Letters

A common soil handling technique can generate incorrect estimates of soil biota effects on plants

Introduction

An active area of research seeks to understand how soil biota effects on plants vary across experimental factors (i.e. regions, treatments). The study biotas are obtained by gathering soil sample(s) from randomly selected location(s) within each experimental unit, with an experimental unit being a site within a study region or a field plot of a manipulative experiment. Then, plant growth is measured in glasshouse containers housing soil and biota from the various sites or plots. Results of these glasshouse bioassays are sensitive to a common soil handling decision. In particular, it is common to either: (1) fill each container with soil from one experimental unit (e.g. Callaway *et al.*, 2004; Hood *et al.*, 2004; Kardol *et al.*, 2006; Wardle *et al.*, 2012), or (2) fill each container with a mixture of soils from multiple experimental units (i.e. all sites within a region, all plots that received the same treatment) (e.g. Van der Putten *et al.*, 1993; Nijjer *et al.*, 2007; Felker-Quinn *et al.*, 2011; Pendergast *et al.*, 2013; Rodríguez-Echeverría *et al.*, 2013; Yang *et al.*, 2013; Gundale *et al.*, 2014; Pizano *et al.*, 2014; Hilbig & Allen, 2015; Larios & Suding, 2015) (Fig. 1). We define samples generated from these two approaches as 'individual soil samples' (ISS) and 'mixed soil samples' (MSS). The term 'individual soil sample' is slightly misleading, as ISS are often formed by mixing multiple samples gathered from the same experimental unit (i.e. pooling subsamples). Combining subsamples from individual experimental units is a perfectly acceptable approach. Conversely, the express purpose of this paper is to illustrate that, without exception, the MSS approach of mixing together soils from multiple experimental units (Fig. 1) is fatally flawed. Hypotheses regarding differences among regions or treatments cannot be legitimately tested by the MSS approach of mixing together soils from multiple sites within regions or multiple plots receiving the same treatment. The importance of this point is clearly underappreciated: we estimate, 52% of published studies use MSS in place of the correct ISS methodology (of 76 evaluated studies using ISS or MSS, 40 used MSS) (K. O. Reinhart & M. J. Rinella, unpublished, 2015).

Estimating treatment (e.g. region, nutrient) differences entails computing residual variance. Residual variance describes variation not explained by treatments, and it is needed to compute relevant statistics (i.e. *P*-values, confidence intervals). In experiments considered here, there are two contributors to residual variance in plant growth: (1) spatial variation in soil biotas (i.e. site-to-site variation not explained by region, plot-to-plot variation not explained by treatment) and (2) glasshouse variation owing to environmental gradients (e.g. temperature) and plant genetics. With the ISS approach (Fig. 1), having two contributors to residual variance poses no unique analytic challenges, and standard regression and analysis of variance (ANOVA) approaches give correct inferences. With the MSS approach, all information regarding residual variation in soil mutualists and pathogens is lost, and if this variation is nonzero, MSS and ISS are guaranteed to give different inferences, with the MSS inferences being incorrect. More specifically, if residual variation in mutualists and/or pathogens is nonzero, statistical estimates from MSS will be falsely precise and evidence for differences among treatments (e.g. regions, nutrients) will be weaker than reported. Assuming only factors being studied cause soil mutualists/pathogens to vary spatially is highly unrealistic, particularly given that plant disease expression (e.g. Martin & Loper, 1999) and soil microbe compositions (e.g. Ettema & Wardle, 2002; Ritz *et al.*, 2004) are known to vary widely across even small spatial scales (i.e. < 1.0 m).

