Supplementary Information for:

The global human impact on biodiversity

François Keck[*,1,2], Tianna Peller[1,2], Roman Alther[1,2], Cécilia Barouillet[3,4], Rosetta Blackman[1,2], Éric Capo[5], Teofana Chonova[6], Marjorie Couton[1,2], Lena Fehlinger[7], Dominik Kirschner[2,8,9], Mara Knüsel[1,2], Lucile Muneret[10,11], Rebecca Oester[1,2,12], Kálmán Tapolczai[13,14], Heng Zhang[1,2], Florian Altermatt[*,1,2]

[1] Department of Evolutionary Biology and Environmental Studies, University of Zurich, Winterthurerstr. 190, CH-8057 Zürich, Switzerland

[2] Eawag, Swiss Federal Institute of Aquatic Science and Technology, Department of Aquatic Ecology, Überlandstrasse 133, CH-8600 Dübendorf, Switzerland

[3] INRAE, Université Savoie Mont Blanc, CARRTEL, 74200, Thonon-les-Bains, France

[4] Pôle R&D ECLA, CARRTEL, 74200, Thonon-les-Bains, France

[5] Department of Ecology and Environmental Science, Umeå University, Linnaeus väg 4-6, 907 36, Umeå, Sweden

[6] Eawag, Swiss Federal Institute of Aquatic Science and Technology, Department Environmental Chemistry, Überlandstrasse 133, CH-8600 Dübendorf, Switzerland

[7] GEA Aquatic Ecology Group, University of Vic - Central University of Catalonia, Carrer de la Laura 13, 08500 Vic, Spain

[8] Department of Environmental Systems Science, Institute of Terrestrial Ecosystems, Ecosystems and landscape evolution, ETH Zürich, CH-8092, Zürich, Switzerland

[9] Department of Landscape Dynamics & Ecology, Swiss Federal Research Institute WSL, Zürcherstrasse 111, 8903 Birmensdorf, Switzerland.

[10] INRAE, Université Paris-Saclay, AgroParisTech, UMR Agronomie, 91120, Palaiseau, France

[11] INRAE, Agroécologie, Institut Agro, Univ. Bourgogne, Univ. Bourgogne Franche-Comté, F-21000, Dijon, France

[12] Institute of Microbiology, University of Applied Sciences and Arts of Southern Switzerland, via Flora Ruchat Roncati 15, CH-6850 Mendrisio, Switzerland

[13] HUN-REN Balaton Limnological Research Institute, Klebelsberg Kuno Street 3, H-8237 Tihany, Hungary

[14] National Laboratory for Water Science and Water Security, HUN-REN Balaton Limnological Research Institute, Klebelsberg Kuno Street 3, H-8237 Tihany, Hungary

[*] Corresponding authors

Supplementary Information

Supplementary Methods

S1. Stratified analysis: observational studies

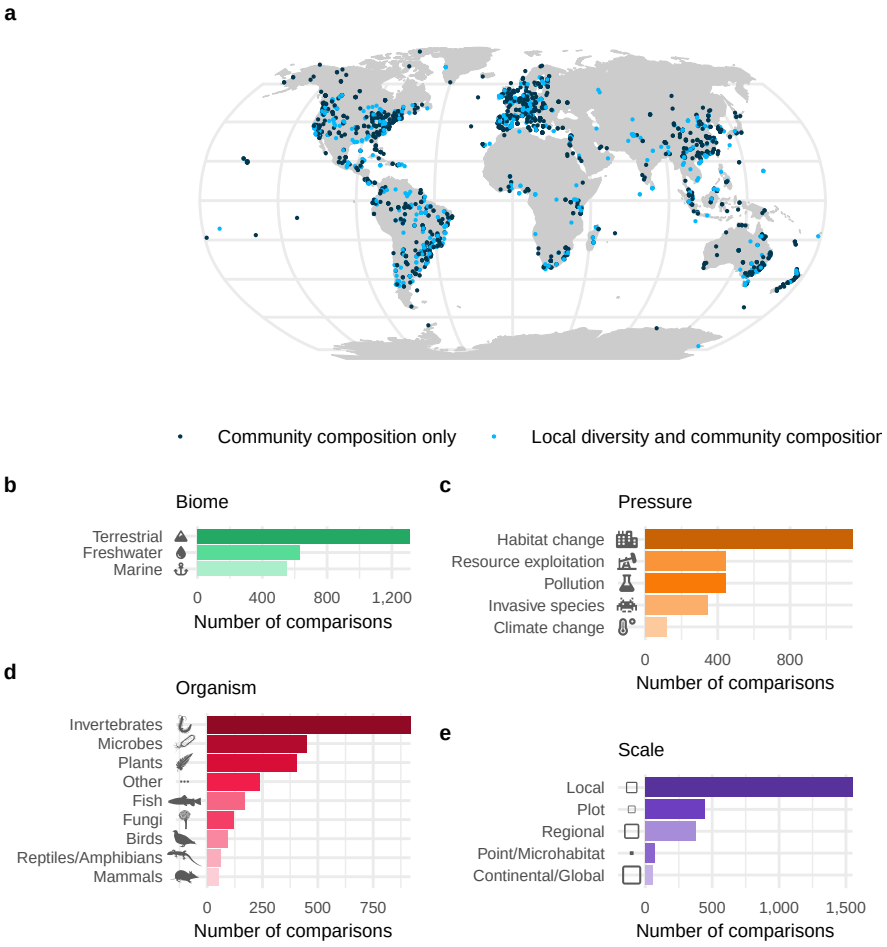

Figure S1 Location of diversity comparisons from observational studies and their distribution across biomes, pressures, organisms and scale. **a**, Global map of the 2,501 comparisons of diversity included. **b**–**e**, Distribution of comparisons of diversity by type of biome, human pressures, groups of organisms and spatial scale. These variables correspond to the four main factors tested in this study.

