctly reject a false null and avoid Type II statistical error. Horizontal dotted lines represent the minimum (n=4), and maximum (n=25) published sample sizes for pea, and n=224 the total number of replicates across the literature. Vertical lines represent conventional small, medium and high effect sizes from left to right respectively.