

Multidimensional β -diversity responses to global change: a meta-analysis highlighting divergent effects on plant communities

Corresponding Author: Professor Peng Luo

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

The manuscript presents a meta-analysis of the responses of plant taxonomic, phylogenetic, and functional beta-diversity to a variety of global change factors. The authors have amassed an impressive database of responses from sites around the world. Their analytical approach is generally sound (see below for minor related comments). However, the paper does not do an adequate job of setting up the questions and weaving a narrative throughout the discussion. These results are interesting and important to share with the scientific community, and pitching the framing correctly will go a long way towards doing so. We feel that the paper needs substantial revision to meet the standards of Nature Communications.

As written, the paper feels like a litany of statistical comparisons with a series of disjunct significant effects that are addressed in a random order in the discussion. A more understandable way to set up this paper would be a framework of (1) is there a beta-diversity response to global change factors (answer: yes, but there is a 50:50 split between homogenizing and differentiating) and then importantly (2) what factors drive the homogenizing vs differentiating (do you see homogenization more frequently in specific ecosystems, global change factors, etc). By framing the paper around explaining the fascinating trends of some responses being homogenization and others being differentiation would be much more compelling than the current approach that maintains the same analytical question (is there a response) throughout the manuscript. Then the discussion could be structured around what is driving the positive or negative response. Instead, it just feels like a deep dive into a bunch of random things because those are the things that wound up being significant.

Additional specific comments:

The abstract could tie concepts together more. It currently reads like an onslaught of information, without key motivations for those results or any underlying threads. Additionally, the statistical methods in the abstract are unclear. The authors use linear regression and random forest analysis to do what exactly or address what specific questions? The summary of their results here suggests that the variance in response to anthropogenic interventions skews results towards non-significance, when in fact the difference in beta-diversity effects is significant. Invasions always led to a decrease in beta-diversity, while other interventions varied in response direction. Climate change was the largest homogenizing force for functional beta-diversity, and land management was most significant for phylogenetic beta diversity. Finally, the use of the term "dimensions" to indicate taxonomic, functional, and phylogenetic beta-diversity is a little confusing because it is not well defined in the abstract. Clarification is needed.

The introductory paragraph relies heavily on one citation (Dornelas et al), which is a great paper. But a more varied set of references would be nice to support these points.

Word choice should also be carefully considered throughout the manuscript. Some specific choices include "dimensions" of beta-diversity (needs a better definition/explanation to use this term as many readers might not be familiar with dimensions of biodiversity in the NSF sense), "direction of variation" in beta-diversity (ln 54; variance implies spread in both directions around a mean, so how can it have a direction?), and human intervention (which sounds like a positive restoration-based term, when really it means a negative thing here). Similarly, "treatment" in the methods should be clarified to be a statistical treatment, rather than an experimental treatment.

The methods section needs to make the process of paring down a massive number of articles to a small subset (don't relegate this information to the supplement). The initial number of papers does not make sense with the numbers that were excluded to get to the final number included. Additionally, it is very unclear how the 47 papers were chosen for the ordination analysis, which needs to be added into the methods.

Ln 304-305: What does it mean that the correlations were predominantly from a subset of the total number of studies? How was this determined?

Figure 1: The color choice for the three dimensions of diversity is not red-green color-blind friendly. A darker shade for taxonomic or brighter shade for phylogenetic would be better. The density plots for panel a are difficult to read and don't seem necessary (don't convey any more information than the map itself). The points in panel b obscure each other. Perhaps do different shapes or sizes stacked from large (bottom) to small (top)? Panel d, the diverging x-axis for the number of studies and tests is confusing. Two bars with all positive values would be better.

Is a regression analysis the right way to look at correlations among the dimensions of beta-diversity? Regression implies directionality of driver-response and has specific requirements for assumptions. A correlation seems like a better choice.

Figure 4, panel a has a fishy regression line. How much influence are the outlier points having on this line?

The distinction between panels a vs b in Figure 5 needs clarification in the legend.

The discussion should include speculation about why some "dimensions" of beta-diversity are responding while others are not. These three dimensions are all tightly linked, so a good null hypothesis is that all should respond. Why don't they?

Minor typos/grammatical issues:

Ln 11 (verb tense), 42 (too many commas confuses sentence structure), 43-47 (overly complicated sentence, split it up?), 50-54 (overly complicated sentence), 59 (is it really numerous if there are only three citations?), 60-63 (what is the difference between a specific vs a global region?), 70 (what does no variations mean?), 75 (jargon heavy sentence), 87 (impacts to what?), 118 (missing word), 122 (unit needed), 161 (tiny sentence, combine with previous), 273 (clarify that it isn't the lnRR that is indicating the significant changes, but rather the confidence interval not overlapping with 0), 413 (missing space), 469 (missing letter)

Reviewer #2

(Remarks to the Author)

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

Reviewer #3

(Remarks to the Author)

Please see attached.

Reviewer #4

(Remarks to the Author)

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

The revised manuscript is substantially improved from the previous iteration and we thank the authors for their careful attention to reviewer comments. However, we feel the revision still remains unclear as to how the different analyses connect to each other and how the analyses relate back to the questions posed at the end of the introduction. Finally, the figures need clarification regarding what they are attempting to show. We outline our primary concerns below.

(1) The introduction brings the reader's attention to three specific questions. Throughout the results and discussion section, the many analyses presented do not seem to directly tie back into these questions. The authors should consider revisions to the methods, results, and discussion in order to highlights how their analyses answer those three questions.

(2) Figure 1: The rebuttal states that the density plots were removed from the borders of the map, this does not appear to be

the case.

(3) Lines 378-381: Refers to Figure 4, which are the bi-variate plots of response ratio correlations. This does not seem related to BI in the Himalayan region or Freshwater. Perhaps the authors meant to cite supplementary Figure 4? Or Figure 3?

(4) Figure 5: Which part of the plots are representing the relative contribution? The bar graphs, which appear to be the primary portion of the figures, do not match with the values listed in the donut graphs, and neither match with the values provided in the text. Specifically, the passage on lines 399-400 which reads: "... when considering the overall effects of GCFs, the dimension contributed 17.27% to the total impact..." This does not appear to match any of the plots provided in figure 5. The same goes for the passage on lines 404-405: "Plant life forms had a substantial impact (14.89%) on the effects of CC on plant β -diversity". The authors should clarify, either in the text or in their figure caption, what is being reported in these plots and how the donut graphs relate to the bar graphs.

(5) Figure 6: It is unclear what is being shown here, and the text connecting to this figure seems unrelated. If the effect is mainly being driven by LUC and LM, where are these factors, and how are they being represented in this model. The authors should consider replacing Figure 6 with their supplementary figure 6, as this helps to clarify their point.

(6) Additionally, authors should consider including an explanation of the statistical assumptions behind the meta-regression within their methods. Then later in the results, explain if their models adequately met those assumptions. Visually it seems to us that assumptions of error normality and homoscedasticity might not be met.

(7) Along the same lines, the authors should consider and add to the discussion the biological relevance of some of their findings, as opposed to only statistical relevance. For example, Figure 6 panels D, E, and G are all statistically significant but with very shallow slopes. How biologically meaningful is this relationship?

(7) Supplementary Figure 5 is mislabeled as Figure 3

(8) Minor syntax and grammatical issues

Line 6: "anthropogenic global change" to "anthropogenic global changes" to keep case consistency. Alternatively, authors could change "have" to "has".

Line 36: "is" to "are".

Line 37: insert "a" before global biodiversity crisis.

Line 47: insert "the" before "introduction"

Line 61: Consider changing "level" to "scales"

Lines 65-66: Consider changing "at the time" to "at a time"

Line 280: "will" to "was"

Line 525: Consider changing "lacks" to "drawbacks"

Reviewer #2

(Remarks to the Author)

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

Reviewer #3

(Remarks to the Author)

SUMMARY

We appreciate the efforts of the authors to improve the manuscript in response to our comments. However, some issues were marked as addressed, yet the corresponding changes were not fully reflected in the text. Additionally, it was difficult to verify where our comments had been incorporated, as the revised sections were not clearly indicated. Moving forward, we kindly ask the authors to specify the exact locations of their revisions by providing line numbers.

Furthermore, we strongly request that the authors complete the PRISMA-EcoEvo checklist (<https://www.prisma-statement.org/ecoevo>) to ensure transparency in their study. Please fill in the checklist with page and line number so we could check them carefully for comprehensive reporting.

Below, we provide our responses to the authors' comments

COMMENTS

L19: What does mean "tree- β diversity"? Please explain.

[Response 12] Thank you very much for your valuable comments. In our study, we classified plants into categories based on life forms, including aquatic plants, herbaceous plants, shrubs, trees, woody plants, combinations of woody and herbaceous plants, mosses or lichens, and so on. Subgroup analysis results showed that GCFs significantly reduced the β -diversity of trees (Figure 3). We acknowledge that the original wording may have caused confusion, and we have made corresponding revisions throughout the manuscript to clarify such issues.

Could the authors clarify what specific revisions were made throughout the manuscript to address this issue?

L64-65: It is unclear what "all three dimensions" refer to at this point (though it becomes apparent later that they are functional, phylogenetic, and taxonomic plant β -diversity). However, we noticed that no explanation or definition of these three different types of plant β -diversity was provided in the whole manuscript. We would suggest including this in the Introduction.

[Response 13] Thank you for your valuable comment. We appreciate the suggestion to clarify the reference to "all three dimensions." We have included a clearer definition in the abstract, where we specify that the three dimensions of biodiversity refer to taxonomic, functional, and phylogenetic diversity. We will also ensure that these terms are further defined and discussed in the Introduction to provide better clarity for readers throughout the manuscript. The authors state that these terms have been further defined in the Introduction to enhance clarity. However, we were unable to locate a clear explanation. If this revision has been made, please specify the relevant sentences.

We could not find the information on metadata, raw data, and scripts used for analyses from this manuscript. We strongly suggest they provide all the information necessary to replicate this study (at the review stage so we can assess them). Sharing them will enhance transparency, allowing others to validate or build on the findings. If there is a reason why they could not do it, the authors should clearly state that. - [18.1-4]

[Response 14] Thank you very much for your suggestion. We have uploaded the raw data into the submission system as Supplementary Material.

While the authors have provided the raw data and R scripts, the corresponding metadata explanation is still missing. Providing this information is important for ensuring that future researchers can properly interpret the data.

The authors should describe what sources of information were sought. For example, was the search limited to only published studies or included grey literature (e.g. personal communications, theses, or unpublished studies)? - [5.1]

L97-99: We would like them to provide the exact search strings used for each database. When we entered the search strings provided by the authors into the search engines of Web of Science, Google Scholar, and China Knowledge Resource, the total number of results was significantly lower than the number of papers reported by the authors in the manuscript. Also, we would like to know whether the authors set any temporal restrictions on the database searches. - [5.3-4]

L99: "35206 unique studies" is not a number that includes duplicate studies.

[Response 15] Thank you for your valuable feedback. We have updated the manuscript to provide more details about the sources of information sought. The literature search was conducted in Web of Science, Google Scholar, and the China Knowledge Resource Integrated Databases between 29 March and 31 August 2023. The search in Web of Science and Google Scholar used the following keyword combinations: ("beta diversity" OR " β diversity") and ("biotic homogen*" OR "biotic differentiat*" OR "biotic heterogen*" OR "biological homogen*" OR "biological differentiat*" OR "biological heterogen*"). In the China Knowledge Resource Integrated Databases, we translated these keywords into Chinese for the search. A total of 35,206 results were retrieved across all three search engines (see Supplementary Fig. 1), including duplicates. We did not apply any temporal restrictions during the search. Additionally, we examined the reference lists of relevant review articles, as well as subsequent studies citing these reviews (e.g., Catano et al. 2017, Kramer et al. 2022, McKinney 2006, Petsch et al. 2022, Rolls et al. 2023, Soininen et al. 2018) (229 studies). The search was limited to published studies.

The authors have not provided the exact search strings used in each database. By "exact search string," we refer to the precise query used, which should return the same or slightly more results as reported in the manuscript when entered directly into the database. We attempted to replicate the search using the provided search string but encountered the following issues:

Web of Science: the search resulted in an error message - "Search Error: invalid field tag."

Google Scholar: the search returned only 3 results, which seems inconsistent with the reported results.

Could the authors clarify the exact search string used for each database, including any field tags, Boolean operators, and filters applied? This would help ensure the reproducibility of the literature search.

L117: How many people were involved, and how did they contribute during screening? And more importantly, how they evaluate the reliability of their screening and extraction (e.g. double-screening, double-checking) - [6.2]

[Response 16] Thank you for your valuable feedback. To clarify, a total of four researchers were involved in the screening and data extraction process. Initially, two researchers (HLL and HY) independently screened all studies based on titles, keywords, and abstracts to identify studies focusing on the effects of human-induced global change on plant β -diversity, excluding 4,766 papers. Next, two researchers (HLL and MN) independently conducted a full-text review of the remaining 2,059 articles. In cases of uncertainties or disagreements, a third reviewer (PL) was consulted to resolve the issues through discussion. To ensure the reliability of the screening and extraction process, double screening and double checking were applied at each stage. The screening was conducted independently by the researchers, and disagreements were resolved by consensus, involving a third reviewer when necessary. This approach helped minimize potential bias and ensure the consistency and accuracy of the study selection process.

We appreciate that the authors have addressed our comments. This part has been adequately resolved.

L118-120: Why the total test number ($n = 294$) from adding functional, phylogenetic, and taxonomic plant β -diversity does not match the number reported in the manuscript ($n = 256$)? We might have overlooked something, but is it the case that some studies tested multiple β -diversity?

L132: How did they handle the missing data or ambiguous information? - [7.4]. Also, it should be better to state who collected data and the number of extractions checked for accuracy by co-authors. - [7.5-6]

[Response 17] Indeed, some studies include multiple dimensions of biodiversity, which could explain the discrepancy in the total test number. We have taken this into account in our analysis, but the discrepancy arises because these studies report separate results for each biodiversity dimension.

[Response 18] We independently extracted measures of central tendency (mean and median), dispersion (standard deviations, standard error, confidence intervals, and amplitude), and sample sizes for both the control (sites not exposed to GCFs) and treatment (sites affected by GCFs) groups in the selected studies by two groups (HLL and WWX combined as group one, and YC and YH combined as group two). In cases where studies did not report outcomes correctly or lacked sufficient original data, they were excluded from our analysis.

This part has been adequately resolved.

L147-153: We would like to provide more specific and detailed explanations for all the variables included in the six categories. For example, what is meant by 'plant functional type,' and what categories are included? Additionally, were any variables among they extracted either not used in the analysis or excluded due to insufficient information? - [8.1-4]

[Response 19] We sincerely thank the reviewers for their insightful suggestions. We have revised the manuscript to address both the request for more specific explanations of the variables and the suggestion regarding the quality assessment of primary studies.

1) In response to your request for more specific details about the variables included in the seven categories, we have provided clearer definitions in the manuscript. We also acknowledge that the term "plant functional type" was not used correctly and have replaced it with "plant life forms," which is a more appropriate description. The seven categories and their specific variables are as follows:

Five GCFs: biological invasion (BI), climate change (CC), land use change (LUC), land management (LM), and multiple factors (MF).

β -diversity dimension: Three dimensions of β -diversity are considered: taxonomic β -diversity (TB), functional β -diversity (FB), and phylogenetic β -diversity (PB).

Geographic location: The geographic variables include latitude, longitude, region, realm, elevation, and whether the study site falls within a biodiversity hotspot.

