Supplementary materials for “A global synthesis of soil microbial effects on plant species coexistence”

Authors: Xinyi Yan\textsuperscript{1,*}, Jonathan M. Levine\textsuperscript{2}, Gaurav S. Kandlikar\textsuperscript{3,4,*}

\textsuperscript{1} Department of Integrative Biology, University of Texas at Austin
\textsuperscript{2} Department of Ecology and Evolutionary Biology, Princeton University
\textsuperscript{3} Division of Biological Sciences, University of Missouri, Columbia
\textsuperscript{4} Division of Plant Sciences and Technology, University of Missouri, Columbia

* These authors contributed equally.

Authors for correspondence:
Xinyi Yan: xinyiyan@utexas.edu
Gaurav S. Kandlikar: gkandlikar@missouri.edu

This supplemental PDF file includes:

1. Supplemental figures and tables (jump to page);
2. Appendix S1: PRISMA chart for meta-analysis (jump to page);
3. Appendix S2: Effect size derivations (jump to page);

Code and data for reproducing all analyses have been deposited at: doi.org/10.5281/zenodo.6513066
Figure S.1: In the four outcomes of the Main Text Fig. 1, we show that plant-soil feedbacks can result in priority effects, coexistence, or competitive exclusion when soil microbes generally suppress plant growth (negative values of $m_i\chi$). Here, we show that all four outcomes even soil microbes promote the growth of both plant species (positive values of $m_i\chi$).
Figure S.2: Likelihood profiles for Q1 model parameters, generated with `metafor::profile()`

(a) live reference soil

(b) sterile reference soil
Figure S.3: Moderator analyses for Q1. Soil points indicate the overall effect sizes for pairs that were studied using the relevant methodological approach, and whiskers indicate the 95% confidence intervals. Significance brackets indicate those pairwise (De)stabilization/Fitness difference comparisons that are significantly different from one another.
Figure S.4: Comparison of (de)stabilization and fitness difference effect sizes (Main text Q1) when analyzed through multivariate models vs. through separate univariate models of the two effect sizes.

Figure S.5: Comparison of (de)stabilization and fitness difference effect sizes (Main text Q1) when analyzed with variance-covariance matrices derived assuming different strengths of correlations between (de)stabilization and fitness differences among species pairs.
Figure S.6: The effect of excluding influential studies. Kandlikar et al. 2019 and Miller et al. 2015 were identified as influential experiments in the live reference dataset, and Hendriks et al. 2013 in the sterile reference dataset. Fitness difference remains significantly larger than (de)stabilization in both reference datasets after removing these studies.
Figure S.7: Effect sizes from meta-analyses of species pairs with stabilizing (A) or destabilizing (B) effects of microbes. Microbially mediated fitness differences are stronger than stabilizing effects, regardless of whether the reference growth was of plants in sterile or live soils (panel A). Among species pairs where microbes destabilize interactions, fitness differences were stronger than destabilizing effects only among plants whose growth in sterile soils provided the reference comparison (panel B). Large points indicate the overall effect sizes; black lines show the 95% confidence interval.
Figure S.8: **Schematic that explains our reasoning behind the species-pair analyses in Q2 with simulated data.** In principle, only comparing the meta-analytical mean strength of stabilization and fitness difference can obscure the effect of microbes on plant competition across individual species pairs. In this simulation, when microbes drive stronger mean fitness differences than stabilization (non-overlapping 95% CI in panel A), which predict species exclusion on average. But, when evaluated across individual species pairs, we see that microbes favor coexistence among a substantial fraction of species pairs because of a negative correlation in the strength of stabilization and fitness differences (panel B).
Figure S.9: Orchard plot showing the results from a meta-analysis of the destabilizing and stabilizing microbial effects across the entire dataset.
Table S1: Proportion of simulation runs (Q2) that resulted in Coexistence/Priority effects/Exclusion among pairs for which growth in uncultivated live soils provide the reference. Top number indicates median, and the range indicates the 95% confidence interval.