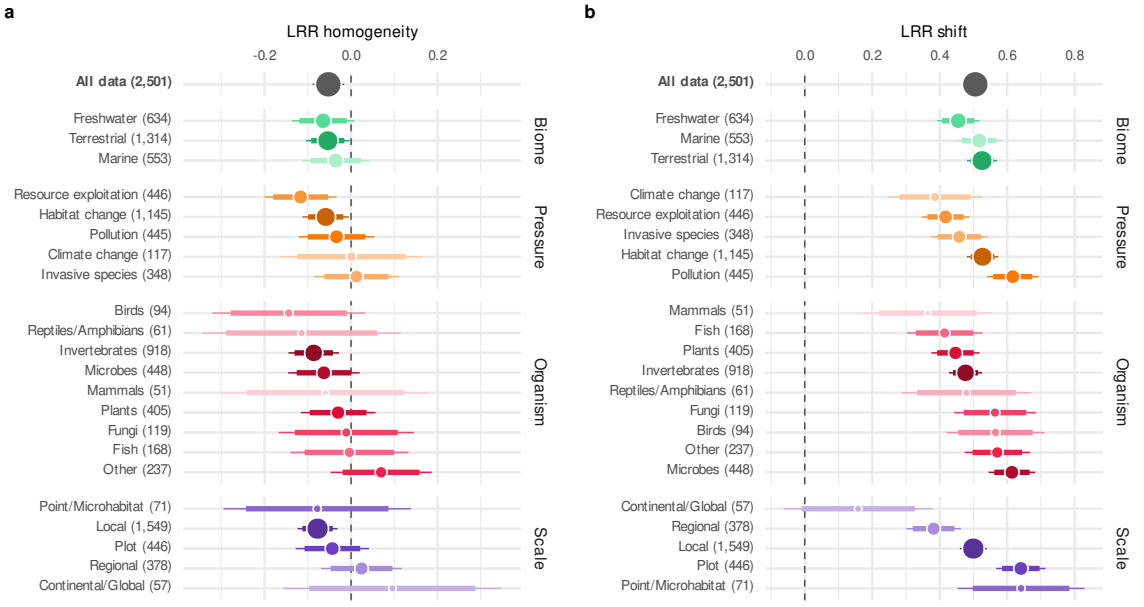

Figure S2 Impacts of human pressures on homogeneity and shifts in composition of biological communities estimated from observational studies. **a**, Log-response ratio of community homogeneity (logarithm-transformed ratio of impacted to reference values, LRR homogeneity). **b**, Log-response ratio of community composition shift (LRR shift). The global response (all data) is shown on the first row of each panel and is separated by factors in the following rows. The numbers between parentheses indicate the number of comparisons. For each category the dot represents the marginal mean computed from the model; dot size is proportional to number of studies included. The larger bar shows the 95 % confidence interval and the thinner bar represents the 99 % confidence interval.

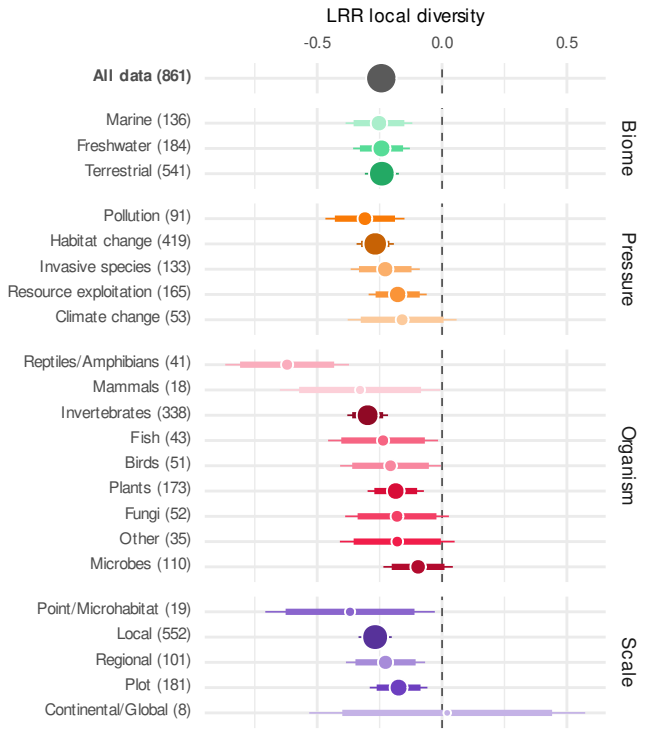

Figure S3 Impacts of human pressures on local diversity estimated from observational studies. Log-response ratio of local diversity (logarithm-transformed ratio of impacted to reference values, LRR local diversity). The global response (all data) shown on the first row is separated by factors in the following rows. The numbers between parentheses indicate the number of comparisons. For each category the dot represents the marginal mean computed from the model; dot size is proportional to number of studies included. The larger bar shows the 95 % confidence interval and the thinner bar represents the 99 % confidence interval.

S2. Stratified analysis: experimental studies

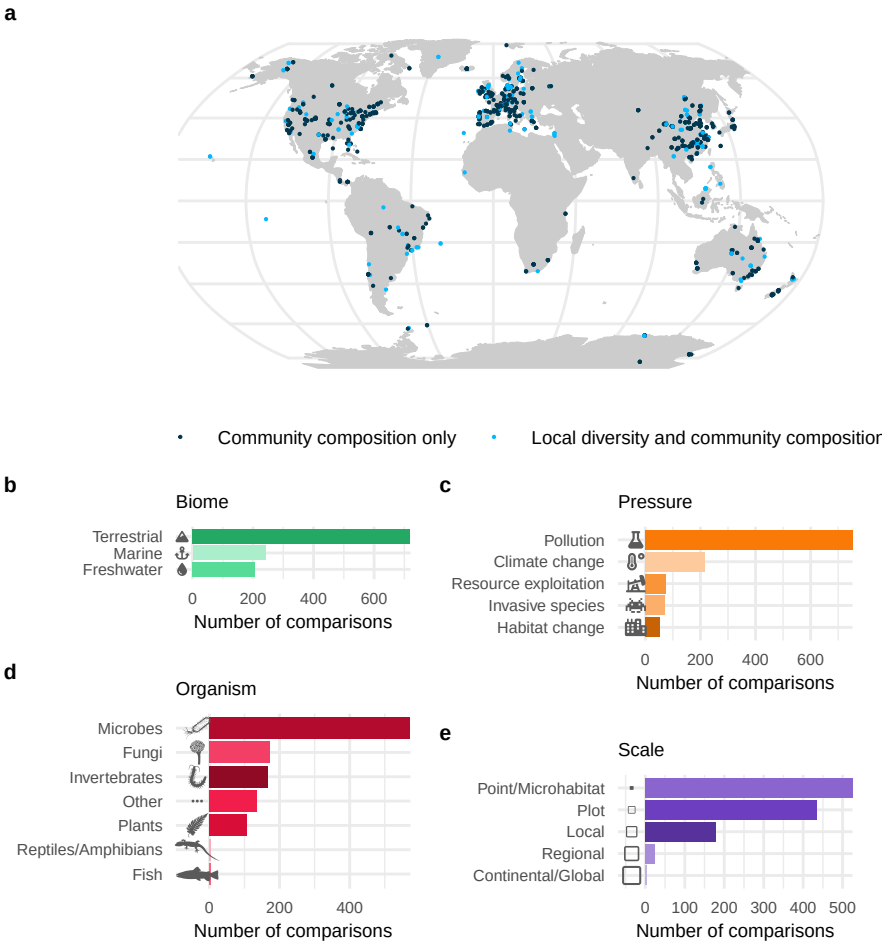

Figure S4 Location of diversity comparisons from experimental studies and their distribution across biomes, pressures, organisms and scale. **a**, Global map of the 1,166 comparisons of diversity included. **b**–**e**, Distribution of comparisons of diversity by type of biome, human pressures, groups of organisms and spatial scale. These variables correspond to the four main factors tested in this study.