Exploring effects of mixing soils

While inferences from MSS are *a priori* expected to be falsely precise (i.e. confidence intervals incorrectly narrow, *P*-values incorrectly small), the magnitude of the inaccuracy is data-dependent. Unfortunately, the magnitude of the inaccuracy arising from MSS cannot typically be quantified, because the necessary information (i.e. soil biota abundance data for each field unit and data describing relationship between soil biota abundance and plant performance) is not collected in most studies. Data from Reinhart & Clay (2009) and Reinhart *et al.* (2010a) provide a rare opportunity to evaluate the magnitude of inferential inaccuracies arising from MSS approaches. Their data describe site-to-site variation in soil pathogen (*Pythium* spp.) densities (propagules/g of soil) and the relationship between these densities and biomass of the host tree *Prunus serotina* in a growth chamber. *Pythium* densities were measured at 10 eastern US sites separated by a maximum distance of 700 km. While *P. serotina* was subjected to only one *Pythium* species in the growth chamber (i.e. *P. sylvaticum*), we believe *P. sylvaticum* is a suitable proxy for what may have been multiple *Pythium* species sampled from the field (Reinhart *et al.*, 2010a). *Pythium sylvaticum* is a widely distributed pathogen with effects on *P. serotina* comparable to other *Pythium* spp. (Reinhart *et al.*, 2010b).

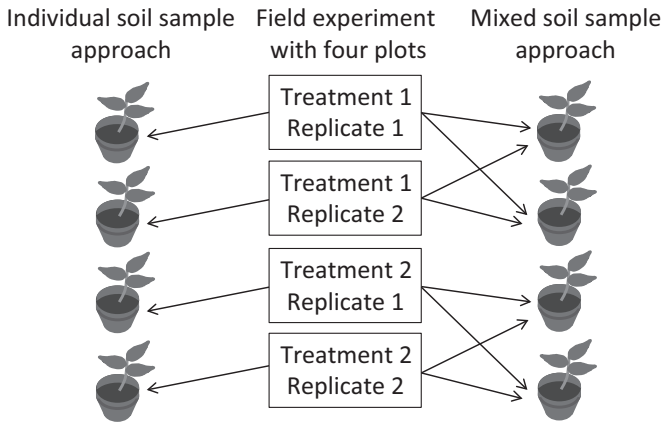

Fig. 1 Diagram of two approaches for testing effects of soil biota on plant growth. Arrows indicate movement of soil from field plots to glasshouse containers.

Simulated experiments

We used the data described in the previous paragraph to parameterize 100 simulated experiments. The experiments were designed to estimate the pathogenicity difference (i.e. plant biomass difference) between two regions, Region 1 and Region 2. Therefore, our simulation is particularly relevant to recent studies that used MSS to test whether soil biota effects varied between two regions (Yang *et al.*, 2013; Gundale *et al.*, 2014). The true pathogenicity difference between Region 1 and Region 2 was zero: for both regions, the pathogen density distribution was that estimated from the 10 sites of Reinhart & Clay (2009) and Reinhart *et al.* (2010a). Because the pathogen density distribution was equivalent for both regions, a proper scheme for handling soils and analyzing data should tend to suggest no pathogenicity difference between regions. For each simulated experiment we followed steps 1–4:

Step 1 Simulated pathogen densities (propagules/g of soil) occurring at 10 sites in Region 1 and 10 sites in Region 2. This entailed drawing 10 values from the pathogen density probability distribution and assigning them to Region 1 and then sampling 10 additional values from the same distribution and assigning them to Region 2.

Step 2 Simulated one glasshouse container pathogen density value for each site (2 regions × 10 sites = 20 values). This was done in two ways corresponding to ISS and MSS. For ISS, the field site and glasshouse container values were identical. For MSS, pathogen densities for Region 1 and 2 containers were set equal to regional means. This replacing of raw values with regional means reflected mixing equal quantities of soil from each site within a region and filling containers with the mixture.

Step 3 Simulated one plant biomass value for each glasshouse container. Plant weights were a function of container pathogen densities (Eqn 1).

Step 4 Computed 95% confidence intervals quantifying the plant biomass difference between Region 1 and Region 2. Steps 1–4 were repeated 100 times.