| category              | Method     | Coexistence | Priority effects | Exclusion  |
|-----------------------|------------|-------------|------------------|------------|
| Overall               | Overall    | 22.2%       | 6.9%             | 69.4%      |
|                       |            | (16.7%-29.2%)| (4.2%-11.1%)     | (62.5%-76.4%) |
| Inoculation fraction  | Low        | 17.6%       | 5.9%             | 76.5%      |
|                       |            | (5.9%-29.4%)| (0.0%-17.6%)     | (58.8%-88.2%) |
|                       | High       | 23.6%       | 7.3%             | 67.3%      |
|                       |            | (18.2%-30.9%)| (3.6%-12.7%)     | (60.0%-76.4%) |
| Training phase location | Field    | 28.6%       | 21.4%            | 50.0%      |
|                       |            | (14.3%-42.9%)| (7.1%-28.6%)     | (35.7%-71.4%) |
|                       | Greenhouse | 22.4%       | 3.4%             | 74.1%      |
|                       |            | (15.5%-27.6%)| (1.7%-8.6%)      | (67.2%-81.0%) |
| Response phase growth type | Individual | 23.6%       | 7.3%             | 69.1%      |
|                       |            | (16.4%-30.9%)| (3.6%-12.7%)     | (60.0%-76.4%) |
|                       | Population | 0.0%        | 0.0%             | 50.0%      |
|                       |            | (0.0%-50.0%)| (0.0%-50.0%)     | (0.0%-100.0%) |
|                       | Community  | 20.0%       | 6.7%             | 73.3%      |
|                       |            | (6.7%-33.3%)| (0.0%-13.3%)     | (60.0%-86.7%) |
Table S2: Proportion of simulation runs (Q2) that resulted in Coexistence/Priority effects/Exclusion among pairs for which growth in sterile soils provide the reference. Top number indicates median, and the range indicates the 95% confidence interval.

| category                              | Method               | Coexistence | Priority effects | Exclusion          |
|----------------------------------------|----------------------|-------------|------------------|--------------------|
| Overall                                | Overall              | 12.8%       | 6.1%             | 81.2%              |
|                                        |                      | (10.8%-14.8%)| (4.5%-7.6%)      | (78.7%-83.6%)      |
| Inoculation fraction                   | Low                  | 11.6%       | 6.7%             | 81.8%              |
|                                        |                      | (9.4%-14.0%)| (4.9%-8.5%)      | (79.0%-84.5%)      |
|                                        | High                 | 16.2%       | 4.3%             | 79.5%              |
|                                        |                      | (12.8%-20.5%)| (1.7%-6.8%)      | (75.2%-83.8%)      |
| Training phase location                | Field                | 12.4%       | 7.3%             | 79.7%              |
|                                        |                      | (9.6%-15.8%)| (4.5%-10.2%)     | (75.7%-83.6%)      |
|                                        | Greenhouse           | 9.4%        | 4.7%             | 86.0%              |
|                                        |                      | (7.2%-11.5%)| (2.6%-6.8%)      | (83.0%-88.9%)      |
| Response phase growth type             | Individual           | 21.2%       | 7.9%             | 70.9%              |
|                                        |                      | (17.5%-24.9%)| (5.3%-11.1%)     | (66.1%-75.1%)      |
|                                        | Population           | 8.1%        | 5.4%             | 86.5%              |
|                                        |                      | (2.7%-13.5%)| (2.6%-13.5%)     | (78.4%-94.6%)      |
|                                        | Community            | 6.4%        | 4.1%             | 89.5%              |
|                                        |                      | (4.5%-8.6%) | (2.3%-5.9%)      | (86.8%-92.3%)      |
| Sterilization method                   | Autoclaving          | 22.9%       | 9.3%             | 67.8%              |
|                                        |                      | (17.8%-28.0%)| (5.9%-13.6%)     | (61.9%-73.7%)      |
|                                        | Gamma irradiation    | 14.8%       | 4.1%             | 80.3%              |
|                                        |                      | (11.5%-18.9%)| (1.6%-7.4%)      | (76.2%-84.4%)      |
|                                        | Heat                 | 4.2%        | 4.2%             | 91.5%              |
|                                        |                      | (2.1%-6.3%) | (2.6%-6.9%)      | (88.4%-94.2%)      |
|                                        | Other                | 23.5%       | 11.8%            | 70.6%              |
|                                        |                      | (11.8%-35.3%)| (0.0%-23.5%)     | (52.9%-82.4%)      |
| Type of soil sterilized                | Greenhouse soil      | 6.0%        | 5.1%             | 88.9%              |
|                                        |                      | (3.7%-8.3%) | (2.8%-7.4%)      | (86.1%-92.1%)      |
|                                        | Field soil           | 6.0%        | 5.1%             | 88.9%              |
|                                        |                      | (3.7%-8.3%) | (2.8%-7.4%)      | (86.1%-92.1%)      |
|                                        | Conditioned soil     | 16.9%       | 6.7%             | 76.4%              |
|                                        | (species-specific)   | (13.3%-20.5%)| (4.1%-9.2%)      | (72.8%-80.5%)      |
|                                        | Conditioned soil     | 16.9%       | 6.7%             | 76.4%              |
|                                        | (mixed)              | (13.3%-20.5%)| (4.1%-9.2%)      | (72.8%-80.5%)      |
Appendix 1: PRISMA chart