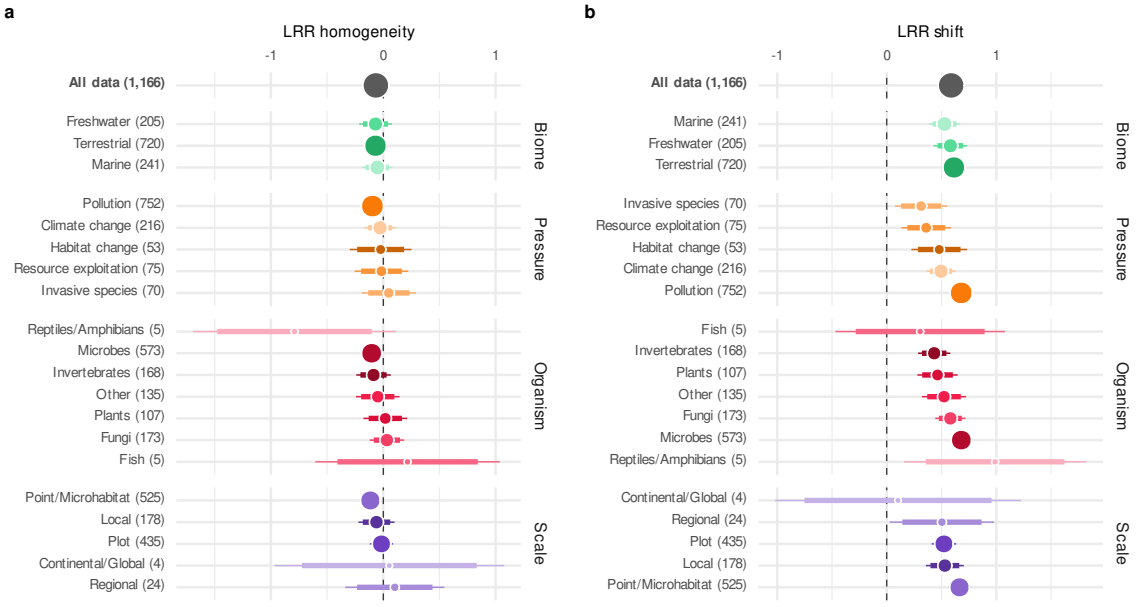

Figure S5 Impacts of human pressures on homogeneity and shifts in composition of biological communities estimated from experimental studies. **a**, Log-response ratio of community homogeneity (logarithm-transformed ratio of impacted to reference values, LRR homogeneity). **b**, Log-response ratio of community composition shift (LRR shift). The global response (all data) is shown on the first row of each panel and is separated by factors in the following rows. The numbers between parentheses indicate the number of comparisons. For each category the dot represents the marginal mean computed from the model; dot size is proportional to number of studies included. The larger bar shows the 95 % confidence interval and the thinner bar represents the 99 % confidence interval.

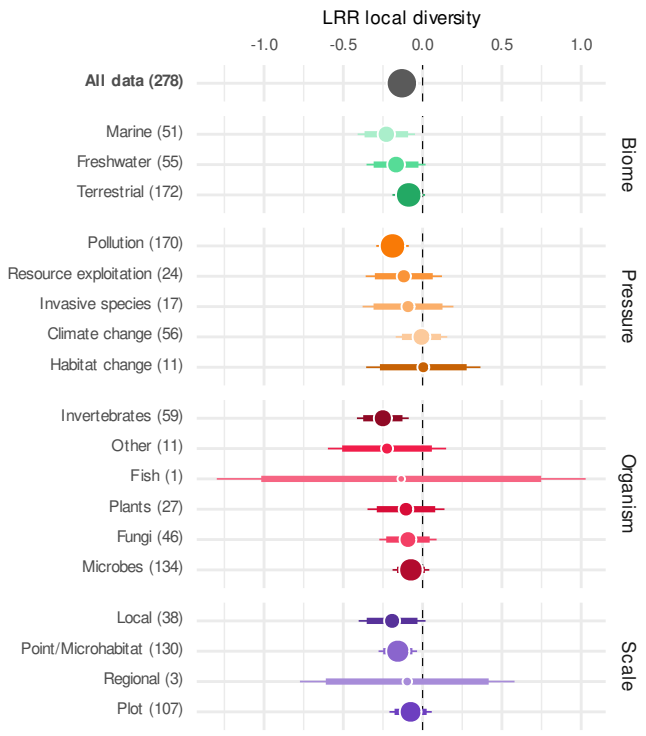

Figure S6 Impacts of human pressures on local diversity estimated from experimental studies. Log-response ratio of local diversity (logarithm-transformed ratio of impacted to reference values, LRR local diversity). The global response (all data) shown on the first row is separated by factors in the following rows. The numbers between parentheses indicate the number of comparisons. For each category the dot represents the marginal mean computed from the model; dot size is proportional to number of studies included. The larger bar shows the 95 % confidence interval and the thinner bar represents the 99 % confidence interval.

S3. Additional information on data collection and extraction

The metadata of bibliographic references in RIS format and their associated full text in PDF format were centralised and managed in a dedicated Zotero database. The text content of the PDF was extracted using the command-line utility tool pdftotext. The extracted text was analysed to detect any references to PCoA and NMDS analyses by searching for the following regular expression :

```
(?i)NMDS|(?i)PCOA|(?i)principal +coordinate|(?i)multidimensional +scaling
```

The documents containing the target terms were then manually evaluated in order to decide on their inclusion on the basis of a simple rule: presence of at least one biplot graph representing the first two axes of a PCoA or NMDS, and including, while differentiating them, at least two treatments ("control" and "impact"). This procedure was carried out by a single operator. A post-hoc check was carried out on a sub-sample of 100 publications by a second operator. This second operator identified 100% of the publications selected by the main operator (37 in total) while being slightly more conservative by selecting 4 additional references. These references would have been eliminated in any case at the data extraction stage due to the absence of some mandatory information (regarding the contrast of control and impact or other criteria used in the extraction step).

The actual data extraction phase was carried out by the 16 project participants. The extraction methodology was communicated to the participants during two workshops (one face-to-face and one online) and through a briefing document. The workshops included a number of practical demonstrations using a variety of examples to review the various extraction stages and special cases. Face-to-face or remote sessions were then organised, where participants worked together on data extraction and the most complex cases.