Ecological setting: This includes climatic zone, ecosystem type, flora realm, mean annual precipitation (MAP, in mm), and mean annual temperature (MAT, in °C).

Study design: This includes information on whether the study focused on temporal or spatial design, sample size, control sample type, temporal extent, and spatial extent.

Data characteristics: data sources and whether the data were abundance-weighted.

Research subject: This includes the examined plant taxa and life forms (instead of "plant functional types").

Each of the aforementioned data points was sourced either directly from the study reports themselves or by utilizing the latitude and longitude of the study sites. In cases where a study did not report the latitude and longitude, we estimated approximate coordinates by geocoding the site name using Google Earth 7.0 and then obtained the elevation based on these coordinates. Similarly, if the study did not provide mean annual temperature (MAT) and mean annual precipitation (MAP), we retrieved these values from WorldClim using the geographic location (latitude and longitude) of the study site. For studies that covered large areas (such as entire countries or the globe) without specific coordinates, we did not assign environmental metrics to them, which represented 6% of the total studies. It's worth noting that during the specific analysis, if a study did not involve a certain type of data (such as sample size), it would not be included in the effect analysis related to that data.

2) In response to the reviewer's comment, we have taken steps to reduce potential biases in our meta-analysis. Specifically, we have only included studies that have undergone peer review and are publicly published, ensuring a level of quality in the primary data. Additionally, based on the reviewer's suggestion and the Prisma-EcoEvo guidelines (O'Dea et al., 2021), we have excluded studies with a sample size of fewer than 3 from our analysis. Consequently, our meta-analysis is now based on 1604 observations from 256 studies, which helps to ensure the reliability and validity of our results.

Thank you for your response. However, we believe the following points have not yet been fully addressed:

How was the quality of individual studies assessed? (e.g., blinding of data collection, reporting quality, differences in study design)

Was the assessment of study quality incorporated into the meta-analysis? (e.g., meta-regression, sensitivity analysis)

A summary of study quality characteristics (e.g., risk of bias assessment)

The following paper may also be helpful as a reference:

Feeley TH. Assessing study quality in meta-analysis. *Human Communication Research*. 2020 Apr;46(2-3):334-42. <https://doi.org/10.1093/hcr/hqaa001>

Hohn RE, Slaney KL, Tafreshi D. Primary study quality in psychological meta-analyses: An empirical assessment of recent practice. *Frontiers in Psychology*. 2019 Jan 9;9:2667. <https://doi.org/10.3389/fpsyg.2018.02667>

Shaheen N, Shaheen A, Ramadan A, Hefnawy MT, Ramadan A, Ibrahim IA, Hassanein ME, Ashour ME, Flouty O. Appraising systematic reviews: a comprehensive guide to ensuring validity and reliability. *Frontiers in research metrics and analytics*. 2023 Dec 21;8:1268045. <https://doi.org/10.3389/frma.2023.1268045>

L187: equation (1) - parentheses for ratio will be needed.

L187-188: These are not the best estimators - see this paper

Lajeunesse MJ. Bias and correction for the log response ratio in ecological meta-analysis. *Ecology*. 2015 Aug;96(8):2056-63.

[Response 20] Thank you for your valuable feedback. In our analysis, we used the log response ratio (ln RR) as an effect size measure, and we are aware of the potential biases associated with this method, as highlighted in Lajeunesse (2015). To address these concerns, we estimated effect sizes using the Delta method recommended by Lajeunesse (2015), which

provides a more reliable and bias-correcting method of calculating In RR, particularly in ecological meta-analyses. In response to your suggestion regarding the equation, we have revised it to include parentheses for the ratio to improve clarity. The updated and additional equations are as follows...

L206: We do not think variance can be calculated using equation (3). Provide an equation for sampling variance.

[Response 21] Thank you for your comment. We have added the formula for calculating the variance. The variances (v) were calculated using the following equation...where n_t and n_c are the sample sizes of the variable in the treatment and control, respectively; s_t and s_c are the standard deviations of the variable in the treatment and control, respectively.

Thank you for addressing our comments. However, it would be helpful to provide a clearer justification for using this new method. The statement "To reduce bias in the effects..." (L196) does not seem to be a sufficient explanation.

The authors ticked the item about Bayesian analysis in the Nature portfolio reporting summary but did not provide information on the choice of priors or Markov Chain Monte Carlo setting - they should have given the information about this in the Results section. Please also provide all packages, functions, and their versions used to run the models. - [13.2-5]

[Response 22] Thanks for your attention. This study indeed does not include the Bayesian analysis. We have revised this choice in the Nature portfolio reporting summary.

The authors have not yet reported the specific functions and package versions used in their analyses. If possible, please include this information in the manuscript.

L215-232: We guess the authors used multilevel meta-analyses from their descriptions. It would be great to see two different levels of heterogeneity (e.g. I^2 at the effect size level and I^2 at the study level).

[Response 23] Thank you for your suggestion. We have included the I^2 statistic in the two-level subgroup analyses in the revised manuscript.

L218: Did the authors consider cohort ID (groups of the same subjects), shared control ID (indicating effect sizes sharing control groups), as well as study ID? - [14.1-3]

Note that the related points were recently explained in this paper with detailed tutorials - see Nakagawa S, Yang Y, Macartney EL, Spake R, Lagisz M. Quantitative evidence synthesis: a practical guide on meta-analysis, meta-regression, and publication bias tests for environmental sciences. *Environmental Evidence*. 2023 Apr 24;12(1):8.

[Response 24] Thank you for your insightful suggestion. In our previous manuscript, we considered the influence of the study ID but did not fully account for the non-independence of data arising from shared control groups within the same study. In response to your comment, we have revised our analysis accordingly. Specifically, we now incorporate both the study ID and control ID as random effects in our mixed-effects models. The control ID is nested within the study ID, meaning that each study has its own one or several sets of control groups, and the effect sizes within a study are potentially correlated due to this structure.

We have also referred to the detailed guidance provided in Nakagawa et al. (2023) to implement these adjustments in the context of environmental science meta-analysis. Below is the updated model:

Effect Size $_{i,j} = \beta_0 + (1 | \text{Study ID}) + (1 | \text{Study ID} : \text{Control ID}) + \epsilon_{i,j}$

Where Effect Size $_{i,j}$ represents the effect size for the i -th observation within the j -th study. β_0 is the fixed effect (the total effect of GCF on plant beta diversity). $(1 | \text{Study ID})$ accounts for variation between studies. $(1 | \text{Study ID} : \text{Control ID})$ accounts for variation due to control groups nested within each study. $\epsilon_{i,j}$ is the residual error term.

We appreciate the authors' revisions. One minor suggestion is to briefly explain the meaning of I^2 and Q_t .

L237: The authors used the fail-safe number, but it is heavily criticized. Consult a recent review of publication bias.

Nakagawa S, Lagisz M, Jennions MD, Koricheva J, Noble DW, Parker TH, Sánchez-Tójar A, Yang Y, O'Dea RE. Methods for testing publication bias in ecological and evolutionary meta-analyses. *Methods in Ecology and Evolution*. 2022 Jan;13(1):4-21.

[Response 25] Thank you for your suggestion. Considering the concerns regarding the reliability of the fail-safe number, we have removed the results based on this metric from the manuscript. Instead, we focused on more robust methods for assessing publication bias, such as the Egger regression test and the trim-and-fill method, as recommended by recent literature (Nakagawa et al., 2022). We believe these methods provide more reliable estimates of the potential effects of publication bias and offer a better evaluation of the robustness of our results.

L264-273: Could the authors clarify what is meant by "marginal publication bias"? Additionally, the relevant results should be reported in the Results section.

L240: We wonder why the authors used linear regression instead of meta-regression. We suspect errors could not be incorporated into Random Forest Models. Actually, unweighted models could be unbiased, but it depends on whether the data set has publication bias. Weighted models will be much better when the dataset contains publication bias.

[Response 26] Thank you for your suggestion. In response to your comment, we have revised our analysis to use meta-regression instead of linear regression. We appreciate your valuable feedback, which has led to an improvement in the statistical approach used in our study. The revised figures are shown in Figure 6 of the revised manuscript and Figure 6 of supplementary information.

Addressed.

Provide the number of effect sizes for the data included in meta-analyses (does the number of tests equal effect size?). - [20.1]

State a summary of key characteristics for reported outcomes. In particular, the authors did not provide any summary of the study design and research subject. - [20.3]

Report the results section based on statistical analysis; do not simply state that there was a significant positive or negative

effect without providing statistical details or reporting only the p-value. Always describe the full statistical results along with the findings. - [21.1, 23.1-4]

[Response 27] Thank you for your valuable feedback. We have carefully revised the manuscript to address your concerns. We have now included the number of effect sizes for the data included in the meta-analyses and clarified that the number of tests corresponds to the number of effect sizes used in the analyses. In the results section, we have ensured that all statistical analyses are presented in full, including effect sizes, confidence intervals, and relevant p values. We have avoided simply stating whether the effects were positive or negative and have provided a more comprehensive description of the statistical results. We hope these revisions address your concerns and improve the clarity of our manuscript. Thank you again for your helpful suggestions.

L311-326: Could the authors specify where in Figure 3A we should look? Furthermore, it would be helpful to provide the values before converting them to percentages, even in areas where percentages are shown.

The relationships between the figures in the Results/Discussion sections and the results are difficult to understand. Please revise to clarify which parts of the figure correspond to specific results.

L261-262: Why did the authors show the number of tests for homogenisation and differentiation in CC, LUC, and BI, but not in LM and MI?

[Response 28] Thank you for your suggestion. In the revised version, we have added figure numbers corresponding to the relevant text in the "Results and Discussion" sections to make it clearer which parts of the figures relate to specific results.

[Response 29] Thank you for your insightful comment. In the revised manuscript, we have now included the relevant data for LM and MF (MI has been revised to MF) as well. The updated sentence is as follows:

"In contrast, LM (with 400 tests for differentiation and 356 tests for homogenization) and MF (with 46 tests for differentiation and 37 tests for homogenization) exhibited the opposite trend."

The authors address these comments on their manuscript.

L274: Please explain the meta-regression analysis in the methods section (e.g. how many and what moderators are used? How select the final model?) - [15.1-3; 20.2,4]

[Response 30] Thank you for your insightful comment. We have made the necessary revisions to the manuscript to clarify the meta-regression analysis, as follows:

First, we conducted a meta-regression analysis to examine the relationship between the effect sizes (ln RR) and the following moderators: latitude, longitude, elevation, mean annual temperature (MAT), mean annual precipitation (MAP), spatial extent, temporal extent.

We performed the analysis using the metafor package in R (Viechtbauer, 2010). The analysis was weighted by the variance of each observation and utilized the Restricted Maximum Likelihood (REML) method to estimate the model parameters.

Regarding missing data, the temporal extent variable had over 50% missing values, so it was removed from the analysis.

For the remaining variables with less than 50% missing data, missing values were imputed using the Random Forest algorithm with default settings, implemented in the missForest package (Stekhoven et al., 2012).

We hope this clarifies the approach taken in the meta-regression analysis. Thank you again for your helpful feedback.

Were there any correlations among these moderators? Additionally, how was the validation of the imputed data evaluated?

L276: What is this percentage value? The back-transformed values of lnRR?

[Response 31] Thank you for your valuable comment. To clarify, the 12% reduction in the β -diversity of tree initially reported refers to the back-transformed values of the natural log response ratio (ln RR). This represents the effect size in the original scale, after converting the ln RR estimates into percentage values. Additionally, we would like to correct this figure: based on our reanalysis, GCFs actually reduced the β -diversity of trees by 15% ($p < 0.05$). We will revise the manuscript to reflect this updated result and provide a clearer explanation of the statistical approach.

The authors have not provided the explanation in their manuscript.

L301-306: Please explain the analyses regarding this paragraph's results in the Methods section.

[Response 32] Thank you for your constructive comment. In response, we have revised the Methods section to provide a more detailed explanation of the statistical analyses used for the results presented in this paragraph. The updated text is as follows: "Additionally, to investigate the consistency of the three dimensions of plant β -diversity (taxonomic, functional, and phylogenetic β -diversity) in response to GCFs, we conducted correlation analyses on the ln RR values of these dimensions using bivariate meta-analysis."

L344: What does binary meta-analysis mean in this context? The authors mention conducting bivariate (or binary?) meta-analyses, but after reviewing the R codes (file: "9.Figure 6-bivariate meta-analysis.R"), we could not determine how such meta-analyses were performed. Could the authors clarify this point?

Also...

L378-381/ L417-422: We may have overlooked something, but we are unsure which parts of the figures correspond to these findings. Could the authors clarify?

- Did they consider limitations and their influence on the generality of their conclusions?

- Although we do not know the number of data points for each variable, should not limitations such as gaps in the available evidence (e.g. geographical research biases or researched plant type bias) be considered? - [25.6]

- Please refer to the comments in the Result section as well.

[Response 33] Thank you for your insightful comments. We have added a discussion on the limitations of the study towards the end of our revised manuscript, as follows: "However, we also note that well-known biases, such as greater sampling effort in the global north compared to the Southern Hemisphere (especially in Australia), persist in the biodiversity data compiled for this study (Figure 1). This significantly limits the generalizability of our findings to well-studied regions of the planet."

The authors addressed our comments on the manuscript. However, we were wondering whether there are any citations here.

Please provide a reference list of all eligible studies for the meta-analysis. Also, please list included studies as referenced sources. - [27.1-2]

[Response 34] Thank you for your helpful suggestion. We have now included the reference list of all eligible studies for the meta-analysis in Supplementary Material 2, which contains the 256 studies that were included in this research.

We hope this addition meets your expectations. Thank you again for your valuable feedback.

Listing only the titles of referenced papers is insufficient. Please include the authors' names, DOIs, publication years, and journal names. Additionally, could the authors provide a full list of articles excluded at the full-text screening stage, along with the reasons for exclusion? This would allow third parties to refer to the information more easily.

Figure 1: Figure 1 was quite challenging to understand. First, we would like clearer separations between sections A and C. Second, in A, please indicate what the insets' X-axis (top) and Y-axis (right) represent. The circles on the world map overlap too much and are hard to interpret. For example, what does this overlap mean? Third, in B, what does each plot represent? (we assumed it was the number of tests, but if so, should not taxonomic β -diversity have the most?) Finally, in C, this might be because it is outside our area of expertise, but in what situation is elevation over 3000 m in freshwater ecosystems?

[Response 35] Thank you for your valuable feedback on Figure 1. We have made several revisions to improve clarity and address your concerns:

1) Clearer Separations Between Sections A and C: To enhance the distinction between sections A and C, we have adjusted the color configuration. Section A now displays the global distribution of observations that include the different dimensions of diversity included in this study. Section C shows the distribution of these diversity dimensions across realms and elevations. This should make the separation between the two sections clearer.

Addressed.

2) Clarification of Axes in Section A: In section A, we have now labeled the axes of the insets for clarity. The X-axis (top) represents longitude, and the Y-axis (right) represents latitude. These labels should provide additional context for interpreting the insets more easily.

Addressed.

3) Overlap of Circles in the World Map: The circles on the world map in section A represent studies that include multiple dimensions of biodiversity (taxonomic, functional, and phylogenetic). The overlap of circles indicates that a single study addresses more than one dimension of biodiversity. To clarify this, we have adjusted the circle sizes to be uniform, removing the reflection of sampling size, which should now make the map easier to interpret.

Addressed.