Figure S.10: PRISMA chart

From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71. For more information, visit: http://www.prisma-statement.org/
Appendix 2: Derivation of effect sizes and variance terms

Here we show the derivation of the microbially mediated stabilization (or destabilization) and fitness difference effect sizes and their variances.

Effect sizes: Stabilization & Fitness Differences

Following (1), the stabilization (or destabilization) and fitness difference of plant species 1 and 2 mediated by their microbial communities A and B can be expressed as the following:

\[
\text{Stabilization} = -\frac{1}{2} (m_{1A} - m_{1B} - m_{2A} + m_{2B}) \quad (1)
\]
\[
\text{Fitness Difference} = \frac{1}{2} [(m_{1A} + m_{1B}) - (m_{2A} + m_{2B})] \quad (2)
\]

By operationalizing each microbial effect \((m_{iX})\) as the natural log growth of plant species \(i\) on microbial community \(X\) minus its log growth in a reference soil community \((\ln(G_{iX}) - \ln(G_{iR}))\), the above equations can be expanded. The reference growth terms cancel out when calculating the Stabilization metric, but not for the Fitness Difference:

\[
\text{Stabilization} = -\frac{1}{2} \left[ (\ln(G_{1A}) - \ln(G_{1B})) - (\ln(G_{1B}) - \ln(G_{1R})) - \\
(\ln(G_{2A}) - \ln(G_{2B})) + (\ln(G_{2B}) - \ln(G_{2R})) \right] \quad (3)
\]
\[
= -\frac{1}{2} [\ln(G_{1A}) - \ln(G_{1B}) - \ln(G_{2A}) + \ln(G_{2B})]
\]
\[
\text{Fitness Difference} = \frac{1}{2} \left[ (\ln(G_{1A}) - \ln(G_{1R})) + (\ln(G_{1B}) - \ln(G_{1R})) - \\
(\ln(G_{2A}) - \ln(G_{2R})) - (\ln(G_{2B}) - \ln(G_{2R})) \right] \quad (4)
\]
\[
= \frac{1}{2} [\ln(G_{1A}) + \ln(G_{1B}) - \ln(G_{2A}) - \ln(G_{2B})] - \ln(G_{1R}) + \ln(G_{2R})
\]

Both summations can also be written in the log response ratio form:

\[
\text{Stabilization} = -\frac{1}{2} \left[ \ln \left( \frac{G_{1A}}{G_{1B}} \right) + \ln \left( \frac{G_{2B}}{G_{2A}} \right) \right] \quad (5)
\]
\[
= -\frac{1}{2} \ln \left( \frac{G_{1A}G_{2B}}{G_{1B}G_{2A}} \right)
\]