Parameterizing the simulation

This section describes technical details of parameterizing the simulation and may be skipped by readers seeking a general understanding. *Pythium* spp. were found at only three of 10 sites, and at these sites, the mean and standard deviation of pathogen density was 3.9 and 1.8 log(propagules/g of soil), respectively. Correspondingly, our system for simulating pathogen densities for regions (Step 1) was to generate 10 *Pythium* presence/absence values with presence probability 0.3. (Our results were similar in an additional simulation that assumed *Pythium* occurred at all sites.) For sites simulated to have *Pythium*, *Pythium* density was simulated as Exp[N(3.9, 1.8)] where $N(\mu,\sigma)$ denotes the normal distribution with, mean μ, standard deviation σ. The growth chamber biomass data from 22 pots of Reinhart *et al.* (2010a) were well-approximated by:

$$\log y_i \sim N(-1.8 - 0.2 \times \log[x_i + 1], 0.13), \quad \text{Eqn 1}$$

where x_i is *P. sylvaticum* density (propagules/g of soil) and y_i is mean *P. serotina* shoot biomass (g/plant) per pot in pot $i = 1, \ldots, 22$, respectively.

Computing the confidence intervals (Step 4) involved fitting the following simple linear regression model to each simulated dataset:

$$\log z_i \sim N(B_0 + B_1 r_i, \sigma), \quad \text{Eqn 2}$$

where z_i is simulated *P. serotina* shoot biomass (g/plant) for pot $i = 1, \ldots, 20$, and r_i is an indicator equaling 0.0 for Region 1 and 1.0 for Region 2. It is easy to show that B_1 is the following log response ratio:

$$\log\left(\frac{\text{region 1 mean biomass/plant}}{\text{region 2 mean biomass/plant}}\right). \quad \text{Eqn 3}$$

The analytical summaries of interest in Step 4 were 95% confidence intervals on this log ratio. Negative log response ratios indicate plant biomass was lower with soil originating from Region 1 than Region 2 and vice versa for positive values. Mathematica 9 code (Wolfram Research Inc., Champaign, IL, USA) for conducting the simulation is provided (Supporting Information Notes S1).

Simulation results

It is critical to recall that the simulation was set up such that there was no pathogenicity difference between the two regions. Thus, for any given simulated experiment, a proper 95% confidence interval on the soil origin effect has a 95% chance of overlapping the zero line (Fig. 2). It follows that if proper methods are used, there should be a tendency to detect no difference in pathogenicity between regions, and 95 of our 100 simulated confidence intervals should overlap zero. (Because of sampling variability, the number need not be exactly 95.) A confidence interval that fails to overlap zero represents a false rejection of the null hypothesis of no difference between regions (i.e. type I error). With the ISS approach, 97 of 100