The extraction phase was carried out using a dedicated web platform developed specifically for the project. The aim of this platform was to maximise the quality of the user experience while minimising the risk of input errors and offering total control of user privileges to prevent data being exposed. Each participant was provided with a personal account and could add data using an interactive web form. This form supported various features to assist with the data extraction, including required fields, data validation (structures, types, pre-defined options, regular expressions), integration of help in the interface (placeholders and tooltips), and dynamic layout (conditional display of components). The website was built on the content management system Drupal (v.9) and was backed by a MySQL database that connected bibliographic references to data added by users in a seamless and secure way.

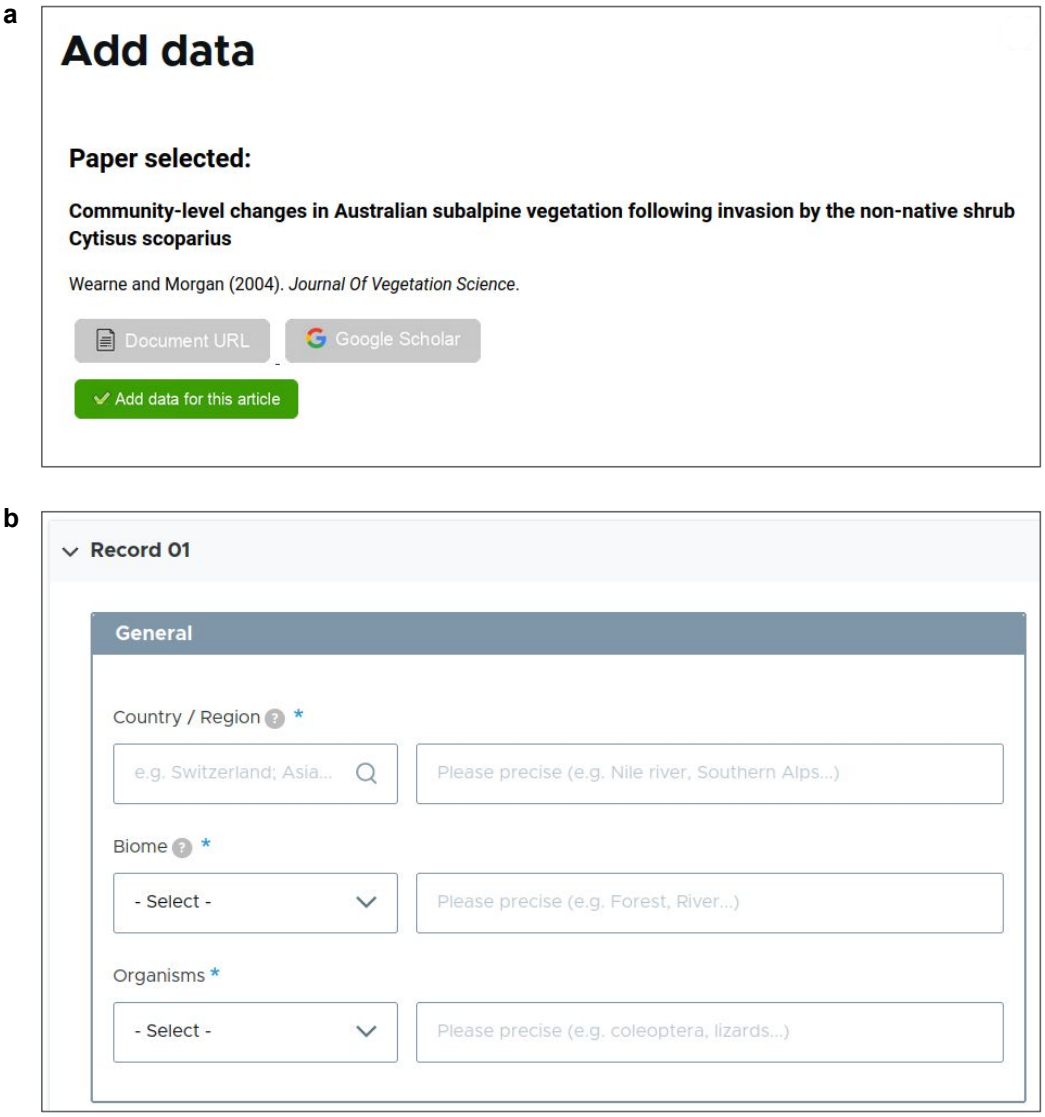

Figure S7 Screenshots of the interface of the web platform developed for data extraction. Connected users were **a**, invited to read and extract data from a publication which was randomly selected from the backend database. If they accepted, they were redirected to **b**, an interactive form designed to assist them with data extraction (only an extract of this form is shown). After they validated the form, data were automatically submitted to the database and users were invited to proceed with a new publication.

S4. Impact of additional samples in PCoA/NMDS on effect size estimation

Our meta-analysis is based on the extraction of sample coordinates from PCoA or NMDS two-dimensional graphs. It is common for several groups of samples or treatments to be represented on the same graph, which led us either to extract several comparisons per graph, or to ignore certain groups of samples (the rules applied for this choice are explained in the "Search strategy" part of the Methods section of the manuscript). The inclusion of additional samples or groups of samples can modify the configuration of the points as calculated by the projection methods, with the possible consequence that the observed distances no longer truly reflect the patterns studied. The impact of other sample groups in the analyses is *a priori* limited since our effect sizes (LRR homogeneity and LRR shift) depend on the relative distances and not the absolute distances between points. We show by simulation that this effect is indeed limited.

Our approach is as follows. We randomly generate a set of communities distributed in n groups ($3 \leq n \leq 5$), each group being defined by its own environmental conditions. We then compute a first PCoA (or NMDS) for the communities belonging to the first two groups, from which we calculate the LRR homogeneity and LRR shift effect sizes. We then compute a second PCoA (or NMDS), this time with all the groups, from which we calculate the LRR homogeneity and LRR shift effect sizes for the first two groups only. The comparison of LRR homogeneity and LRR shift between the two PCoAs allows us to estimate the bias induced by the presence of other groups in the analysis. Details of the community simulation are given in the following pseudo-code (the R code used is available on GitHub).

Listing S1 Algorithm for community simulation

```
Determine the number of groups, n, an integer randomly chosen in [3, 4, 5].
Determine the number of communities for each group, where each value is an
    integer randomly chosen in [20, ..., 200].
Determine the number of environmental dimensions, m, an integer randomly
    chosen in [3, 4, 5].
Determine the number of species in the pool, s, an integer randomly chosen in
    [15, ..., 100], and create S an indexed collection of species of size s.
Create a grouping vector, G, where each element G_i is an integer in [1, ..., n] and
     represents the group membership of each communities C_i.
Generate an s × m random matrix M with each element independently sampled from
    a uniform distribution in the range [−1, 1], representing the means of the
    normal distributions of each species along each environmental dimension.
```

Generate an $s \times m$ random matrix V with each element independently sampled from a uniform distribution in the range $[0, 1]$, representing the standard deviation of the normal distributions of each species along each environmental dimension.