Yan et al.
Fitness Difference \(= \frac{1}{2}\left[\ln(G_{1A}) - \ln(G_{1R}) + \ln(G_{1B}) - \ln(G_{1R}) - \ln(G_{2A}) - \ln(G_{2R}) - \ln(G_{2B}) - \ln(G_{2R})\right]\)

\(= \frac{1}{2}\ln\left(\frac{G_{1A}}{G_{1R}}\right) + \ln\left(\frac{G_{1B}}{G_{1R}}\right) - \ln\left(\frac{G_{2A}}{G_{2R}}\right) - \ln\left(\frac{G_{2B}}{G_{2R}}\right)\)  

(6)

Note that the \(rr\) (Stabilization) is \(-\frac{1}{2}\) of the \(rr\) (\(I_S\)) effect size used by Crawford et al., and can likewise be interpreted as the difference of the two plant’s species in their responses to conspecific and heterospecific soils. The \(rr\) (Fitness Difference) is the ratio of the average mictobially mediated fitness of plant species 1 over that of plant species 2.

**Variance of the (de)stabilization & fitness differences**

The variance of the effect sizes can be calculated using the following equation ((2), (3)):

\[\sigma^2(\text{Stabilization/fitness Difference}) = A^T\Sigma A\]  

(7)

where \(A\) is the column vector of partial derivatives of Stabilization (SD) or Fitness Difference (FD) with respect to each of its \(G\) terms.

Therefore,

\[A_{SD}^T = \frac{1}{2}\left[\frac{\partial(\text{de})st}{\partial G_{1A}}, \frac{\partial(\text{de})st}{\partial G_{1B}}, \frac{\partial FD}{\partial G_{2A}}, \frac{\partial FD}{\partial G_{2B}}\right] = \left[\frac{1}{G_{1A}}, -\frac{1}{G_{1B}}, \frac{1}{G_{2A}}, \frac{1}{G_{2B}}\right]\]  

(8)

\[A_{FD}^T = \frac{1}{2}\left[\frac{\partial FD}{\partial G_{1A}}, \frac{\partial FD}{\partial G_{1B}}, \frac{\partial FD}{\partial G_{2A}}, \frac{\partial FD}{\partial G_{2B}}\right] = \left[\frac{1}{G_{1A}}, -\frac{1}{G_{1B}}, -\frac{1}{G_{2A}}, -\frac{1}{G_{2B}}\right]\]  

(9)

Assuming that the growth terms \((G_{iX})\) are independent of each other, the large-sample variance covariance matrix for niche difference \(\Sigma_{SD}\) is simply a diagonal matrix; However, due to the repeated terms for growth in reference soils \((G_{iR})\) in the equation for fitness difference, \(\Sigma_{FD}\) contains covariance terms between the duplicated controls (see Appendix Fig. A1b in ref (3)).
Note that the above variance for stabilization is 1 of that for $I_S$ used by Crawford et al. Those variances can also be calculated following the statistical definition of standard error of the mean. Here we show the derivation for fitness difference:

Putting together, the respective variance for the stabilization and fitness difference matrices are:

$$
\sigma_{SD}^2 = \frac{1}{4} \left( \frac{\sigma_{1A}^2}{G_{1A}^2 N_{1A}} + \frac{\sigma_{1B}^2}{G_{1B}^2 N_{1B}} + \frac{\sigma_{2A}^2}{G_{2A}^2 N_{2A}} + \frac{\sigma_{2B}^2}{G_{2B}^2 N_{2B}} \right)
$$

(10)

$$
\sigma_{FD}^2 = \frac{1}{4} \left( \frac{\sigma_{1A}^2}{G_{1A}^2 N_{1A}} + \frac{\sigma_{1B}^2}{G_{1B}^2 N_{1B}} + \frac{\sigma_{2A}^2}{G_{2A}^2 N_{2A}} + \frac{\sigma_{2B}^2}{G_{2B}^2 N_{2B}} + 4 \frac{\sigma_{2R}^2}{G_{2R}^2 N_{2R}} \right)
$$