4) Interpretation of Plots in Section B: Section B displays the distribution of diversity dimensions with respect to mean annual temperature (MAT) and mean annual precipitation (MAP). We have adjusted the sizes of the circles for each dimension to prevent overlap and enhance visibility. Additionally, we reviewed the representation of taxonomic β -diversity to ensure accuracy, as it is expected to have the largest number of studies. This has been corrected in the figure.

It would be helpful to include an explanation of the circle sizes in the figure legend.

5) Elevation Over 3000 m in Freshwater Ecosystems (Section C): We acknowledge that the example in section C, which includes an elevation over 3000 m, may seem unusual for freshwater ecosystems. The study referenced (Rosero et al., 2021) was conducted in a small glacierized catchment in Ecuador, located at 5018 m a.s.l. at the Carihuairazo volcano (01°24'25"S, 78°45'00"W). This site is an exceptional case relevant to our study, and we have updated the figure caption to provide this additional context for clarity.

Understood.

Figure 3: What do the numeric values in parentheses within the figure represent? Also, please explain both Fig. 3A and Fig. 3B in the legend separately.

[Response 36] Thank you for your comment. The numeric values in parentheses within the figure represent the number of studies and the number of observations, respectively. We have also provided separate explanations for both Fig. 3A and Fig. 3B in the legend. The revised descriptions are as follows. Additionally, we have made corresponding modifications to other figures of the same type for consistency.

"Figure 3 The effect size (lnRR) of GCFs on plant β -diversity varies across different GCFs, dimensions, hotspot status, hotspots, regions, elevations, realms, and climate zones (A), as well as across ecosystems, floristic realms, plant growth forms, taxa types, abundance status, data sources, and study design aspects such as temporal extent and spatial extent (B). Data are presented as mean $\pm$ 95% CIs of the estimated effect sizes. Nonsignificant changes are represented by gray dots, while significantly negative and positive effects are represented by red and green dots, respectively. The numbers in parentheses indicate the number of studies and the number of observations, respectively."

The authors have addressed our comments in the manuscript.

Figure 4: We could not find the method related to this. One cannot just use the correlation to analyse the relationships between 2 sets of effect sizes (e.g. pseudo-replication and also there are two correlations that can be calculated at the between- and within-study levels). One should use bivariate meta-analysis. See this link (but your model will be more complex): <https://www.metafor-project.org/doku.php/analyses:vanhouwelingen2002>

[Response 37] Thank you for your valuable feedback. You are correct that simply using correlation analysis may not be appropriate for analyzing the relationships between two sets of effect sizes, as it could lead to pseudo-replication issues. In response to your comment, we have updated the methods section of the manuscript to include the following clarification:

"Additionally, to investigate the consistency of the three dimensions of plant β -diversity (taxonomic, functional, and phylogenetic β -diversity) in response to GCFs, we conducted Pearson correlation analyses on the ln RR values of these dimensions. This analysis was performed only for studies that included data from at least two dimensions of β -diversity."

As you suggested, we acknowledge that bivariate meta-analysis would be more appropriate for this type of analysis.

However, due to the limitations of the available data, we opted to conduct the correlation analysis and present the results graphically. We used this approach to display the consistency of the direction of change in β -diversity dimensions due to

GCFs, based on data where at least two dimensions were reported.

We have updated the figure legend and manuscript accordingly to reflect these changes.

The legend of Figure 4 and the main text state that a bivariate meta-analysis was conducted. However, the authors mention that Pearson correlation analysis was used instead due to data limitations. Could the authors clarify whether a bivariate meta-analysis was performed? If so, please provide the corresponding results.

Figure 5: Why are the marks on the y-axis of Fig. 5b shown in both black and grey, and what is the meaning of this difference? Also, what do the numbers in parentheses mean (e.g. TB×MI (16, 69))?

[Response 38] Thank you for your comment. In order to improve the clarity of the graphs and better illustrate the results, we have rearranged the graphs for this part of the results. Please see Figure 5 in the revised supplementary information.

Supplementary Figure 1: In this PRISMA-like flowchart, the authors should report brief reasons for exclusion from the full-text screening stage. Also, the number of Reports screened is 6825, not 7825. [19.1-4]

[Response 39] Thank you for pointing out these issues. In the revised version, we have added a clear summary of the reasons for excluding studies during the full-text screening stage. Additionally, we have corrected the number of reports screened. The correct number is indeed 6825, not 7825, and we have updated this figure accordingly in the revised manuscript and flowchart.

Supplementary Figure 2: We wondered how the authors compared the responses of plant β -diversity to global change drivers using data from studies that used different analytical methods. Supplementary Figure 2b: How many studies are included in each of these four analytical methods?

[Response 40] Thank you for your valuable comment. We appreciate your concern regarding the comparison of plant β -diversity responses to GCFs using data from studies that employed different analytical methods. To address this, we have conducted a meta-analysis to assess how different analytical methods influence the results.

We have included the number of studies corresponding to each analytical method in the figure, as requested. This allows for a clearer understanding of how the different analytical methods are distributed across studies and their effect on plant β -diversity response to GCFs.

Supplementary Figures 3 and 4: Please see the comment on Figure 3.

[Response 41] Thank you for your comment. As per our response to the reviewer's feedback on Figure 3, we have made the corresponding revisions to Supplementary Figures 3 and 4. The updated descriptions are as follows:

Supplementary Figure 3: The effect size (lnRR) of GCFs on the three dimensions of plant β -diversity—taxonomic (TB), functional (FB), and phylogenetic (PB)—varies across different GCFs, hotspot status, hotspots, regions, elevations, realms, and climate zones (left panel), as well as across ecosystems, floristic realms, plant life forms, taxa types, abundance status, data sources, and study design aspects such as temporal extent and spatial extent (right panel).

Supplementary Figure 4: The effect size (lnRR) of GCFs on the five GCFs—biological invasion (BI), climate change (CC), land use change (LUC); LM, land management (LM) and multiple factors (MF)—varies across different GCFs, hotspot status, hotspots, regions, elevations, realms, and climate zones (left panel), as well as across ecosystems, floristic realms, plant life forms, taxa types, abundance status, data sources, and study design aspects such as temporal extent and spatial extent (right panel).

Supplementary Figure 5: Please see the comment on Figure 5.

[Response 42] Thank you for your comment. In order to improve the clarity of the graphs and better illustrate the results, we have rearranged the graphs for this part of the results. Please see Supplementary Figure 5 in the revised version of the manuscript for the updated layout.

Supplementary Figure 8: There is no explanation regarding to this Fig.

[Response 43] Thank you for your comment. We have revised the presentation of these types of figures and added the necessary textual explanation in the updated figure legend. Please refer to Figure 6 for the revised supplementary information, which now includes a clearer explanation of the figure. The updated text is as following:

Supplementary Figure 6: Meta-regression results (red line) showing the relationships between mean annual temperature, mean annual precipitation, elevation, latitude, longitude, sample size, spatial extent, temporal extent, and the response ratios (ln RR) of plant β -diversity in response to specific types of global change factors (GCFs).

The authors addressed our comments above in the manuscript.

Reviewer #4

(Remarks to the Author)

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

Version 2:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

We appreciate the author's work in revising this manuscript. While the majority of the changes are acceptable, the figures are all very difficult to read. Figures should all be revised to increase text size for readability when printed on a single page.

Reviewer #3

(Remarks to the Author)

SUMMARY

Overall, the authors have responded well to the comments, and most of the points have been appropriately addressed. We appreciate the careful revisions and clarifications made throughout the manuscript. However, a few minor issues still remain, which we outline below. In addition, the figures in the supplementary materials are currently very small and difficult to read. We strongly encourage the authors to enlarge them to improve visibility and accessibility for readers.

[Response 12 – Second Round] Thank you very much for your careful review and comments. We have carefully examined the issues raised and believe that some misunderstandings may have occurred regarding how we presented our search strategy.

For example, when we searched for β diversity and biological homogen* in Google Scholar without quotation marks, the search returned approximately 18,000 results (as shown in Screenshot A, which we have now provided). To further clarify, we have also included Screenshot B, which displays a portion of the records retrieved from Web of Science during our searches conducted in 2023.

The confusion likely stems from the fact that in our original manuscript, we enclosed the keyword combinations within quotation marks (e.g., “beta diversity” AND “biological homogen*”). We now realize that including these quotation marks may have misled readers to enter them directly into the search boxes, which can result in very limited or even erroneous search results (e.g., only 2–3 records in Google Scholar). In response to this, we have removed quotation marks from the search strings in the revised manuscript to avoid any further confusion.

To reiterate, in the China Knowledge Resource Integrated Databases, we used translated Chinese terms that correspond to the English keywords used in Web of Science and Google Scholar. The searches were all conducted manually, with different keyword combinations tested and recorded systematically.

We thank the reviewers once again for their attention to detail and their commitment to ensuring the reproducibility of our methodology.

Thank you for your detailed response and for clarifying the issues regarding the search strategy. We appreciate the added screenshots and the explanation about the quotation marks possibly causing confusion. However, we would still strongly recommend that the exact search strings used for each database be provided in full, including any field tags, Boolean operators, and filters applied. This is essential to ensure full reproducibility of your search.

For example, for Web of Science, it would be helpful to specify whether you used tags such as TS=, TI=, or others, and how the keywords were combined. Similarly, for Google Scholar and the China Knowledge Resource Integrated Database, we encourage you to include the full queries (as entered) that yielded the reported results.

(13) Thank you for your response. However, we believe the following points have not yet been fully addressed:

How was the quality of individual studies assessed? (e.g., blinding of data collection, reporting quality, differences in study design)

Was the assessment of study quality incorporated into the meta-analysis? (e.g., meta-regression, sensitivity analysis)

A summary of study quality characteristics (e.g., risk of bias assessment)

The following paper may also be helpful as a reference:

Feeley TH. Assessing study quality in meta-analysis. *Human Communication Research*. 2020 Apr;46(2-3):334-42. <https://doi.org/10.1093/hcr/hqaa001>

Hohn RE, Slaney KL, Tafreshi D. Primary study quality in psychological meta-analyses: An empirical assessment of recent practice. *Frontiers in Psychology*. 2019 Jan 9;9:2667. <https://doi.org/10.3389/fpsyg.2018.02667>

Shaheen N, Shaheen A, Ramadan A, Hefnawy MT, Ramadan A, Ibrahim IA, Hassanein ME, Ashour ME, Flouty O. Appraising systematic reviews: a comprehensive guide to ensuring validity and reliability. *Frontiers in research metrics and analytics*. 2023 Dec 21;8:1268045. <https://doi.org/10.3389/frma.2023.1268045>

[Response 13 – Second Round] Thank you very much for your thoughtful comments. We acknowledge that our initial response did not fully address the important points you raised regarding the appraisal of individual study quality.

To address this, we have added a new section (2.3 Critical Appraisal) in the Methods part of the revised manuscript, where we describe our approach to assessing the quality of the included studies. Specifically, we evaluated study quality based on a set of criteria adapted from Mupepele et al. (2016), which are designed for assessing evidence strength in studies on ecosystem services and biodiversity. In determining the final criteria used in this study, we also referred to the methods applied by Huynh et al. (2024) and Su et al. (2021).

Each study was scored based on whether it met individual quality criteria, and the overall scores were used to classify studies into four levels of evidence: weak, moderate, strong, and very strong. A detailed breakdown of quality scores for each study is now provided in a newly added Supplementary file (Study quality assessment.xlsx), and a summary of the assessment criteria and scoring system is included in Supplementary Table 1.

We hope that these revisions improve the transparency of our methodology and help to strengthen the validity and reliability of our findings. Please let us know if any further clarification is needed—we would be happy to revise further in accordance with your suggestions.

Thank you for adding a detailed quality assessment framework and providing the checklist in Supplementary Table 1. We appreciate your effort to enhance transparency in study appraisal. However, we would kindly like to ask for further clarification on two points:

Was study quality (e.g., quality score or evidence level) used as a moderator in any sensitivity analyses or meta-regressions? If so, we would appreciate a pointer to the relevant results (main text or Supplementary).

While the checklist is detailed, would it be possible to include a brief summary in the main text (e.g., number or proportion of studies classified as weak, moderate, strong, very strong) to improve understanding for readers who may not check the supplementary files?

(18) L311-326: Could the authors specify where in Figure 3A we should look? Furthermore, it would be helpful to provide the values before converting them to percentages, even in areas where percentages are shown.

[Response 18 – Second Round] We sincerely thank the reviewer for this helpful suggestion.

In response, we have added the original $\ln RRR$ values along with their corresponding 95% confidence intervals throughout the relevant parts of the text to enhance transparency.

Additionally, we have clarified the specific subpanels and categories in Figure 3A where each result is displayed, enabling readers to easily locate the relevant data. These revisions are reflected in Lines 386–394 of the revised manuscript. To further improve clarity, we have also added the formula used for converting $\ln RRR$ values to percentage change in the Methods section (see Lines 259–260).

Thank you for providing the detailed explanation regarding collinearity and the rationale for retaining all moderators in the random forest analysis. However, we noticed that the variable “growth form”, which was mentioned in the author's response as being highly correlated with other variables, is not actually included in Figure 6.

Could the authors clarify whether growth form was excluded from the random forest models? If so, we would suggest briefly explaining why it was excluded, especially given its reported correlation with taxa type and realm. Additionally, even if retained variables are correlated, a short discussion of how multicollinearity might affect variable importance rankings (e.g., through shared explanatory power) would be helpful for readers interpreting the results.

[Response 35] Thank you for your valuable feedback on Figure 1. We have made several revisions to improve clarity and address your concerns:

4) Interpretation of Plots in Section B: Section B displays the distribution of diversity dimensions with respect to mean annual temperature (MAT) and mean annual precipitation (MAP). We have adjusted the sizes of the circles for each dimension to prevent overlap and enhance visibility. Additionally, we reviewed the representation of taxonomic β -diversity to ensure accuracy, as it is expected to have the largest number of studies. This has been corrected in the figure.

(24) It would be helpful to include an explanation of the circle sizes in the figure legend.

[Response 24 – Second Round] Thank you very much for your valuable comments. As the three dimensions of β -diversity are already differentiated by color in Figure 1A, B, and C, we used different circle sizes only in Figure 1B to reduce visual overlap and improve readability. In response to your suggestion, we have added an explanation of the circle sizes to the figure legend of Figure 1B to enhance clarity. Specifically, circle sizes were manually adjusted to represent the three dimensions of β -diversity: Taxonomic (small), Functional (medium), and Phylogenetic (large). Please refer to the updated figure legend on Lines 149–151 of the revised manuscript.

Thank you for clarifying the rationale for using different circle sizes in Figure 1B. However, we would like to respectfully suggest reconsidering this design choice. In general, varying symbol size is interpreted as representing quantitative information (e.g., magnitude, frequency), and using size purely to distinguish categorical variables can be visually misleading. Since the three beta-diversity dimensions are already clearly differentiated by colour, using size as an additional encoding may add unnecessary visual complexity or confusion.

Reviewer #4

(Remarks to the Author)

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

Open Access This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

Reviewer #1 & #2 (Remarks to the Author)

The manuscript presents a meta-analysis of the responses of plant taxonomic, phylogenetic, and functional beta-diversity to a variety of global change factors. The authors have amassed an impressive database of responses from sites around the world. Their analytical approach is generally sound (see below for minor related comments). However, the paper does not do an adequate job of setting up the questions and weaving a narrative throughout the discussion. These results are interesting and important to share with the scientific community, and pitching the framing correctly will go a long way towards doing so. We feel that the paper needs substantial revision to meet the standards of Nature Communications.