Generate an $n \times m$ random matrix E containing environmental conditions for each group along each environmental dimension, with each element independently sampled from a normal distribution $\mathcal{N}(\mu = 0, \sigma^2 = 0.05^2)$.

For each community C_i belonging to the group G_i, generate species abundances by randomly drawing 300 individuals from S, where for each species S_j the probability to be sampled is weighted by the product of the densities of its normal distributions along each environmental dimension $\forall k \in \{1, \dots, m\}$, $\mathcal{N}(\mu_{j,k} = M_{j,k}, \sigma^2_{j,k} = V^2_{j,k})$ corresponding to the environmental conditions of the community $E_{G_i,k}$.

In total, we simulated 1000 independent sets of communities. Overall, we found a strong correlation between LRR homogeneity derived from the 2 groups analysis and LRR homogeneity derived from the >2 groups analysis (Figure S8a,c). This correlation was $\rho = 0.85$ (t = 51.53, df = 998, P < 0.001) for PCoA and $\rho = 0.70$ (t = 31.176, df = 998, P < 0.001) for NMDS. Figures S9a and S9c show the distribution of the difference between the two variables for the PCoA and NMDS analyses. The mean of this difference is 0.004 for PCoA and -0.007 for NMDS. One sample t-tests show that neither value is statistically different from zero (PCoA: t = 0.984, df = 999, P = 0.325; NMDS: t = -1.371, df = 999, P = 0.170), indicating the absence of bias.

We also found a very strong correlation between the LRR shift values derived from the 2 groups and the >2 groups analyses (Figure S8b,d). This correlation was $\rho = 0.99$ (t = 218.27, df = 998, P < 0.001) for PCoA and $\rho = 0.83$ (t = 31.176, df = 998, P < 0.001) for NMDS. Figures S9b and S9d show the distribution of the difference between the two variables for the PCoA and NMDS analyses. The mean of this difference is -0.035 for PCoA and 0.063 for NMDS. One sample t-tests show that both values are statistically different from zero (PCoA: t = -9.419, df = 999, P < 0.001; NMDS: t = 3.447, df = 999, P < 0.001), indicating a very small but consistent bias, possibly driven in the case of the NMDS by the series of outliers observed in Figure S8d and S9d.

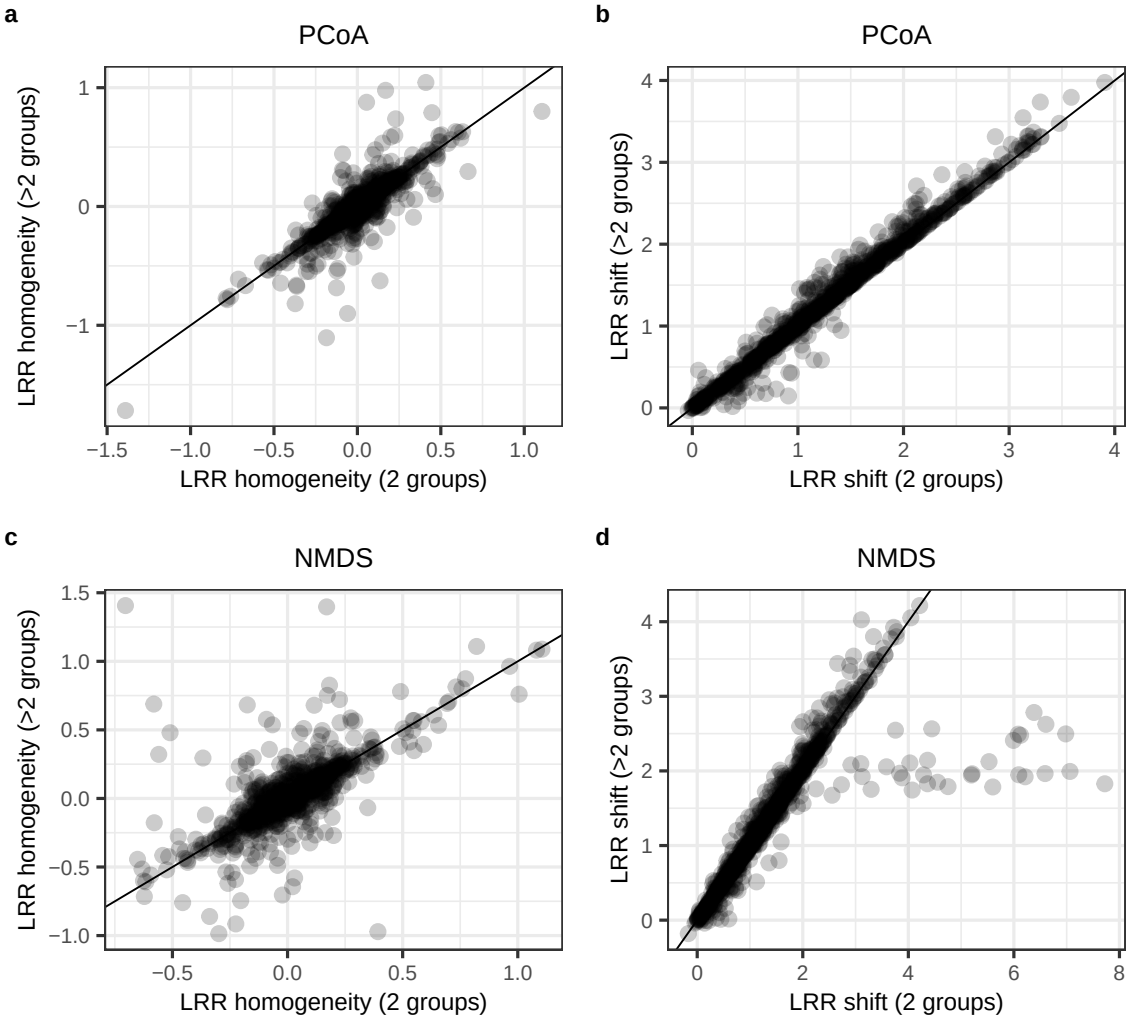

Figure S8 Relationship between the effect-sizes estimated from an analysis based on two groups of points and the same effect-sizes estimated from the same analysis but including additional groups of points. **a**. LRR homogeneity from PCoA, **b**. LRR shift from PCoA, **c**. LRR homogeneity from NMDS and **d**. LRR shift from NMDS. The black line delineates the 1:1 line, corresponding to perfect correspondence between the two estimates.