(11)

Note that the above variance for stabilization is 1 of that for $I_S$ used by Crawford et al. Those variances can also be calculated following the statistical definition of standard error of the mean. Here we show the derivation for fitness difference:

$$
\sigma_{FD}^2 = \sigma^2 \left( \frac{1}{2} \left[ \ln(G_{1A}) + \ln(G_{1B}) - 2 \ln(G_{1R}) - \ln(G_{2A}) - \ln(G_{2B}) + 2 \ln(G_{2R}) \right] \right)
$$

$$
= \frac{1}{4} \left[ \frac{\sigma^2(\ln(G_{1A}))}{N_{1A}} + \frac{\sigma^2(\ln(G_{1B}))}{N_{1B}} + \frac{\sigma^2(\ln(G_{1R}))}{N_{1R}} + \frac{\sigma^2(\ln(G_{2A}))}{N_{2A}} + \frac{\sigma^2(\ln(G_{2B}))}{N_{2B}} + 4 \frac{\sigma^2(\ln(G_{2R}))}{N_{2R}} \right]
$$

and using propagation error theory to get the variance of a natural log of a random variable:

$$
= \frac{1}{4} \left[ \frac{\sigma^2(G_{1A})}{N_{1A} G_{1A}^2} + \frac{\sigma^2(G_{1B})}{N_{1B} G_{1B}^2} + \frac{\sigma^2(G_{1R})}{N_{1R} G_{1R}^2} + \frac{\sigma^2(G_{2A})}{N_{2A} G_{2A}^2} + \frac{\sigma^2(G_{2B})}{N_{2B} G_{2B}^2} + 4 \frac{\sigma^2(G_{2R})}{N_{2R} G_{2R}^2} \right]
$$

(12)
Do we expect larger mean value for Fitness Difference than Stabilization?

Based on the equations for calculating stabilization and fitness differences, we would expect larger variance for abs(Fitness Difference) than for abs(Stabilization). This is because of the extra variance contributed by the reference growth data to the Fitness Difference, which is not included in the Stabilization calculation (comparing Eqn. 6 and 7 in the main text). At least qualitatively, this is what we observe in the data (e.g. see the larger confidence intervals for abs(Fitness Difference) in Fig. 3 in the main text).

We wanted to confirm that we should not expected any such systematic/statistical tendency for the mean magnitude of stabilization and fitness differences. We explored two scenarios to confirm that this is the case.

First, we considered the null case where plant growth is unaffected by the composition of the soil community. In this case, all $G_{iX}$ terms in equation 4 and 5 have the same expected value, and $\text{abs(Fitness difference)} = \text{abs(Stabilization)} = \frac{1}{2} \ln(1)$. As expected, the two metrics have the exact same value.

Next, we considered a species pair where both species have the same growth in the reference soil ($G_{1,\text{Ref}} = G_{2,\text{Ref}}$), but respond differently to cultivated microbes. The two metrics become: $\text{abs(Fitness difference)} = \frac{1}{2} \ln(\frac{G_{1A}G_{1B}}{G_{2A}G_{2B}})$ and $\text{abs(Stabilization)} = \frac{1}{2} \ln(\frac{G_{1A}G_{2B}}{G_{1B}G_{2A}})$. Any difference between the two metrics will be based on the biological differences in how plants grow with the cultivated microbes, rather than due to any statistical tendency.

We found this to be true in all other cases we explored - while the two metrics can vary considerably, these differences arise directly due to systematic differences in how plants grow in different microbial contexts rather than due to any statistical tendency or mathematical artifact.

References for Appendix 2

1. G. S. Kandlikar, C. A. Johnson, X. Yan, N. J. B. Kraft, J. M. Levine, Winning and losing with microbes: How microbially mediated fitness differences influence plant diversity. *Ecology Letters* 22, 1178–1191 (2019).

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

3. M. J. Lajeunesse, On the meta-analysis of response ratios for studies with correlated and multi-group designs. *Ecology* 92, 2049–2055 (2011).