1. As written, the paper feels like a litany of statistical comparisons with a series of disjunct significant effects that are addressed in a random order in the discussion. A more understandable way to set up this paper would be a framework of (1) is there a beta-diversity response to global change factors (answer: yes, but there is a 50:50 split between homogenizing and differentiating) and then importantly (2) what factors drive the homogenizing vs differentiating (do you see homogenization more frequently in specific ecosystems, global change factors, etc). By framing the paper around explaining the fascinating trends of some responses being homogenization and others being differentiation would be much more compelling than the current approach that maintains the same analytical question (is there a response) throughout the manuscript. Then the discussion could be structured around what is driving the positive or negative response. Instead, it just feels like a deep dive into a bunch of random things because those are the things that wound up being significant.

[Response 1] Thanks so much for your constructive comments. We tried our best to improve the manuscript and made some changes to the manuscript. And here we did not list the changes but marked in blue on the revised paper.

2. The abstract could tie concepts together more. It currently reads like an onslaught

of information, without key motivations for those results or any underlying threads. Additionally, the statistical methods in the abstract are unclear. The authors use linear regression and random forest analysis to do what exactly or address what specific questions? The summary of their results here suggests that the variance in response to anthropogenic interventions skews results towards non-significance, when in fact the difference in beta-diversity effects is significant. Invasions always led to a decrease in beta-diversity, while other interventions varied in response direction. Climate change was the largest homogenizing force for functional beta-diversity, and land management was most significant for phylogenetic beta diversity. Finally, the use of the term “dimensions” to indicate taxonomic, functional, and phylogenetic beta-diversity is a little confusing because it is not well defined in the abstract. Clarification is needed.

[Response 2] Thank you for your valuable comment. We have completely revised the abstract in the revised version. In addition, we appreciate the suggestion to clarify the reference to "all three dimensions." We have included a clearer definition in the abstract, where we specify that the three dimensions of biodiversity refer to taxonomic, functional, and phylogenetic diversity. We will also ensure that these terms are further defined and discussed in the Introduction to provide better clarity for readers throughout the manuscript.

3. The introductory paragraph relies heavily on one citation (Dornelas et al), which is a great paper. But a more varied set of references would be nice to support these points.

[Response 3] We appreciate your suggestion and have addressed it by replacing the over-reliance on a single citation (Dornelas et al.) with a more diverse set of recent studies. In line with your feedback on the overall writing logic of the manuscript, we have thoroughly revised the introduction to ensure a broader and more balanced set of references to support the key points.

4. Word choice should also be carefully considered throughout the manuscript. Some specific choices include “dimensions” of beta-diversity (needs a better definition/explanation to use this term as many readers might not be familiar with

dimensions of biodiversity in the NSF sense), “direction of variation” in beta-diversity (ln 54; variance implies spread in both directions around a mean, so how can it have a direction?), and human intervention (which sounds like a positive restoration-based term, when really it means a negative thing here). Similarly, “treatment” in the methods should be clarified to be a statistical treatment, rather than an experimental treatment.

[Response 4] Thanks for your careful checks. As suggested, we have replaced "variation" with "change", "human intervention" with "human activities" and "treatment" with "statistical treatment" to ensure clarity. We have also reviewed the manuscript to ensure these terms are used consistently throughout.

5. The methods section needs to make the process of paring down a massive number of articles to a small subset (don’t relegate this information to the supplement). The initial number of papers does not make sense with the numbers that were excluded to get to the final number included. Additionally, it is very unclear how the 47 papers were chosen for the ordination analysis, which needs to be added into the methods.

[Response 5] Thank you for your constructive comments. We have now incorporated the full process of reference screening and data abstraction into the methods section of the revised manuscript, as you suggested. Regarding the 47 papers that used ordination analysis to represent β -diversity, they were selected through the same screening process described earlier. However, unlike the other studies, these 47 papers were analyzed using a different method to calculate the log-response ratio (see Equations 5 and 6), which is now clearly explained in the methods section.

6. Ln 304-305: What does it mean that the correlations were predominantly from a subset of the total number of studies? How was this determined?

[Response 6] Thank you for your careful review. As described in lines 343-355 of the revised manuscript, to further explore the consistency of the responses of the three dimensions of plant β -diversity to GCFs, we conducted a binary meta-analysis of the ln *RR* values to determine the correlation coefficients (*r*) between these dimensions. The results indicated that TB was positively correlated with both FB and PB, with

correlation coefficients of 0.6306 and 0.537, respectively (Figure 4). The correlation coefficient between FB and PB was 0.2511. This analysis was based on the full dataset.

Additionally, we selected data from studies that included at least two dimensions of β -diversity in order to analyze the consistency of responses to GCFs within the same study (as represented by the points in Figure 4). Among the 75 paired observations, TB and PB both increased in 21 cases under the influence of GCFs, while in 98 paired observations, TB and FB both increased in 53 cases. Despite 15 out of 30 observations showing a simultaneous decrease, no significant correlation was found between FB and PB (Figure 4).

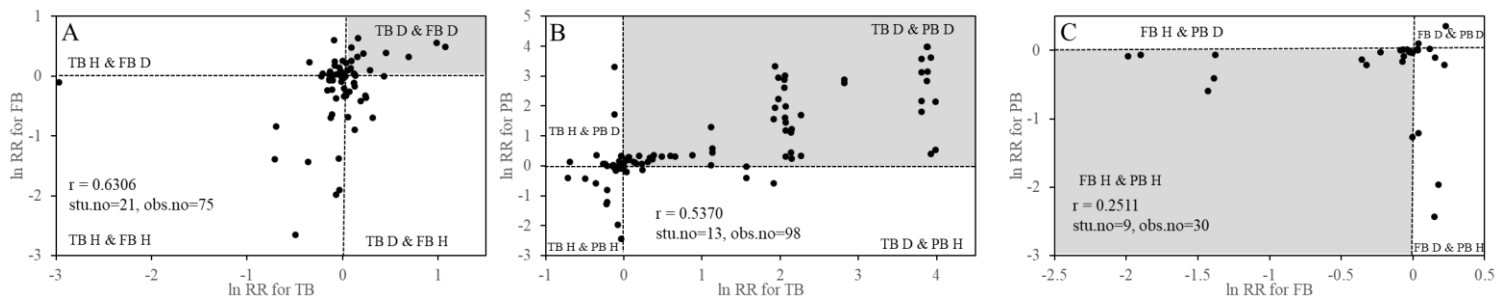


Figure 4 Bivariate meta-analysis of log-response ratios (ln RR) illustrating the impact of GCFs on plant β -diversity. (A) Taxonomic versus functional ln RR, (B) Taxonomic versus phylogenetic ln RR, and (C) Functional versus phylogenetic ln RR. The quadrants delineate the patterns of homogenization (H) or differentiation (D) for taxonomic (TB), functional (FB), and phylogenetic (PB) diversity. The gray shaded areas highlight the predominant trends. The correlation coefficient (r) is the result of a bivariate meta-analysis and represents the strength of the relationship between the variables in each panel. The points in the figure then represent data from the same study containing multiple dimensions of diversity. The sample size is indicated by the number of studies (stu.no) and observations (obs.no).

7. Figure 1: The color choice for the three dimensions of diversity is not red-green color-blind friendly. A darker shade for taxonomic or brighter shade for phylogenetic would be better. The density plots for panel a are difficult to read and don't seem necessary (don't convey any more information than the map itself). The points in panel b obscure each other. Perhaps do different shapes or sizes stacked from large (bottom) to small (top)? Panel d, the diverging x-axis for the number of studies and tests is confusing. Two bars with all positive values would be better.

[Response 7] Thank you for your constructive feedback. I have addressed each point as follows:

- 1) **Color Choice for Diversity Dimensions:** I appreciate the suggestion regarding color choices for the three dimensions of diversity. I agree that the current colors may not be the most suitable for individuals with red-green color blindness. To improve accessibility, I have modified the color scheme by using blue for the taxonomic diversity, ensuring better contrast and readability for all viewers.
- 2) **Density Plots in Panel A:** I understand that the density plots in panel A are difficult to interpret and do not add significant value beyond the map. I have removed the density plots and now rely solely on the map to convey the relevant information, as the map itself provides a clear representation of the data distribution.
- 3) **Overlapping Points in Panel B:** I note the issue of overlapping points in panel B. To improve clarity, I have adjusted the visualization by using different point sizes to avoid overlapping.
- 4) **Diverging X-axis in Panel D:** I recognize that the diverging x-axis for the number of studies and tests may be confusing. I have revised this by presenting two bars with all positive values, which will help avoid confusion and make the figure easier to interpret.

Thank you again for your insightful suggestions. I believe these adjustments will significantly improve the clarity and accessibility of the figure.

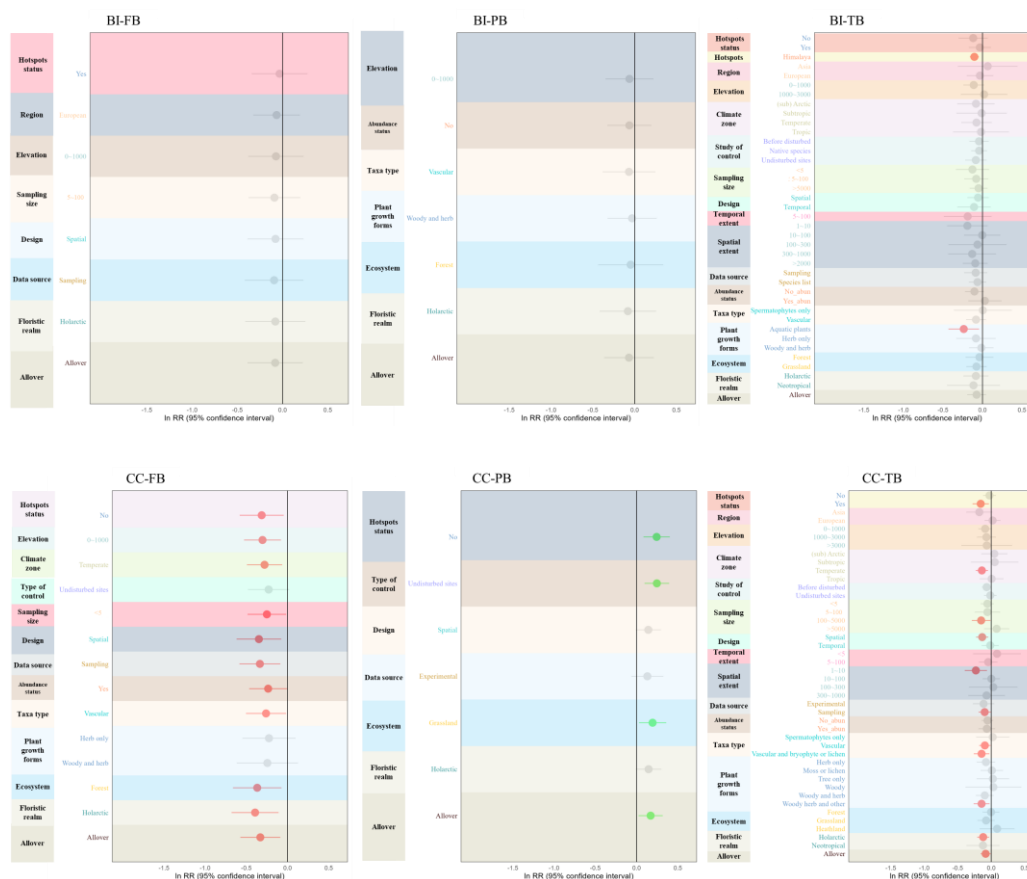
8. Is a regression analysis the right way to look at correlations among the dimensions of beta-diversity? Regression implies directionality of driver-response and has specific requirements for assumptions. A correlation seems like a better choice.

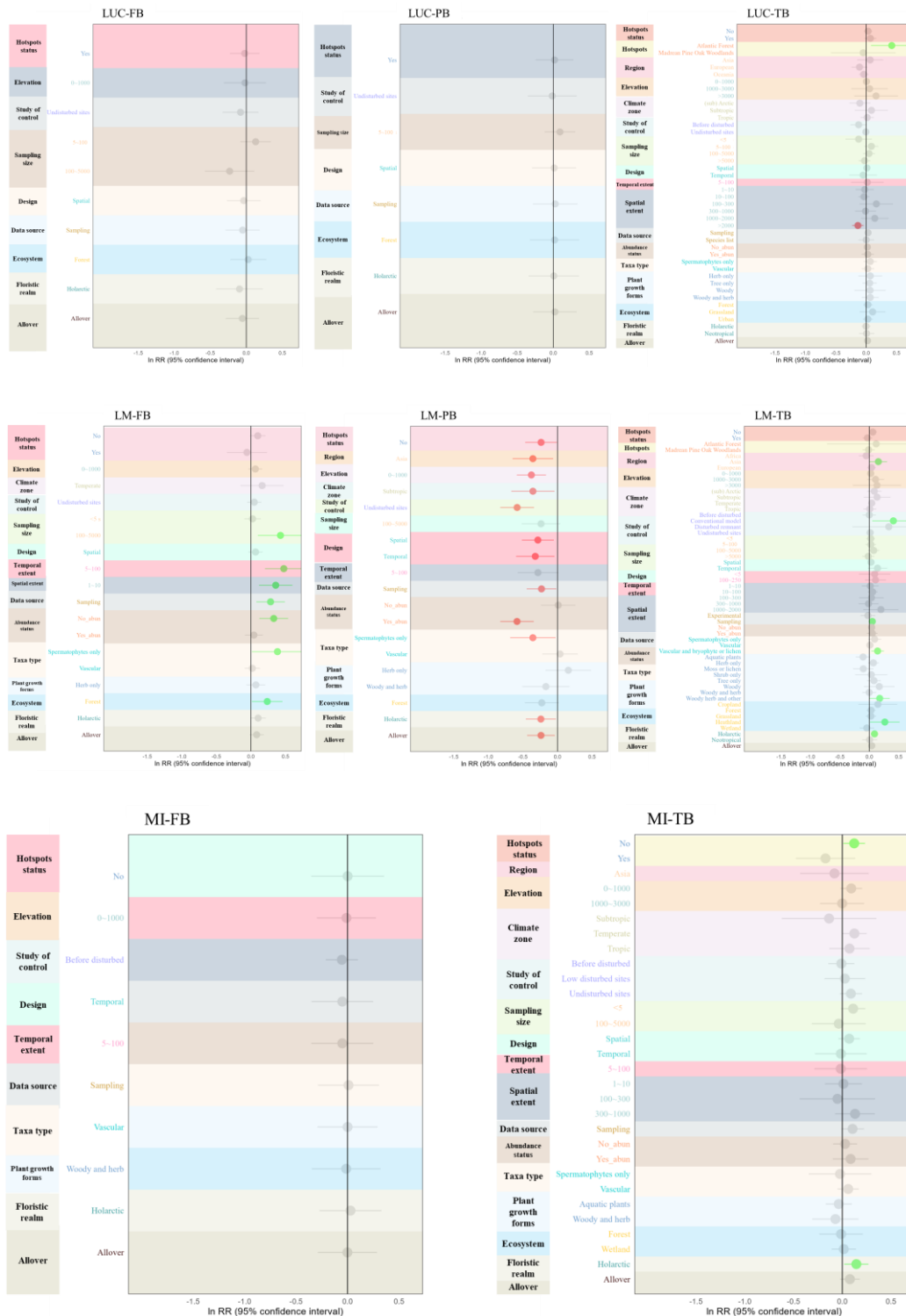
Figure 4, panel a has a fishy regression line. How much influence are the outlier points having on this line?