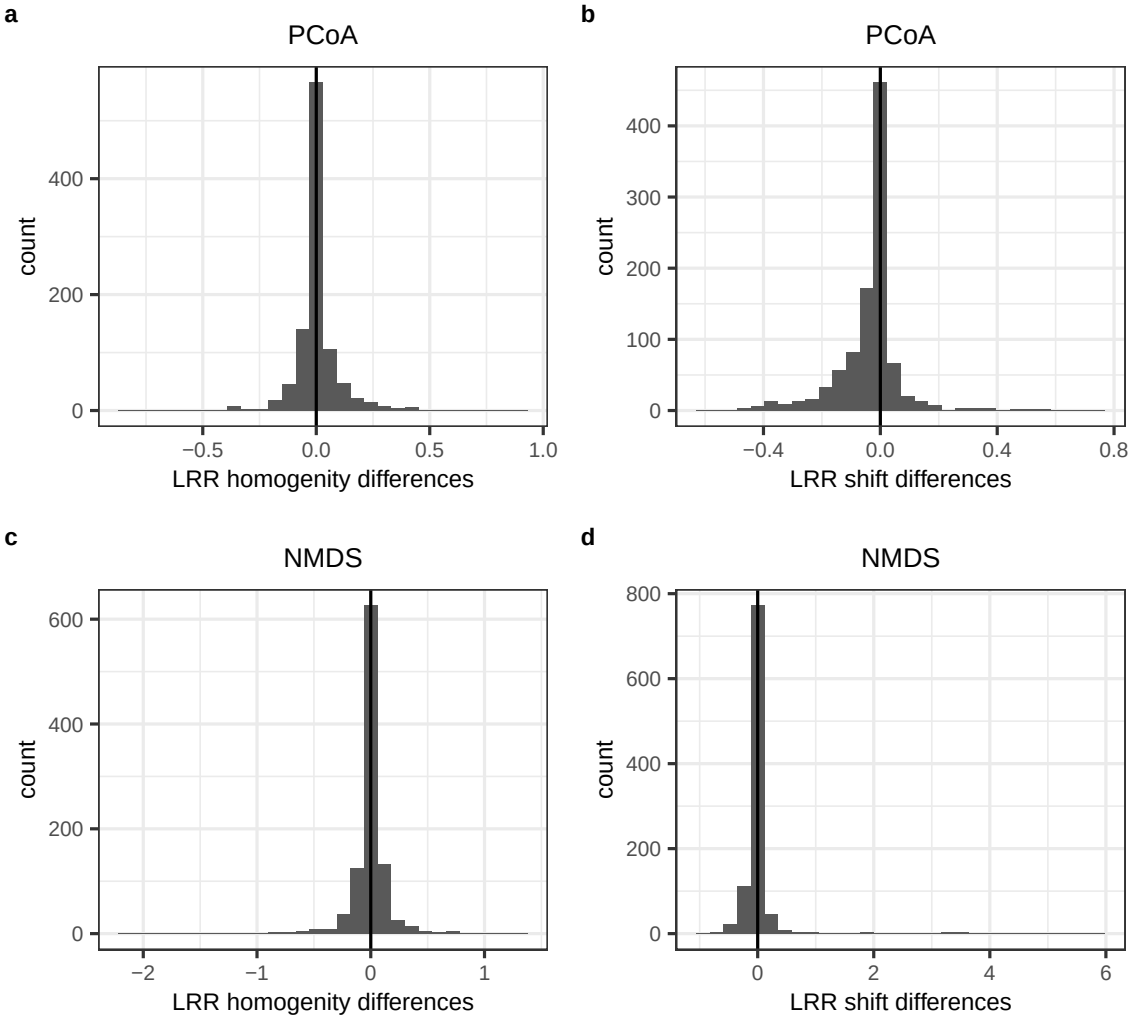

Figure S9 Distributions of the differences between the effect-sizes estimated from an analysis based on two groups of points and the same effect-sizes estimated from the same analysis but including additional groups of points. **a**. LRR homogeneity from PCoA, **b**. LRR shift from PCoA, **c**. LRR homogeneity from NMDS and **d**. LRR shift from NMDS.

S5. Replication of Figure 2 and 3 using weighted regressions

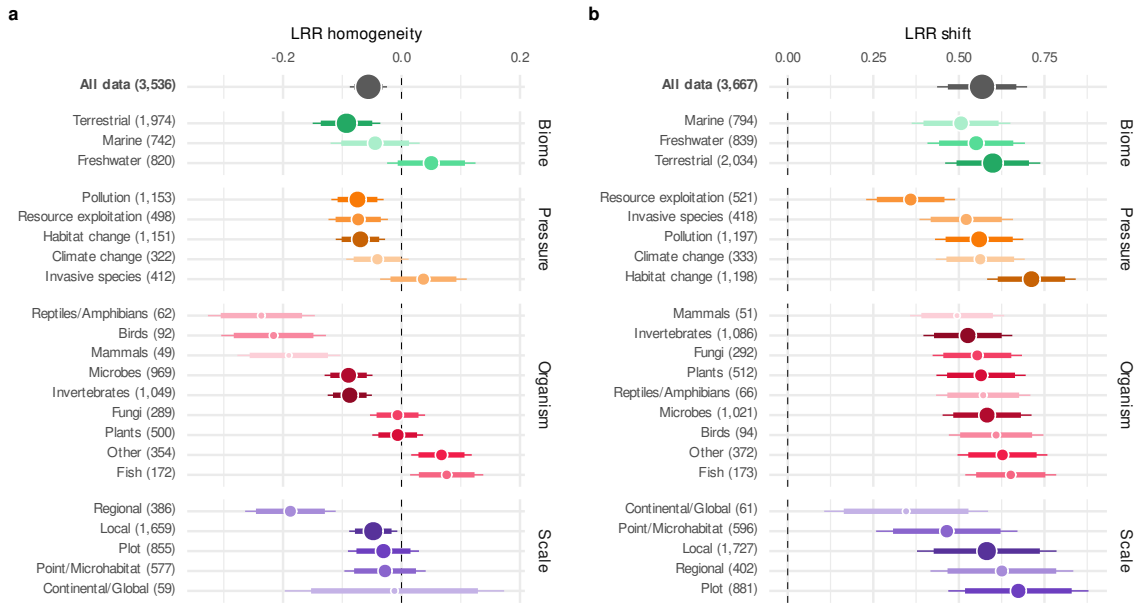

Figure S10 Impacts of human pressures on homogeneity and shifts in composition of biological communities estimated using weighted regressions. The weighting scheme applied is the inverse of the square root of the variance of the effect sizes, computed following Zhou et al.[1]. **a**, Log-response ratio of community homogeneity (logarithm-transformed ratio of impacted to reference values, LRR homogeneity). **b**, Log-response ratio of community composition shift (LRR shift). The global response (all data) is shown on the first row of each panel and is separated by factors in the following rows. The numbers between parentheses indicate the number of comparisons. For each category the dot represents the marginal mean computed from the model; dot size is proportional to number of studies included. The larger bar shows the 95 % confidence interval and the thinner bar represents the 99 % confidence interval.