New Phytologist

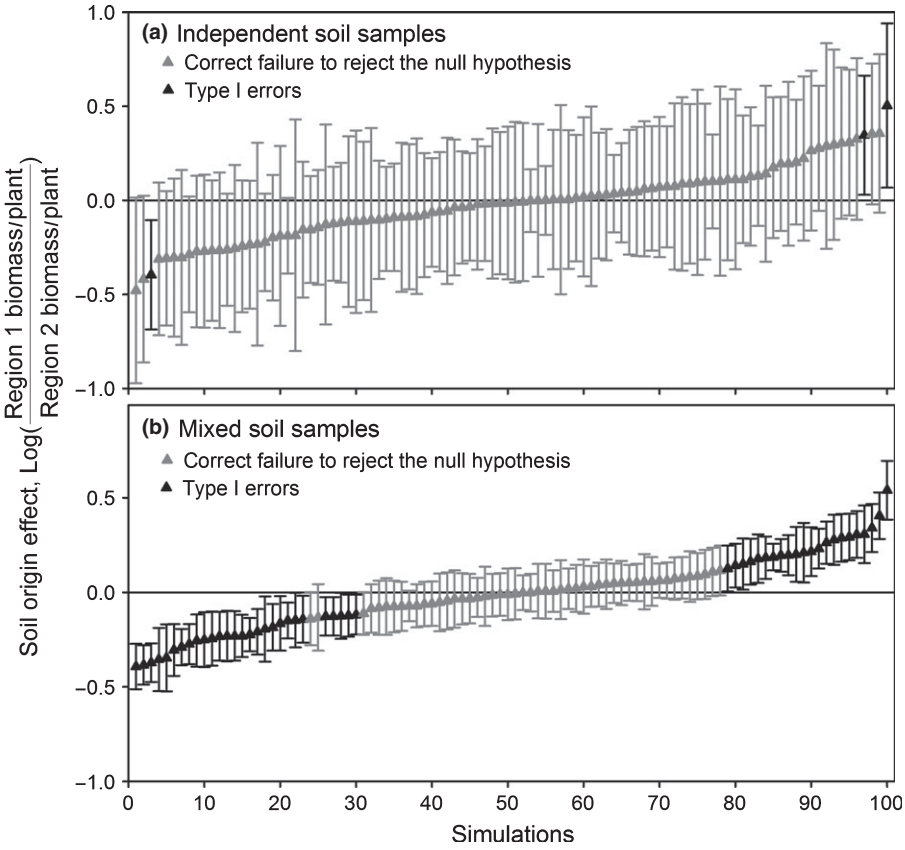

Fig. 2 Point estimates and 95% confidence intervals on log response ratios from 100 simulated soil biota experiments. The simulated experiments were parameterized from real data, and they compared how plant growth responses to pathogens differed between two regions. The true difference between simulated regions was zero, so proper experimental and analytical methods should generate confidence intervals with a 95% chance of overlapping zero (i.e. *c.* 95 of the 100 confidence intervals should overlap zero). (a) Results from proper methods in which each plant was subjected to soil from one of 20 simulated sites, 10 from each region. (b) Results obtained by mixing together soils from all 10 sites within a region and subjecting plants to these regional mixtures. This latter approach led to falsely narrow confidence intervals that grossly overestimated evidence for regional differences.

confidence intervals overlapped zero (Fig. 2a), thereby indicating the reliability of this approach. Conversely, the MSS approach performed very poorly, leading to a false rejection of the null hypothesis in 50 of 100 experiments (Fig. 2b). This poor performance occurred because mixing soil samples eliminated site-to-site variation in soil biota, thereby leading to falsely precise estimates. This can be seen by comparing the narrow MSS confidence intervals with the broader ISS confidence intervals (Fig. 2).

Conclusions

Soil handling decisions may dramatically affect conclusions of experiments seeking to quantify effects of soil biota on plants. Our results lead us to question results of replicated field studies that use MSS approaches to assess effects of soil biota on plant growth (e.g. Van der Putten *et al.*, 1993; Nijjer *et al.*, 2007; Felker-Quinn *et al.*, 2011; Pendergast *et al.*, 2013; Rodríguez-Echeverría *et al.*, 2013; Yang *et al.*, 2013; Gundale *et al.*, 2014; Pizano *et al.*, 2014; Hilbig & Allen, 2015; Larios & Suding, 2015). In the types of studies we describe in this paper, a one-to-one correspondence should be maintained between field units (i.e. plots, sites) and glasshouse containers. Finally, for the reasons described in this paper, inferences from all studies (e.g. soil chemistry, soil microbial community composition), not just soil biota effects studies, are nearly certain to be invalid if they are derived by performing tests on mixtures of soils from multiple experimental units.

Author contributions

K.O.R. and M.T.R. identified the methodological problem, M.T.R. designed the simulation and performed the analysis, K.O.R. provided data to parameterize the simulation, and M.T.R. and K.O.R. wrote the manuscript.

Kurt O. Reinhart* and Matthew J. Rinella

United States Department of Agriculture – Agricultural Research Service, Fort Keogh Livestock & Range Research Laboratory, 243 Fort Keogh Road, Miles City, MT 59301-4016, USA
(*Author for correspondence: tel +1 406 874 8211; email kurt.reinhart@ars.usda.gov)

References

Callaway RM, Thelen G, Rodriguez A, Holben WE. 2004. Soil biota and exotic plant invasion. *Nature* **427**: 731–733.

Ettema CH, Wardle DA. 2002. Spatial soil ecology. *Trends in Ecology & Evolution* **17**: 177–183.