[Response 8] We agree with your point that regression analysis may not be the most suitable method for examining the relationships between β -diversity dimensions, as regression implies a directional driver-response relationship, which is not our aim. Our focus is on assessing the correlations between changes in these dimensions in response

to GCFs, without assuming causality. Therefore, we have revised the analysis to use a binary meta-analysis instead of regression to more accurately reflect the relationships among the changes in the dimensions of β -diversity. Please refer to our response to your sixth question for further details.

9. The distinction between panels a vs b in Figure 5 needs clarification in the legend. [Response 9] We sincerely thank the reviewers for their careful reading and valuable suggestions. To improve the clarity and understanding of the figure, we have revised both panels and updated the legend. The revised version of Figure 5 is now included in the Supplementary Information.





Supplementary Figure 5 Comparisons of the responses of TB, FB, and PB to five GCFs, categorized by hotspot status, hotspots, regions, elevations, realms, and climate zones (**left panel**), as well as by ecosystems, floristic realms, plant life forms, taxa types, abundance status, data sources, and study design factors such as temporal and spatial extent (**right panel**). To ensure the robustness and reliability of the findings by focusing on more data points, only the results with a study count of 3 or more are shown in the analysis for the second-level subgroups. Key: TB, taxonomic β -

diversity; FB, functional β -diversity; and PB, phylogenetic β -diversity. BI, biological invasion; CC, climate change; LUC, land use change; LM, land management; and MF: multiple factors.

10. The discussion should include speculation about why some “dimensions” of beta-diversity are responding while others are not. These three dimensions are all tightly linked, so a good null hypothesis is that all should respond. Why don’t they?

[Response 10] Thank you for your valuable suggestion. We have completely revised the discussion in the light of your comments and in conjunction with the revised introduction.

11. Minor typos/grammatical issues:

Ln 11 (verb tense), 42 (too many commas confuses sentence structure), 43-47 (overly complicated sentence, split it up?), 50-54 (overly complicated sentence), 59 (is it really numerous if there are only three citations?), 60-63 (what is the difference between a specific vs a global region?), 70 (what does no variations mean?), 75 (jargon heavy sentence), 87 (impacts to what?), 118 (missing word), 122 (unit needed), 161 (tiny sentence, combine with previous), 273 (clarify that it isn’t the lnRR that is indicating the significant changes, but rather the confidence interval not overlapping with 0), 413 (missing space), 469 (missing letter)

[Response 11] Thank you for your detailed comments and for pointing out these important issues. We have addressed each of the points raised, making the necessary corrections throughout the revised manuscript.

Reviewer #3 & #4 (Remarks to the Author):

This manuscript is an interesting study that evaluates the impact of human activities on plant β -diversity through a global meta-analysis. The authors examined the magnitude of the effects of different anthropogenic interventions on phylogenetic, taxonomic, and functional β -diversity, as well as the effects of other various factors on these β -diversities. While we are not experts in this field (plant ecology), we are confident in our expertise in meta-analysis. We found that the manuscript requires revisions - particularly regarding the transparency and clarity of the analyses, results, and figures, among other issues. We reviewed this manuscript using the Prisma-EcoEvo guidelines (O'Dea et al., 2021; <https://onlinelibrary.wiley.com/doi/full/10.1111/brev.12721>) and left comments below - mainly focusing on the points that seem to deviate the guidelines. For comments based on the guidelines, the corresponding Sub-item number from the guidelines is noted alongside using square brackets.

As you see below, the authors' reporting standards are suboptimal. Given what was reported, we could not assess the reliability of their results and, therefore, conclusions.

Title and abstract

12. L19: What does mean "tree- β diversity"? Please explain.

[Response 12] Thank you very much for your valuable comments. In our study, we classified plants into categories based on life forms, including aquatic plants, herbaceous plants, shrubs, trees, woody plants, combinations of woody and herbaceous plants, mosses or lichens, and so on. Subgroup analysis results showed that GCFs significantly reduced the β -diversity of trees (Figure 3). We acknowledge that the original wording may have caused confusion, and we have made corresponding revisions throughout the manuscript to clarify such issues.

13. L64-65: It is unclear what "all three dimensions" refer to at this point (though it becomes apparent later that they are functional, phylogenetic, and taxonomic plant β -diversity). However, we noticed that no explanation or definition of these three different

types of plant β -diversity was provided in the whole manuscript. We would suggest including this in the Introduction.

[Response 13] Thank you for your valuable comment. We appreciate the suggestion to clarify the reference to "all three dimensions." We have included a clearer definition in the abstract, where we specify that the three dimensions of biodiversity refer to taxonomic, functional, and phylogenetic diversity. We will also ensure that these terms are further defined and discussed in the Introduction to provide better clarity for readers throughout the manuscript.

Methods

14. We could not find the information on metadata, raw data, and scripts used for analyses from this manuscript. We strongly suggest they provide all the information necessary to replicate this study (at the review stage so we can assess them). Sharing them will enhance transparency, allowing others to validate or build on the findings. If there is a reason why they could not do it, the authors should clearly state that. - [18.1-4]

[Response 14] Thank you very much for your suggestion. We have uploaded the raw data into the submission system as Supplementary Material.

15. L95: The authors should describe what sources of information were sought. For example, was the search limited to only published studies or included grey literature (e.g. personal communications, theses, or unpublished studies)? - [5.1]

L97-99: We would like them to provide the exact search strings used for each database. When we entered the search strings provided by the authors into the search engines of Web of Science, Google Scholar, and China Knowledge Resource, the total number of results was significantly lower than the number of papers reported by the authors in the manuscript. Also, we would like to know whether the authors set any temporal restrictions on the database searches. - [5.3-4]

L99: "35206 unique studies" is not a number that includes duplicate studies.

[Response 15] Thank you for your valuable feedback. We have updated the manuscript

to provide more details about the sources of information sought. The literature search was conducted in Web of Science, Google Scholar, and the China Knowledge Resource Integrated Databases between 29 March and 31 August 2023. The search in Web of Science and Google Scholar used the following keyword combinations: (“beta diversity” OR “ β diversity”) and (“biotic homogen*” OR “biotic differentiat*” OR “biotic heterogen*” OR “biological homogen*” OR “biological differentiat*” OR “biological heterogen*”). In the China Knowledge Resource Integrated Databases, we translated these keywords into Chinese for the search.

A total of 35,206 results were retrieved across all three search engines (see Supplementary Fig. 1), including duplicates. We did not apply any temporal restrictions during the search. Additionally, we examined the reference lists of relevant review articles, as well as subsequent studies citing these reviews (e.g., Catano et al. 2017, Kramer et al. 2022, McKinney 2006, Petsch et al. 2022, Rolls et al. 2023, Soininen et al. 2018) (229 studies). The search was limited to published studies.

16. L117: How many people were involved, and how did they contribute during screening? And more importantly, how they evaluate the reliability of their screening and extraction (e.g. double-screeing, double-checking) - [6.2]

[Response 16] Thank you for your valuable feedback. To clarify, a total of four researchers were involved in the screening and data extraction process. Initially, two researchers (HLL and HY) independently screened all studies based on titles, keywords, and abstracts to identify studies focusing on the effects of human-induced global change on plant β -diversity, excluding 4,766 papers. Next, two researchers (HLL and MN) independently conducted a full-text review of the remaining 2,059 articles. In cases of uncertainties or disagreements, a third reviewer (PL) was consulted to resolve the issues through discussion.

To ensure the reliability of the screening and extraction process, double screening and double checking were applied at each stage. The screening was conducted independently by the researchers, and disagreements were resolved by consensus, involving a third reviewer when necessary. This approach helped minimize potential

bias and ensure the consistency and accuracy of the study selection process.

17. L118-120: Why the total test number ($n = 294$) from adding functional, phylogenetic, and taxonomic plant β -diversity does not match the number reported in the manuscript ($n = 256$)? We might have overlooked something, but is it the case that some studies tested multiple β -diversity?

[Response 17] Indeed, some studies include multiple dimensions of biodiversity, which could explain the discrepancy in the total test number. We have taken this into account in our analysis, but the discrepancy arises because these studies report separate results for each biodiversity dimension.

18. L132: How did they handle the missing data or ambiguous information? - [7.4]. Also, it should be better to state who collected data and the number of extractions checked for accuracy by co-authors. - [7.5-6]

[Response 18] We independently extracted measures of central tendency (mean and median), dispersion (standard deviations, standard error, confidence intervals, and amplitude), and sample sizes for both the control (sites not exposed to GCFs) and treatment (sites affected by GCFs) groups in the selected studies by two groups (HLL and WWX combined as group one, and YC and YH combined as group two). In cases where studies did not report outcomes correctly or lacked sufficient original data, they were excluded from our analysis.

19. L147-153: We would like to provide more specific and detailed explanations for all the variables included in the six categories. For example, what is meant by 'plant functional type,' and what categories are included? Additionally, were any variables among they extracted either not used in the analysis or excluded due to insufficient information? - [8.1-4]

Also, we suggest the authors assess the quality of the primary studies included in their meta-analysis and consider potential biases in the results. Because assessing them is crucial for ensuring the reliability and validity of the meta-analysis results. Without

considering these factors, the conclusions could be skewed by low-quality/biased studies. - [9.1-2; 20.5]

[Response 19] We sincerely thank the reviewers for their insightful suggestions. We have revised the manuscript to address both the request for more specific explanations of the variables and the suggestion regarding the quality assessment of primary studies.

1) In response to your request for more specific details about the variables included in the seven categories, we have provided clearer definitions in the manuscript. We also acknowledge that the term "plant functional type" was not used correctly and have replaced it with "plant life forms," which is a more appropriate description. The seven categories and their specific variables are as follows:

- Five GCFs: biological invasion (BI), climate change (CC), land use change (LUC), land management (LM), and multiple factors (MF).
- β -diversity dimension: Three dimensions of β -diversity are considered: taxonomic β -diversity (TB), functional β -diversity (FB), and phylogenetic β -diversity (PB).
- Geographic location: The geographic variables include latitude, longitude, region, realm, elevation, and whether the study site falls within a biodiversity hotspot.
- Ecological setting: This includes climatic zone, ecosystem type, flora realm, mean annual precipitation (MAP, in mm), and mean annual temperature (MAT, in °C).
- Study design: This includes information on whether the study focused on temporal or spatial design, sample size, control sample type, temporal extent, and spatial extent.
- Data characteristics: data sources and whether the data were abundance-weighted.
- Research subject: This includes the examined plant taxa and life forms (instead of "plant functional types").

Each of the aforementioned data points was sourced either directly from the study reports themselves or by utilizing the latitude and longitude of the study sites. In cases where a study did not report the latitude and longitude, we estimated approximate coordinates by geocoding the site name using Google Earth 7.0 and then obtained the elevation based on these coordinates. Similarly, if the study did not provide mean annual temperature (MAT) and mean annual precipitation (MAP), we retrieved these

values from WorldClim using the geographic location (latitude and longitude) of the study site.

For studies that covered large areas (such as entire countries or the globe) without specific coordinates, we did not assign environmental metrics to them, which represented 6% of the total studies. It's worth noting that during the specific analysis, if a study did not involve a certain type of data (such as sample size), it would not be included in the effect analysis related to that data.

2) In response to the reviewer's comment, we have taken steps to reduce potential biases in our meta-analysis. Specifically, we have only included studies that have undergone peer review and are publicly published, ensuring a level of quality in the primary data. Additionally, based on the reviewer's suggestion and the Prisma-EcoEvo guidelines (O'Dea et al., 2021), we have excluded studies with a sample size of fewer than 3 from our analysis. Consequently, our meta-analysis is now based on 1604 observations from 256 studies, which helps to ensure the reliability and validity of our results.

20. L187: equation (1) - parentheses for ratio will be needed.

L187-188: These are not the best estimators - see this paper

Lajeunesse MJ. Bias and correction for the log response ratio in ecological meta-analysis. *Ecology*. 2015 Aug;96(8):2056-63.

[Response 20] Thank you for your valuable feedback. In our analysis, we used the log response ratio ($\ln RR$) as an effect size measure, and we are aware of the potential biases associated with this method, as highlighted in Lajeunesse (2015). To address these concerns, we estimated effect sizes using the Delta method recommended by Lajeunesse (2015), which provides a more reliable and bias-correcting method of calculating $\ln RR$, particularly in ecological meta-analyses. In response to your suggestion regarding the equation, we have revised it to include parentheses for the ratio to improve clarity. The updated and additional equations are as follows:

$$\ln RR = \ln\left(\frac{\bar{X}_t}{\bar{X}_c}\right) \quad (1)$$

$$v_i = \frac{s_t^2}{n_t \bar{x}_t^2} + \frac{s_c^2}{n_c \bar{x}_c^2} \quad (2)$$

$$RR^\Delta = \ln RR + \frac{1}{2} \left[\frac{(SD_T)^2}{n_T \bar{x}_T^2} - \frac{(SD_C)^2}{n_C \bar{x}_C^2} \right] \quad (3)$$

$$var(RR^\Delta) = var(v_i) + \frac{1}{2} \left[\frac{(SD_T)^4}{n_T^2 \bar{x}_T^4} + \frac{(SD_C)^4}{n_C^2 \bar{x}_C^4} \right] \quad (4)$$

21. L206: We do not think variance can be calculated using equation (3). Provide an equation for sampling variance.

[Response 21] Thank you for your comment. We have added the formula for calculating the variance. The variances (v) were calculated using the following equation:

$$v = \frac{s_t^2}{n_t \bar{D}_t^2} + \frac{s_c^2}{n_c \bar{D}_c^2} \quad (4)$$

where n_t and n_c are the sample sizes of the variable in the treatment and control, respectively; s_t and s_c are the standard deviations of the variable in the treatment and control, respectively.

22. L214-249: The authors ticked the item about Bayesian analysis in the Nature portfolio reporting summary but did not provide information on the choice of priors or Markov Chain Monte Carlo setting - they should have given the information about this in the Result section. Please also provide all packages, functions, and their versions used to run the models. - [13.2-5]

[Response 22] Thanks for your attention. This study indeed does not include the Bayesian analysis. We have revised this choice in the Nature portfolio reporting summary.

23. L215-232: We guess the authors used multilevel meta-analyses from their descriptions. It would be great to see two different levels of heterogeneity (e.g. I^2 at the effect size level and I^2 at the study level).

[Response 23] Thank you for your suggestion. We have included the I^2 statistic in the two-level subgroup analyses in the revised manuscript.

24. L218: Did the authors consider cohort ID (groups of the same subjects), shared control ID (indicating effect sizes sharing control groups), as well as study ID? - [14.1-3]

Note that the related points were recently explained in this paper with detailed tutorials - see Nakagawa S, Yang Y, Macartney EL, Spake R, Lagisz M. Quantitative evidence synthesis: a practical guide on meta-analysis, meta-regression, and publication bias tests for environmental sciences. *Environmental Evidence*. 2023 Apr 24;12(1):8.

[Response 24] Thank you for your insightful suggestion. In our previous manuscript, we considered the influence of the study ID but did not fully account for the non-independence of data arising from shared control groups within the same study. In response to your comment, we have revised our analysis accordingly. Specifically, we now incorporate both the study ID and control ID as random effects in our mixed-effects models. The control ID is nested within the study ID, meaning that each study has its own one or several sets of control groups, and the effect sizes within a study are potentially correlated due to this structure.