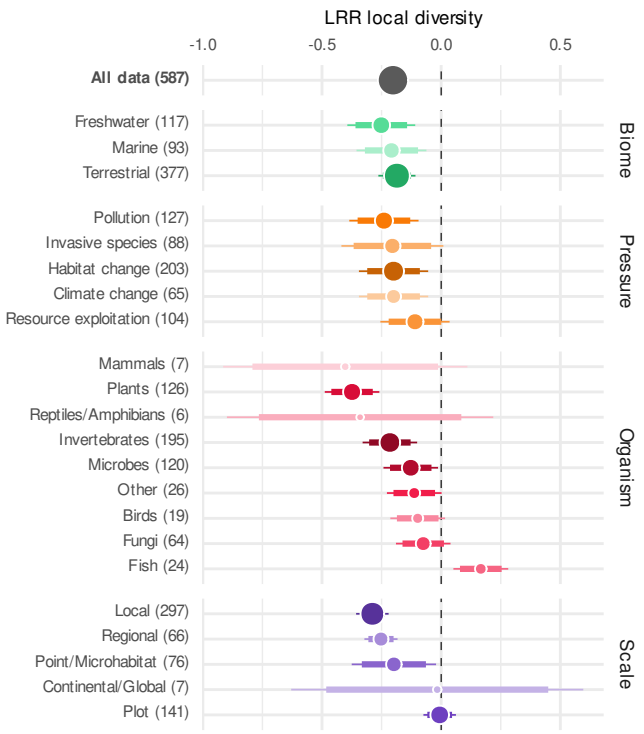

Figure S11 Impacts of human pressures on local diversity estimated from experimental studies using weighted regression. The weighting scheme applied is the inverse of the variance of the effect size, computed following Hedges et al.[2]. Log-response ratio of local diversity (logarithm-transformed ratio of impacted to reference values, LRR local diversity). The global response (all data) shown on the first row is separated by factors in the following rows. The numbers between parentheses indicate the number of comparisons. For each category the dot represents the marginal mean computed from the model; dot size is proportional to number of studies included. The larger bar shows the 95 % confidence interval and the thinner bar represents the 99 % confidence interval.

S6. Assessment of publication bias

We assessed potential publication bias for each outcome measure, namely LRR homogeneity, LRR shift and LRR local diversity. Our strategy was based on a combination of funnel plot visual inspection[3], file drawer analysis[4] and P-curve analysis[5].

We observe no evidence of asymmetry on the funnel plots with regard to LRR homogeneity (Figure S12a) and LRR local diversity (Figure S12c) while for LRR shift a small asymmetry can be observed (Figure S12b). This asymmetry, however, appears to be linked to the effect size distribution and has no apparent relationship with the standard error. A small-study effect, on the contrary, would have resulted in an over-representation of studies in the lower right-hand part of the funnel plot only.

To ascertain the robustness of our findings against possible publication bias, we performed a File Drawer Analysis and calculated the Fail-Safe N, a measure estimating the number of additional studies with null findings required to render observed effects non-significant. The File Drawer Analysis was conducted with the function `fsn` from the R package `metafor` using the generalised method of Orwin and Rosenberg[4,6]. We found Fail-Safe N values of 19,689 for LRR homogeneity, 2,757,894 for LRR shift and 12,052 for LRR local diversity. Overall these large to very large values (5 to 1000 times more studies than included in our analyses) indicate that our results are robust to publication bias.

Finally, we performed a test for p-values right-skewness (P-Curve analysis) to further validate the robustness of our results, specifically with regards to p-hacking. For every tested effect sizes, we found that both the half and full p-curve tests clearly indicated a right-skewed shape: LRR homogeneity (Full: $Z = -21.8$, $P < 0.001$, Half: $Z = -20.3$, $P < 0.001$), LRR shift (Full: $Z = -62.0$, $P < 0.001$, Half: $Z = -51.5$, $P < 0.001$) and LRR local diversity (Full: $Z = -66.4$, $P < 0.001$, Half: $Z = -67.2$, $P < 0.001$). Together these results indicate that there is strong evidential value behind our data.

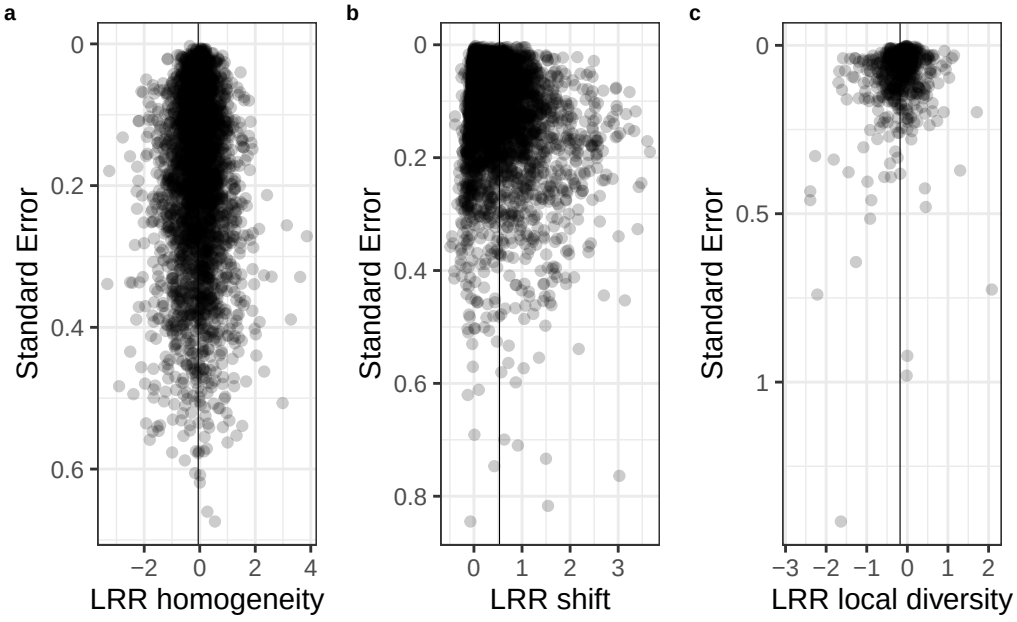

Figure S12 Funnel plots of effect sizes versus standard error for **a**. LRR homogeneity, **b**. LRR shift and **c**. LRR local diversity.