Felker-Quinn E, Bailey JK, Schweitzer JA. 2011. Soil biota drive expression of genetic variation and development of population-specific feedbacks in an invasive plant. *Ecology* **92**: 1208–1214.

Gundale MJ, Kardol P, Nilsson MC, Nilsson U, Lucas RW, Wardle DA. 2014. Interactions with soil biota shift from negative to positive when a tree species is moved outside its native range. *New Phytologist* **202**: 415–421.

Hilbig BE, Allen EB. 2015. Plant-soil feedbacks and competitive interactions between invasive *Bromus diandrus* and native forb species. *Plant and Soil* **392**: 191–203.

Hood LA, Swaine MD, Mason PA. 2004. The influence of spatial patterns of damping-off disease and arbuscular mycorrhizal colonization on tree seedling establishment in Ghanaian tropical forest soil. *Journal of Ecology* **92**: 816–823.

Kardol P, Bezemer TM, Van der Putten WH. 2006. Temporal variation in plant–soil feedback controls succession. *Ecology Letters* **9**: 1080–1088.

Larios L, Suding KN. 2015. Competition and soil resource environment alter plant–soil feedbacks for native and exotic grasses. *AoB Plants* **7**: plu077.

Martin FN, Loper JE. 1999. Soilborne plant diseases caused by *Pythium* spp.: ecology, epidemiology, and prospects for biological control. *Critical Reviews in Plant Sciences* **18**: 111–181.

Nijjer S, Rogers WE, Siemann E. 2007. Negative plant-soil feedbacks may limit persistence of an invasive tree due to rapid accumulation of soil pathogens. *Proceedings of the Royal Society of London Series B: Biological Sciences* **274**: 2621–2627.

Pendergast TH, Burke DJ, Carson WP. 2013. Belowground biotic complexity drives aboveground dynamics: a test of the soil community feedback model. *New Phytologist* **197**: 1300–1310.

Pizano C, Mangan SA, Graham JH, Kitajima K. 2014. Habitat-specific positive and negative effects of soil biota on seedling growth in a fragmented tropical montane landscape. *Oikos* **123**: 846–856.

Reinhart KO, Clay K. 2009. Spatial variation in soil-borne disease dynamics of a temperate tree, *Prunus serotina*. *Ecology* **90**: 2984–2993.

Reinhart KO, Royo AA, Kageyama SA, Clay K. 2010a. Canopy gaps decrease microbial densities and disease risk for a shade-intolerant tree species. *Acta Oecologica* **36**: 530–536.

Reinhart KO, Tytgat T, Van der Putten WH, Clay K. 2010b. Virulence of soil-borne pathogens and invasion by *Prunus serotina*. *New Phytologist* **186**: 484–495.

Ritz K, McNicol JW, Nunan N, Grayston S, Millard P, Atkinson D, Gollotte A, Habeshaw D, Boag B, Clegg CD *et al.* 2004. Spatial structure in soil chemical and microbiological properties in an upland grassland. *FEMS Microbiology Ecology* **49**: 191–205.

Rodríguez-Echeverría S, Armas C, Pistón N, Hortal S, Pugnaire FI. 2013. A role for below-ground biota in plant-plant facilitation. *Journal of Ecology* **101**: 1420–1428.

Van der Putten WH, Van Dijk C, Peters BAM. 1993. Plant-specific soil-borne diseases contribute to succession in foredune vegetation. *Nature* **362**: 53–56.

Wardle DA, Bellingham PJ, Fukami T, Bonner KI. 2012. Soil-mediated indirect impacts of an invasive predator on plant growth. *Biology Letters* **8**: 574–577.

Yang Q, Carrillo J, Jin H, Shang L, Hovick SM, Nijjer S, Gabler CA, Li B, Siemann E. 2013. Plant–soil biota interactions of an invasive species in its native and introduced ranges: implications for invasion success. *Soil Biology & Biochemistry* **65**: 78–85.

Supporting Information

Additional supporting information may be found in the online version of this article.

Notes S1 Mathematica code for simulation.

Please note: Wiley Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.

Key words: microbial ecology, plant–soil biota experiment, plant–soil feedback, soil handling technique, type I error.