We have also referred to the detailed guidance provided in Nakagawa et al. (2023) to implement these adjustments in the context of environmental science meta-analysis. Below is the updated model:

$$\text{Effect Size}_{i,j} = \beta_0 + (1|\text{Study ID}) + (1|\text{Study ID} : \text{Control ID}) + \epsilon_{i,j}$$

Where $\text{Effect Size}_{i,j}$ represents the effect size for the i -th observation within the j -th study. β_0 is the fixed effect (the total effect of GCF on plant beta diversity). $(1|\text{Study ID})$ accounts for variation between studies. $(1|\text{Study ID} : \text{Control ID})$ accounts for variation due to control groups nested within each study. $\epsilon_{i,j}$ is the residual error term.

25. L237: The authors used the fail-safe number, but it is heavily criticized. Consult a recent review of publication bias.

Nakagawa S, Lagisz M, Jennions MD, Koricheva J, Noble DW, Parker TH, Sánchez-Tójar A, Yang Y, O'Dea RE. Methods for testing publication bias in ecological and evolutionary meta-analyses. *Methods in Ecology and Evolution*. 2022 Jan;13(1):4-21.

[Response 25] Thank you for your suggestion. Considering the concerns regarding the reliability of the fail-safe number, we have removed the results based on this metric from the manuscript. Instead, we focused on more robust methods for assessing publication bias, such as the Egger regression test and the trim-and-fill method, as recommended by recent literature (Nakagawa et al., 2022). We believe these methods provide more reliable estimates of the potential effects of publication bias and offer a better evaluation of the robustness of our results.

26. L240: We wonder why the authors used linear regression instead of meta-regression. We suspect errors could not be incorporated into Random Forest Models. Actually, unweighted models could be unbiased, but it depends on whether the data set has publication bias. Weighted models will be much better when the dataset contains publication bias.

[Response 26] Thank you for your suggestion. In response to your comment, we have revised our analysis to use meta-regression instead of linear regression. We appreciate your valuable feedback, which has led to an improvement in the statistical approach used in our study. The revised figures are shown in Figure 6 of the revised manuscript and Figure 6 of supplementary information.

Results

27. Please...

provide the number of effect sizes for the data included in meta-analyses (does the number of tests equal effect size?). - [20.1]

state a summary of key characteristics for reported outcomes. In particular, the authors did not provide any summary of the study design and research subject. - [20.3]

report the results section based on statistical analysis; do not simply state that there was a significant positive or negative effect without providing statistical details or reporting only the p-value. Always describe the full statistical results along with the findings. - [21.1, 23.1-4]

[Response 27] Thank you for your valuable feedback. We have carefully revised the

manuscript to address your concerns. We have now included the number of effect sizes for the data included in the meta-analyses and clarified that the number of tests corresponds to the number of effect sizes used in the analyses. In the results section, we have ensured that all statistical analyses are presented in full, including effect sizes, confidence intervals, and relevant *p* values. We have avoided simply stating whether the effects were positive or negative and have provided a more comprehensive description of the statistical results. We hope these revisions address your concerns and improve the clarity of our manuscript. Thank you again for your helpful suggestions.

28. The relationships between the figures in the Results/Discussion sections and the results are difficult to understand. Please revise to clarify which parts of the figure correspond to specific results.

[Response 28] Thank you for your suggestion. In the revised version, we have added figure numbers corresponding to the relevant text in the "Results and Discussion" sections to make it clearer which parts of the figures relate to specific results.

29. L261-262: Why did the authors show the number of tests for homogenisation and differentiation in CC, LUC, and BI, but not in LM and MI?

[Response 29] Thank you for your insightful comment. In the revised manuscript, we have now included the relevant data for LM and MF (MI has been revised to MF) as well. The updated sentence is as follows:

"In contrast, LM (with 400 tests for differentiation and 356 tests for homogenization) and MF (with 46 tests for differentiation and 37 tests for homogenization) exhibited the opposite trend."

30. L274: Please explain the meta-regression analysis in the methods section (e.g. how many and what moderators are used? How select the final model?) - [15.1-3; 20.2,4]

[Response 30] Thank you for your insightful comment. We have made the necessary revisions to the manuscript to clarify the meta-regression analysis, as follows:

First, we conducted a meta-regression analysis to examine the relationship between the

effect sizes ($\ln RR$) and the following moderators: latitude, longitude, elevation, mean annual temperature (MAT), mean annual precipitation (MAP), spatial extent, temporal extent.

We performed the analysis using the *metafor* package in R (Viechtbauer, 2010). The analysis was weighted by the variance of each observation and utilized the Restricted Maximum Likelihood (REML) method to estimate the model parameters. Regarding missing data, the temporal extent variable had over 50% missing values, so it was removed from the analysis. For the remaining variables with less than 50% missing data, missing values were imputed using the Random Forest algorithm with default settings, implemented in the *missForest* package (Stekhoven et al., 2012).

We hope this clarifies the approach taken in the meta-regression analysis. Thank you again for your helpful feedback.

31. L276: What is this percentage value? The back-transformed values of $\ln RR$?

[Response 31] Thank you for your valuable comment. To clarify, the 12% reduction in the β -diversity of tree initially reported refers to the back-transformed values of the natural log response ratio ($\ln RR$). This represents the effect size in the original scale, after converting the $\ln RR$ estimates into percentage values. Additionally, we would like to correct this figure: based on our reanalysis, GCFs actually reduced the β -diversity of trees by 15% ($p < 0.05$). We will revise the manuscript to reflect this updated result and provide a clearer explanation of the statistical approach.

32. L301-306: Please explain the analyses regarding this paragraph's results in the **Methods section**.

[Response 32] Thank you for your constructive comment. In response, we have revised the Methods section to provide a more detailed explanation of the statistical analyses used for the results presented in this paragraph. The updated text is as follows: “Additionally, to investigate the consistency of the three dimensions of plant β -diversity (taxonomic, functional, and phylogenetic β -diversity) in response to GCFs, we conducted correlation analyses on the $\ln RR$ values of these dimensions

using bivariate meta-analysis.”

33. Discussion

- Did they consider limitations and their influence on the generality of their conclusions?
- Although we do not know the number of data points for each variable, should not limitations such as gaps in the available evidence (e.g. geographical research biases or researched plant type bias) be considered? - [25.6]
- Please refer to the comments in the Result section as well.

[Response 33] Thank you for your insightful comments. We have added a discussion on the limitations of the study towards the end of our revised manuscript, as follows: “However, we also note that well-known biases, such as greater sampling effort in the global north compared to the Southern Hemisphere (especially in Australia), persist in the biodiversity data compiled for this study (Figure 1). This significantly limits the generalizability of our findings to well-studied regions of the planet.”

34. References

Please provide a reference list of all eligible studies for the meta-analysis. Also, Please list included studies as referenced sources. - [27.1-2]

[Response 34] Thank you for your helpful suggestion. We have now included the reference list of all eligible studies for the meta-analysis in Supplementary Material 2, which contains the 256 studies that were included in this research.

We hope this addition meets your expectations. Thank you again for your valuable feedback.

Figures

Overall, we are sorry, but the figures in this manuscript are not easy for us to understand. Additionally, the colour scheme in some figures/supplementary figures does not appear considerate to people with colour vision deficiencies.

35. Figure 1: Figure 1 was quite challenging to understand. First, we would like clearer separations between sections A and C.

Second, in A, please indicate what the insets' X-axis (top) and Y-axis (right) represent. The circles on the world map overlap too much and are hard to interpret. For example, what does this overlap mean? Third, in B, what does each plot represent? (we assumed it was the number of tests, but if so, should not taxonomic β -diversity have the most?) Finally, in C, this might be because it is outside our area of expertise, but in what situation is elevation over 3000 m in freshwater ecosystems?

[Response 35] Thank you for your valuable feedback on Figure 1. We have made several revisions to improve clarity and address your concerns:

- 1) **Clearer Separations Between Sections A and C:** To enhance the distinction between sections A and C, we have adjusted the color configuration. Section A now displays the global distribution of observations that include the different dimensions of diversity included in this study. Section C shows the distribution of these diversity dimensions across realms and elevations. This should make the separation between the two sections clearer.
- 2) **Clarification of Axes in Section A:** In section A, we have now labeled the axes of the insets for clarity. The X-axis (top) represents longitude, and the Y-axis (right) represents latitude. These labels should provide additional context for interpreting the insets more easily.
- 3) **Overlap of Circles in the World Map:** The circles on the world map in section A represent studies that include multiple dimensions of biodiversity (taxonomic, functional, and phylogenetic). The overlap of circles indicates that a single study addresses more than one dimension of biodiversity. To clarify this, we have adjusted the circle sizes to be uniform, removing the reflection of sampling size, which should now make the map easier to interpret.
- 4) **Interpretation of Plots in Section B:** Section B displays the distribution of diversity dimensions with respect to mean annual temperature (MAT) and mean annual precipitation (MAP). We have adjusted the sizes of the circles for each dimension to prevent overlap and enhance visibility. Additionally, we reviewed the representation of taxonomic β -diversity to ensure accuracy, as it is expected to have

the largest number of studies. This has been corrected in the figure.

5) Elevation Over 3000 m in Freshwater Ecosystems (Section C): We acknowledge that the example in section C, which includes an elevation over 3000 m, may seem unusual for freshwater ecosystems. The study referenced (Rosero et al., 2021) was conducted in a small glacierized catchment in Ecuador, located at 5018 m a.s.l. at the Carihuairazo volcano (01°24'25"S, 78°45'00"W). This site is an exceptional case relevant to our study, and we have updated the figure caption to provide this additional context for clarity.

We hope these revisions effectively address your concerns and improve the clarity of Figure 1. Thank you once again for your constructive comments.

36. Figure 3: What do the numeric values in parentheses within the figure represent? Also, please explain both Fig. 3A and Fig. 3B in the legend separately.

[Response 36] Thank you for your comment. The numeric values in parentheses within the figure represent the number of studies and the number of observations, respectively. We have also provided separate explanations for both Fig. 3A and Fig. 3B in the legend. The revised descriptions are as follows. Additionally, we have made corresponding modifications to other figures of the same type for consistency.

“Figure 3 The effect size (lnRR) of GCFs on plant β -diversity varies across different GCFs, dimensions, hotspot status, hotspots, regions, elevations, realms, and climate zones (A), as well as across ecosystems, floristic realms, plant growth forms, taxa types, abundance status, data sources, and study design aspects such as temporal extent and spatial extent (B). Data are presented as mean $\pm$ 95% CIs of the estimated effect sizes. Nonsignificant changes are represented by gray dots, while significantly negative and positive effects are represented by red and green dots, respectively. The numbers in parentheses indicate the number of studies and the number of observations, respectively.”

37. Figure 4: We could not find the method related to this. One cannot just use the correlation to analyse the relationships between 2 sets of effect sizes (e.g. pseudo-

replication and also there are two correlations that can be calculated at the between- and within-study levels). One should use bivariate meta-analysis. See this link (but your model will be more complex): <https://www.metafor-project.org/doku.php/analyses:vanhouwelingen2002>

[Response 37] Thank you for your valuable feedback. You are correct that simply using correlation analysis may not be appropriate for analyzing the relationships between two sets of effect sizes, as it could lead to pseudo-replication issues. In response to your comment, we have updated the methods section of the manuscript to include the following clarification:

"Additionally, to investigate the consistency of the three dimensions of plant β -diversity (taxonomic, functional, and phylogenetic β -diversity) in response to GCFs, we conducted Pearson correlation analyses on the $\ln RR$ values of these dimensions. This analysis was performed only for studies that included data from at least two dimensions of β -diversity."

As you suggested, we acknowledge that bivariate meta-analysis would be more appropriate for this type of analysis. However, due to the limitations of the available data, we opted to conduct the correlation analysis and present the results graphically. We used this approach to display the consistency of the direction of change in β -diversity dimensions due to GCFs, based on data where at least two dimensions were reported.

We have updated the figure legend and manuscript accordingly to reflect these changes.

38. Figure 5: Why are the marks on the y-axis of Fig. 5b shown in both black and grey, and what is the meaning of this difference? Also, what do the numbers in parentheses mean (e.g. TB×MI (16, 69))?

[Response 38] Thank you for your comment. In order to improve the clarity of the graphs and better illustrate the results, we have rearranged the graphs for this part of the results. Please see Figure 5 in the revised supplementary information.

39. Supplementary Figure 1: In this PRISMA-like flowchart, the authors should report brief reasons for exclusion from the full-text screening stage. Also, the number of Reports screened is 6825, not 7825. [19.1-4]

[Response 39] Thank you for pointing out these issues. In the revised version, we have added a clear summary of the reasons for excluding studies during the full-text screening stage. Additionally, we have corrected the number of reports screened. The correct number is indeed 6825, not 7825, and we have updated this figure accordingly in the revised manuscript and flowchart.

40. Supplementary Figure 2: We wondered how the authors compared the responses of plant β -diversity to global change drivers using data from studies that used different analytical methods.

Supplementary Figure 2b: How many studies are included in each of these four analytical methods?

[Response 40] Thank you for your valuable comment. We appreciate your concern regarding the comparison of plant β -diversity responses to GCFs using data from studies that employed different analytical methods. To address this, we have conducted a meta-analysis to assess how different analytical methods influence the results.

We have included the number of studies corresponding to each analytical method in the figure, as requested. This allows for a clearer understanding of how the different analytical methods are distributed across studies and their effect on plant β -diversity response to GCFs.

41. Supplementary Figures 3 and 4: Please see the comment on Figure 3.

[Response 41] Thank you for your comment. As per our response to the reviewer's feedback on Figure 3, we have made the corresponding revisions to Supplementary Figures 3 and 4. The updated descriptions are as follows:

Supplementary Figure 3: The effect size (lnRR) of GCFs on the three dimensions of plant β -diversity—taxonomic (TB), functional (FB), and phylogenetic (PB)—varies across different GCFs, hotspot status, hotspots, regions, elevations, realms, and climate

zones (left panel), as well as across ecosystems, floristic realms, plant life forms, taxa types, abundance status, data sources, and study design aspects such as temporal extent and spatial extent (right panel).

Supplementary Figure 4: The effect size ($\ln RR$) of GCFs on the five GCFs—biological invasion (BI), climate change (CC), land use change (LUC); LM, land management (LM) and multiple factors (MF)—varies across different GCFs, hotspot status, hotspots, regions, elevations, realms, and climate zones (left panel), as well as across ecosystems, floristic realms, plant life forms, taxa types, abundance status, data sources, and study design aspects such as temporal extent and spatial extent (right panel).

42. Supplementary Figure 5: Please see the comment on Figure 5.

[Response 42] Thank you for your comment. In order to improve the clarity of the graphs and better illustrate the results, we have rearranged the graphs for this part of the results. Please see Supplementary Figure 5 in the revised version of the manuscript for the updated layout.

43. Supplementary Figure 8: There is no explanation regarding to this Fig.

[Response 43] Thank you for your comment. We have revised the presentation of these types of figures and added the necessary textual explanation in the updated figure legend. Please refer to Figure 6 for the revised supplementary information, which now includes a clearer explanation of the figure. The updated text is as following:

Supplementary Figure 6: Meta-regression results (red line) showing the relationships between mean annual temperature, mean annual precipitation, elevation, latitude, longitude, sample size, spatial extent, temporal extent, and the response ratios ($\ln RR$) of plant β -diversity in response to specific types of global change factors (GCFs).