S7. Spatial distribution of reference and impacted sites

We analysed the spatial distribution of reference and impacted study sites in order to compare geographical patterns with those observed in biological communities.

To this end, we extracted the geographical coordinates of reference and impacted sites for a subset of 200 articles. The coordinates of impacted sites were extracted directly from the maps published in the articles, using the same digitising technique described in the manuscript for biological communities. We also used the same effect sizes (LRR homogeneity and LRR shift) here applied to geographic distance matrices to analyse whether 1) distances between samples within treatments were equivalent (LRR homogeneity) and 2) whether distances between treatments were greater than distances within treatments (LRR shift).

Spatial LRR homogeneity exhibited a distribution closely centred around zero (Figure S13a) and a small (0.17) yet significant mean value ($t = 2.979$, $df = 199$, $P = 0.003$). Similarly we found the mean of spatial LRR shift (0.32) significant ($t = 7.087$, $df = 199$, $P < 0.001$) and its distribution right skewed (Figure S13b). However, neither of the two space-based effect-sizes seems clearly correlated with their equivalent community-based measures (LRR homogeneity: $\rho = 0.08$, $t = 1.142$, $df = 198$, $P = 0.25$, Figure S13c; LRR shift: $\rho = 0.13$, $t = 1.888$, $df = 198$, $P = 0.06$, Figure S13d).

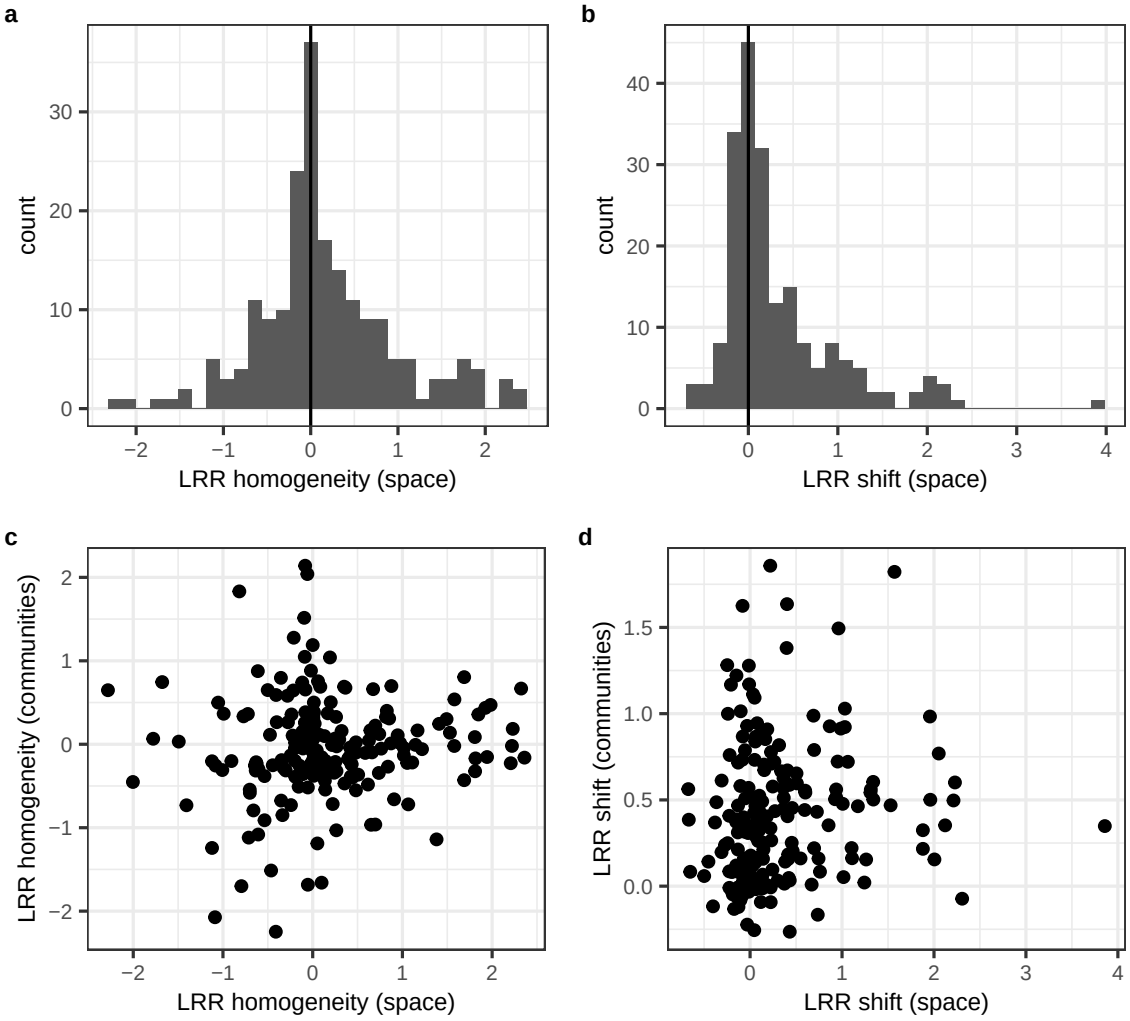

Figure S13 Distributions of **a**. LRR homogeneity and **b**. LRR shift values computed from spatial distances for a subset of 186 studies. Relationship between **c.** LRR homogeneity and **d.** LRR shift values computed from spatial and community distances.

Supplementary References

1. Zhou, Z., Wang, C. & Luo, Y. Meta-analysis of the impacts of global change factors on soil microbial diversity and functionality. *Nature Communications* **11**, 3072 (2020).

2. Hedges, L. V., Gurevitch, J. & Curtis, P. S. The meta-analysis of response ratios in experimental ecology. *Ecology* **80**, 11501156 (1999).

3. Møller, A. P. & Jennions, M. D. Testing and adjusting for publication bias. *Trends in Ecology & Evolution* **16**, 580–586 (2001).

4. Rosenberg, M. S. The file-drawer problem revisited: A general weighted method for calculating fail-safe numbers in meta-analysis. *Evolution* **59**, 464–468 (2005).

5. Simonsohn, U., Simmons, J. P. & Nelson, L. D. Better p-curves: Making p-curve analysis more robust to errors, fraud, and ambitious p-hacking, a reply to ulrich and miller (2015). *Journal of Experimental Psychology: General* **144**, 1146–1152 (2015).

6. Orwin, R. G. A Fail-SafeN for Effect Size in Meta-Analysis. *Journal of Educational Statistics* **8**, 157–159 (1983).