References

- Nakagawa S., Lagisz M., Jennions M.D., et al. (2022) Methods for testing publication bias in ecological and evolutionary meta-analyses. *Methods Ecol Evol* 13(1):4-21.
- Nakagawa S., Yang Y., Macartney E.L., et al. (2023) Quantitative evidence synthesis: a practical guide on meta-analysis, meta-regression, and publication bias tests for environmental sciences. *Environ Evid* 12: 8.
- O'Dea R.E., Lagisz M., Jennions M.D., et al. (2021) Preferred reporting items for systematic reviews and meta-analyses in ecology and evolutionary biology: a PRISMA extension. *Biol Rev* 96: 1695-1722.
- Rosero P., Crespo-Pérez V., Espinosa R., et al. (2021) Multi-taxa colonisation along the foreland of a vanishing equatorial glacier. *Ecography* 44: 1010-1021.

For all reviewers:

We sincerely apologize for the oversight in our previous revision, where we failed to indicate the line numbers corresponding to the changes made. We fully recognize that this may have caused inconvenience during the review process. In this revision, all modifications have been carefully reviewed and are now clearly annotated with the appropriate line numbers for ease of reference.

We are deeply grateful for the reviewers' thoughtful and constructive feedback. In response, we have undertaken a thorough and substantive revision of the manuscript. This includes methodological improvements, structural adjustments, and refinement of key arguments throughout the text. These changes have enhanced the clarity, analytical rigor, and overall focus of the study.

Reflecting these substantial revisions, we have also modified the manuscript title from "Global homogenization of plant communities in the Anthropocene: A meta-analysis" to "How did global homogenization of Anthropocene plant communities occur? A meta-analysis". The revised title more accurately reflects the updated scope of the manuscript and underscores its emphasis on mechanisms and context-dependent outcomes, rather than assuming a uniform global trend.

We hope the revised manuscript better aligns with the reviewers' suggestions and expectations, and we again thank all reviewers for their valuable input in strengthening this work.

Reviewer #1 & #2 (Remarks to the Author):

The revised manuscript is substantially improved from the previous iteration and we thank the authors for their careful attention to reviewer comments. However, we feel the revision still remains unclear as to how the different analyses connect to each other and how the analyses relate back to the questions posed at the end of the introduction. Finally, the figures need clarification regarding what they are attempting to show. We outline our primary concerns below.

(1) The introduction brings the reader's attention to three specific questions. Throughout the results and discussion section, the many analyses presented do not seem to directly tie back into these questions. The authors should consider revisions to the methods, results, and discussion in order to highlights how their analyses answer those three questions.

[Response 1] We sincerely thank the reviewer for this insightful comment. We agree that in the original version of the manuscript, the three research questions outlined in the Introduction exhibited some conceptual overlap, which may have obscured the connection between the questions and the analyses presented in the Methods, Results, and Discussion sections.

To improve conceptual clarity and ensure that each research question is clearly addressed throughout the manuscript, we have revised and reformulated the three questions as follows:

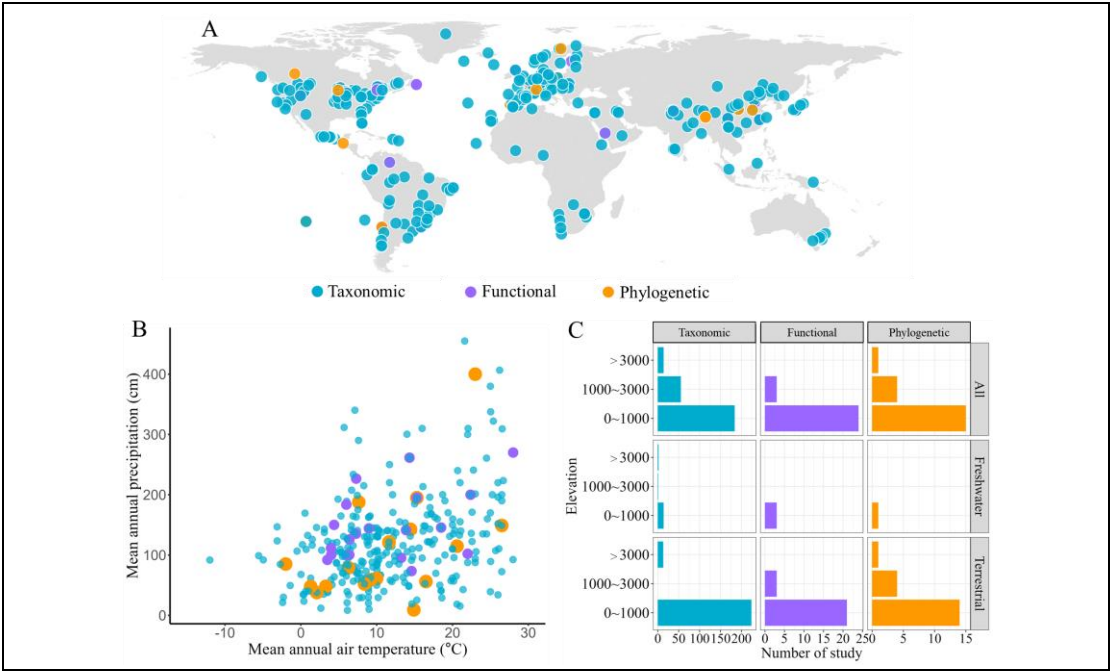
1. Is there a global trend toward biotic homogenization of plant communities under GCFs?
2. How do biogeography, environmental context, focal taxa, and study design shape the direction and strength of GCF effects on plant β -diversity?
3. Are the impacts of different GCF types consistent across the three dimensions of plant β -diversity?

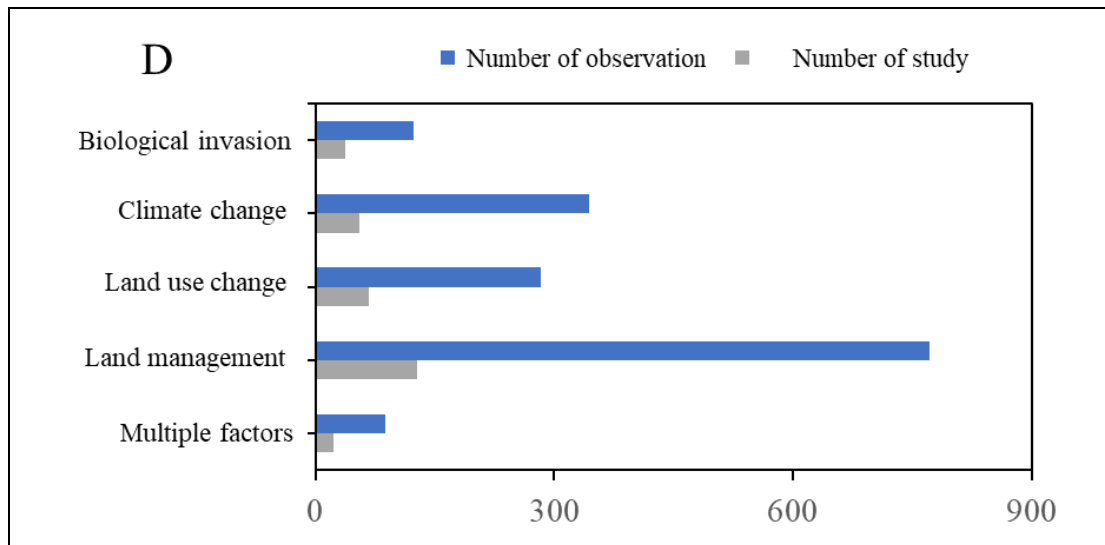
In the *Methods* section, we have added three clearly delineated subsections under

Section 2.6 (Statistical analyses) to clarify the purpose of each analysis and how it aligns with the corresponding research questions (see Lines 265–333). In the *Results* section, following the general literature patterns (Section 3.1), we have explicitly organized the remaining results to follow the structure of these three research questions. Similarly, in the *Discussion*, the first question is addressed in the first two paragraphs, the second is in the third paragraph. The subsequent paragraphs integrate findings relevant to both the second and third questions, thereby ultimately addressing the overarching theme of the manuscript: “How did global homogenization of Anthropocene plant communities occur?”. We believe that these revisions have significantly improved the logical flow and coherence of the manuscript, while also making clearer how our analyses respond to the central research objectives.

(2) Figure 1: The rebuttal states that the density plots were removed from the borders of the map, this does not appear to be the case.

[Response 2] Thank you for your constructive feedback. You are absolutely right — we overlooked removing the density plots from the borders of the map as stated in the rebuttal. We sincerely apologize for this oversight and appreciate your careful review. We have now corrected this issue, and the revised Figure 1 is provided below.





(3) Lines 378-381: Refers to Figure 4, which are the bi-variate plots of response ratio correlations. This does not seem related to BI in the Himalayan region or Freshwater. Perhaps the authors meant to cite supplementary Figure 4? Or Figure 3?

[Response 3] Thank you for your careful observation. As you correctly pointed out, the text refers to the second-level subgroup analysis, which was originally presented as Supplementary Figure 5. We have now moved this analysis into the main text and updated it as the new Figure 4. Accordingly, we have revised the in-text citation to ensure it accurately refers to the correct figure and analysis (see Lines 416–444).

(4) Figure 5: Which part of the plots are representing the relative contribution? The bar graphs, which appear to be the primary portion of the figures, do not match with the values listed in the donut graphs, and neither match with the values provided in the text. Specifically, the passage on lines 399-400 which reads: “... when considering the overall effects of GCFs, the dimension contributed 17.27% to the total impact...” This does not appear to match any of the plots provided in figure 5. The same goes for the passage on lines 404-405: “Plant life forms had a substantial impact (14.89%) on the effects of CC on plant β -diversity”. The authors should clarify, either in the text or in their figure caption, what is being reported in these plots and how the donut graphs relate to the bar graphs.

[Response 4] We thank the reviewer for the careful reading and helpful comment. As outlined in the Methods section (Lines 181–200), we grouped the 21 individual predictor variables influencing GCF effects into seven broader thematic categories.

In Figure 6 of the revised manuscript, the bar graphs display the relative importance of each individual predictor variable, as determined by the random forest analysis. In contrast, the donut charts present the aggregated contributions of the broader predictor groups, obtained by summing the relative importance values of all variables within respective category.

In response to your observation regarding the apparent inconsistencies between the numerical values in the figure and those reported in the main text (e.g., Lines 399–400 and 404–405 of previous manuscript), we have carefully reviewed the calculations and corrected the numerical values in both the figure and the text. The text and figure have now been fully harmonized to ensure consistency and accuracy.

To improve clarity, we have revised the caption of Figure 6 (now on Lines 486–494 in the revised manuscript) as follow:

“The bar graphs display the relative importance (in %) of the 21 individual predictors based on random forest analysis. The donut charts summarize the contributions of the seven predictor groups, calculated as the percentage of total importance, obtained by summing the relative importance values of all individual variables within each group.”

We hope this revision resolves the confusion and enhances the interpretability of the figure.

(5) Figure 6: It is unclear what is being shown here, and the text connecting to this figure seems unrelated. If the effect is mainly being driven by LUC and LM, where are these factors, and how are they being represented in this model. The authors should consider replacing Figure 6 with their supplementary figure 6, as this helps to clarify their point.

Additionally, authors should consider including an explanation of the statistical assumptions behind the meta-regression within their methods. Then later in the results,

explain if their models adequately met those assumptions. Visually it seems to us that assumptions of error normality and homoscedasticity might not be met.

Along the same lines, the authors should consider and add to the discussion the biological relevance of some of their findings, as opposed to only statistical relevance. For example, Figure 6 panels D, E, and G are all statistically significant but with very shallow slopes. How biologically meaningful is this relationship?

[Response 5] Following the reviewers' suggestion, we checked the statistical assumptions of our meta-regression models. We performed diagnostic analyses including Q-Q plots, Shapiro-Wilk normality tests, and residuals vs. fitted values plots (see Supplementary Figure 5).

The diagnostics indicated that the assumptions of normality and homoscedasticity were not fully satisfied. However, such deviations are common in meta-analytic data due to inherent between-study heterogeneity. To address potential violations, we employed multilevel models and robust variance estimation using cluster-robust standard errors. These approaches are recommended for ensuring the reliability of meta-regression results under imperfect model assumptions (Hedges, 2010; Hua et al., 2024). However, none of the moderators showed statistically significant effects ($p > 0.05$; see Supplementary Figure 5).

(6) Supplementary Figure 5 is mislabeled as Figure 3

[Response 6] Thank you for your careful review and for pointing out this error. We have thoroughly checked all figure and table references throughout the main text and supplementary materials to ensure consistency and have corrected any similar errors. We have also ensured that no similar errors remain.

(7) Minor syntax and grammatical issues

Line 6: “anthropogenic global change” to “anthropogenic global changes” to keep case consistency. Alternatively, authors could change “have” to “has”.

Line 36: “is” to “are”.

Line 37: insert “a” before global biodiversity crisis.

Line 47: insert “the” before “introduction”

Line 61: Consider changing “level” to “scales”

Lines 65-66: Consider changing “at the time” to “at a time”

Line 280: “will” to “was”

Line 525: Consider changing “lacks” to “drawbacks”

[Response 7] We sincerely thank the reviewer for the close reading and helpful suggestions regarding grammatical and syntactic refinements. We greatly appreciate your attention to detail. In response, we have implemented all suggested edits as follows:

- Line 6: “Anthropogenic global change” has been revised to “anthropogenic global changes” to ensure case consistency (see Line 8 of the revised manuscript).
- Line 36: The verb “is” has been corrected to “are” to agree with the plural subject (see Line 44).
- Line 37: We have inserted “a” before “global biodiversity crisis” for grammatical accuracy (see Line 45).
- Line 47: “The” has been added before “introduction” (see Line 56).
- Line 61: This sentence has been fully revised for clarity and precision; see updated wording in Lines 83–84.
- Lines 65–66: “At the time” has been changed to “at a time” (see Line 89).
- Line 280: The modal verb “will” have been corrected to “was” to reflect past tense (see Line 314).
- Line 525: This sentence has also been fully revised; “lacks” has been replaced with “drawbacks” as suggested, and the revised sentence appears in Lines 609–612.

All changes have been carefully made in the corresponding locations. We are grateful again for your thoughtful and constructive input, which helped improve the manuscript’s clarity and readability.

For Reviewer #3 and Reviewer #4:

Revisions corresponding to comments that were confirmed as fully addressed in the previous round are not repeated in this version. Only the remaining issues that were considered unresolved are included in the current response.

Reviewer #3 & #4 (Remarks to the Author):

SUMMARY

We appreciate the efforts of the authors to improve the manuscript in response to our comments. However, some issues were marked as addressed, yet the corresponding changes were not fully reflected in the text. Additionally, it was difficult to verify where our comments had been incorporated, as the revised sections were not clearly indicated. Moving forward, we kindly ask the authors to specify the exact locations of their revisions by providing line numbers.

(8) Furthermore, we strongly request that the authors complete the PRISMA-EcoEvo checklist (<https://www.prisma-statement.org/ecoevo>) to ensure transparency in their study. Please fill in the checklist with page and line number so we could check them carefully for comprehensive reporting.

[Response 8– Second Round] In response to the reviewer's suggestion, we have completed the PRISMA-EcoEvo checklist and provided it as a supplementary file titled *PRISMA-EcoEvo_ExcelChecklist.xlsx*. To ensure transparency and facilitate comprehensive review, we have included the corresponding line numbers for each item in the checklist.

Since each document (revised manuscript and supplementary files) uses a continuous and unique line numbering system, we have not included page numbers separately. However, as some information is located in the supplementary files, we have added an additional column to indicate the file name for each item, making it easier to locate the relevant content.

Below, we provide our responses to the authors' comments

COMMENTS

L19: What does mean “tree- β diversity”? Please explain.

[Response 12] Thank you very much for your valuable comments. In our study, we classified plants into categories based on life forms, including aquatic plants, herbaceous plants, shrubs, trees, woody plants, combinations of woody and herbaceous plants, mosses or lichens, and so on. Subgroup analysis results showed that GCFs significantly reduced the β -diversity of trees (Figure 3). We acknowledge that the original wording may have caused confusion, and we have made corresponding revisions throughout the manuscript to clarify such issues.

(9) Could the authors clarify what specific revisions were made throughout the manuscript to address this issue?

[Response 9– Second Round] We sincerely apologize for not specifying the exact revisions in our previous response. The term “tree β -diversity” was originally intended to mean “ β -diversity of trees.” To improve clarity, we revised this phrasing to “ β -diversity of trees” in the first revised version of the manuscript. In the current version, we further enhanced the clarity and logical consistency of our findings by referring to “tree communities” in the Abstract (see Line 25), which more accurately reflects the scope and structure of the analysis.

L64-65: It is unclear what "all three dimensions" refer to at this point (though it becomes apparent later that they are functional, phylogenetic, and taxonomic plant β -diversity). However, we noticed that no explanation or definition of these three different types of plant β -diversity was provided in the whole manuscript. We would suggest including this in the Introduction.

[Response 13] Thank you for your valuable comment. We appreciate the suggestion to clarify the reference to "all three dimensions." We have included a clearer definition in

the abstract, where we specify that the three dimensions of biodiversity refer to taxonomic, functional, and phylogenetic diversity. We will also ensure that these terms are further defined and discussed in the Introduction to provide better clarity for readers throughout the manuscript.

(10) The authors state that these terms have been further defined in the Introduction to enhance clarity. However, we were unable to locate a clear explanation. If this revision has been made, please specify the relevant sentences.

[Response 10 – Second Round] We sincerely apologize for the oversight in our previous revision. You are absolutely right that, while we initially stated the three dimensions of β -diversity in the Introduction, we did not provide explicit definitions for them. In response to your helpful suggestion, we have now added a clear explanation of these three dimensions of plant β -diversity—taxonomic, functional, and phylogenetic—to the Introduction section to enhance clarity for readers. Specifically, the following sentence has been added on Lines 84–87 of the revised manuscript:

"Moreover, β -diversity extends beyond taxonomic turnover (i.e., species replacement) to include phylogenetic and functional dimensions, which reflect differences in evolutionary histories and ecological strategies across communities."

We hope this revision fully addresses your concern, and we are grateful for your careful and constructive feedback.

We could not find the information on metadata, raw data, and scripts used for analyses from this manuscript. We strongly suggest they provide all the information necessary to replicate this study (at the review stage so we can assess them). Sharing them will enhance transparency, allowing others to validate or build on the findings. If there is a reason why they could not do it, the authors should clearly state that. - [18.1-4]

[Response 14] Thank you very much for your suggestion. We have uploaded the raw data into the submission system as Supplementary Material.

(11) While the authors have provided the dataset and R scripts, the corresponding metadata explanation is still missing. Providing this information is important for

ensuring that future researchers can properly interpret the data.

[Response 11 – Second Round] Thank you for pointing this out. In response to your suggestion, we have now provided a description for database file describing the structure and content of database, including variable definitions, units, data sources, and any necessary notes for interpretation. This description file has been uploaded as Supplementary file alongside database and R code. We hope this will help ensure full reproducibility and facilitate future use of our dataset.

The authors should describe what sources of information were sought. For example, was the search limited to only published studies or included grey literature (e.g. personal communications, theses, or unpublished studies)? - [5.1]

L97-99: We would like them to provide the exact search strings used for each database. When we entered the search strings provided by the authors into the search engines of Web of Science, Google Scholar, and China Knowledge Resource, the total number of results was significantly lower than the number of papers reported by the authors in the manuscript. Also, we would like to know whether the authors set any temporal restrictions on the database searches. - [5.3-4]

L99: “35206 unique studies” is not a number that includes duplicate studies.

[Response 15] Thank you for your valuable feedback. We have updated the manuscript to provide more details about the sources of information sought. The literature search was conducted in Web of Science, Google Scholar, and the China Knowledge Resource Integrated Databases between 29 March and 31 August 2023. The search in Web of Science and Google Scholar used the following keyword combinations: (“beta diversity” OR “ β diversity”) and (“biotic homogen*” OR “biotic differentiat*” OR “biotic heterogen*” OR “biological homogen*” OR “biological differentiat*” OR “biological heterogen*”). In the China Knowledge Resource Integrated Databases, we translated these keywords into Chinese for the search. A total of 35,206 results were retrieved across all three search engines (see Supplementary Fig. 1), including duplicates. We did not apply any temporal restrictions during the search. Additionally, we examined the

reference lists of relevant review articles, as well as subsequent studies citing these reviews (e.g., Catano et al. 2017, Kramer et al. 2022, McKinney 2006, Petsch et al. 2022, Rolls et al. 2023, Soininen et al. 2018) (229 studies). The search was limited to published studies.

(12) The authors have not provided the exact search strings used in each database. By “exact search string,” we refer to the precise query used, which should return the same or slightly more results as reported in the manuscript when entered directly into the database. We attempted to replicate the search using the provided search string but encountered the following issues:

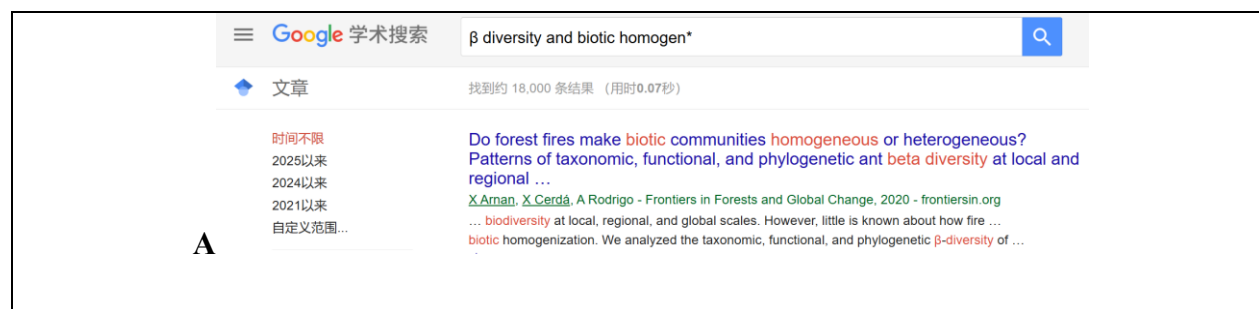
Web of Science: the search resulted in an error message - “Search Error: invalid field tag.”

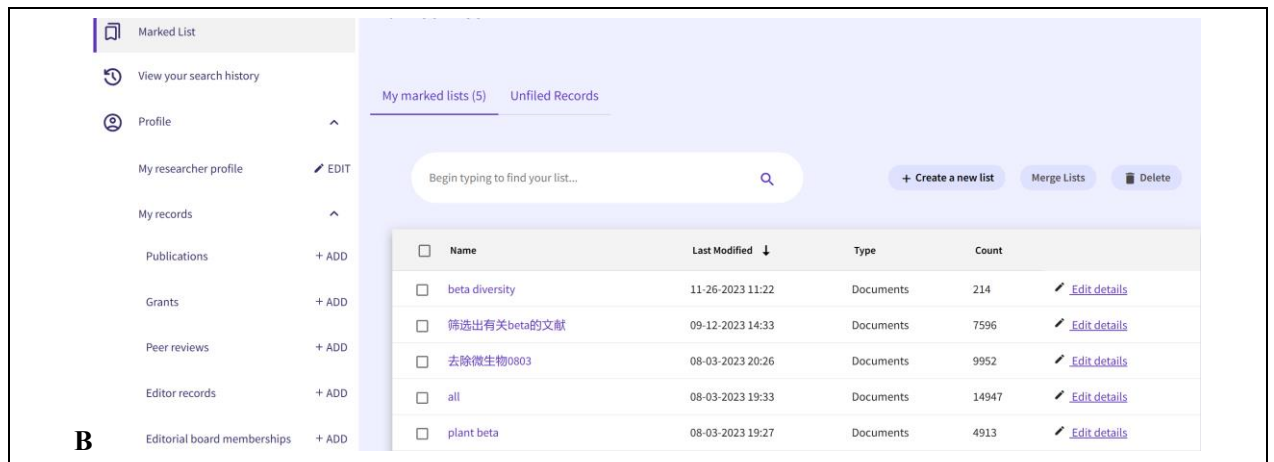
Google Scholar: the search returned only 3 results, which seems inconsistent with the reported results.

Could the authors clarify the exact search string used for each database, including any field tags, Boolean operators, and filters applied? This would help ensure the reproducibility of the literature search.

[Response 12 – Second Round] Thank you very much for your careful review and comments. We have carefully examined the issues raised and believe that some misunderstandings may have occurred regarding how we presented our search strategy.

For example, when we searched for β diversity and biological homogen* in Google Scholar without quotation marks, the search returned approximately 18,000 results (as shown in Screenshot A, which we have now provided). To further clarify, we have also included Screenshot B, which displays a portion of the records retrieved from Web of Science during our searches conducted in 2023.





Screenshot A Search results in Google Scholar using the query *β diversity and biological homogen** without quotation marks, conducted by the authors on April 9, 2025. **Screenshot B** A sample of the records retrieved by the authors from Web of Science during the literature search in 2023.

The confusion likely stems from the fact that in our original manuscript, we enclosed the keyword combinations within quotation marks (e.g., “beta diversity” AND “biological homogen*”). We now realize that including these quotation marks may have misled readers to enter them directly into the search boxes, which can result in very limited or even erroneous search results (e.g., only 2–3 records in Google Scholar). In response to this, we have removed quotation marks from the search strings in the revised manuscript to avoid any further confusion.

To reiterate, in the China Knowledge Resource Integrated Databases, we used translated Chinese terms that correspond to the English keywords used in Web of Science and Google Scholar. The searches were all conducted manually, with different keyword combinations tested and recorded systematically.

We thank the reviewers once again for their attention to detail and their commitment to ensuring the reproducibility of our methodology.

L147-153: We would like to provide more specific and detailed explanations for all the variables included in the six categories. For example, what is meant by 'plant functional

type,' and what categories are included? Additionally, were any variables among they extracted either not used in the analysis or excluded due to insufficient information? - [8.1-4]

[Response 19] We sincerely thank the reviewers for their insightful suggestions. We have revised the manuscript to address both the request for more specific explanations of the variables and the suggestion regarding the quality assessment of primary studies.

1) In response to your request for more specific details about the variables included in the seven categories, we have provided clearer definitions in the manuscript. We also acknowledge that the term "plant functional type" was not used correctly and have replaced it with "plant life forms," which is a more appropriate description. The seven categories and their specific variables are as follows:

Five GCFs: biological invasion (BI), climate change (CC), land use change (LUC), land management (LM), and multiple factors (MF).

β -diversity dimension: Three dimensions of β -diversity are considered: taxonomic β -diversity (TB), functional β -diversity (FB), and phylogenetic β -diversity (PB).

Geographic location: The geographic variables include latitude, longitude, region, realm, elevation, and whether the study site falls within a biodiversity hotspot.

Ecological setting: This includes climatic zone, ecosystem type, flora realm, mean annual precipitation (MAP, in mm), and mean annual temperature (MAT, in °C).

Study design: This includes information on whether the study focused on temporal or spatial design, sample size, control sample type, temporal extent, and spatial extent.

Data characteristics: data sources and whether the data were abundance-weighted.

Research subject: This includes the examined plant taxa and life forms (instead of "plant functional types").

Each of the aforementioned data points was sourced either directly from the study reports themselves or by utilizing the latitude and longitude of the study sites. In cases where a study did not report the latitude and longitude, we estimated approximate coordinates by geocoding the site name using Google Earth 7.0 and then obtained the elevation based on these coordinates. Similarly, if the study did not provide mean annual temperature (MAT) and mean annual precipitation (MAP), we retrieved these

values from WorldClim using the geographic location (latitude and longitude) of the study site.

For studies that covered large areas (such as entire countries or the globe) without specific coordinates, we did not assign environmental metrics to them, which represented 6% of the total studies. It's worth noting that during the specific analysis, if a study did not involve a certain type of data (such as sample size), it would not be included in the effect analysis related to that data.

2) In response to the reviewer's comment, we have taken steps to reduce potential biases in our meta-analysis. Specifically, we have only included studies that have undergone peer review and are publicly published, ensuring a level of quality in the primary data. Additionally, based on the reviewer's suggestion and the Prisma-EcoEvo guidelines (O'Dea et al., 2021), we have excluded studies with a sample size of fewer than 3 from our analysis. Consequently, our meta-analysis is now based on 1604 observations from 256 studies, which helps to ensure the reliability and validity of our results.

(13) Thank you for your response. However, we believe the following points have not yet been fully addressed:

How was the quality of individual studies assessed? (e.g., blinding of data collection, reporting quality, differences in study design)

Was the assessment of study quality incorporated into the meta-analysis? (e.g., meta-regression, sensitivity analysis)

A summary of study quality characteristics (e.g., risk of bias assessment)

The following paper may also be helpful as a reference:

Feeley TH. Assessing study quality in meta-analysis. *Human Communication Research*. 2020 Apr;46(2-3):334-42. <https://doi.org/10.1093/hcr/hqaa001>

Hohn RE, Slaney KL, Tafreshi D. Primary study quality in psychological meta-analyses: An empirical assessment of recent practice. *Frontiers in Psychology*. 2019 Jan 9;9:2667. <https://doi.org/10.3389/fpsyg.2018.02667>

Shaheen N, Shaheen A, Ramadan A, Hefnawy MT, Ramadan A, Ibrahim IA, Hassanein ME, Ashour ME, Flouty O. Appraising systematic reviews: a comprehensive guide to ensuring validity and reliability. *Frontiers in research metrics and analytics*. 2023 Dec

21;8:1268045. <https://doi.org/10.3389/frma.2023.1268045>

[Response 13 – Second Round] Thank you very much for your thoughtful comments. We acknowledge that our initial response did not fully address the important points you raised regarding the appraisal of individual study quality.

To address this, we have added a new section (2.3 Critical Appraisal) in the Methods part of the revised manuscript, where we describe our approach to assessing the quality of the included studies. Specifically, we evaluated study quality based on a set of criteria adapted from Mupepele et al. (2016), which are designed for assessing evidence strength in studies on ecosystem services and biodiversity. In determining the final criteria used in this study, we also referred to the methods applied by Huynh et al. (2024) and Su et al. (2021).

Each study was scored based on whether it met individual quality criteria, and the overall scores were used to classify studies into four levels of evidence: weak, moderate, strong, and very strong. A detailed breakdown of quality scores for each study is now provided in a newly added Supplementary file (Study quality assessment.xlsx), and a summary of the assessment criteria and scoring system is included in Supplementary Table 1.

We hope that these revisions improve the transparency of our methodology and help to strengthen the validity and reliability of our findings. Please let us know if any further clarification is needed—we would be happy to revise further in accordance with your suggestions.

L187: equation (1) - parentheses for ratio will be needed.

L187-188: These are not the best estimators - see this paper

Lajeunesse MJ. Bias and correction for the log response ratio in ecological meta-analysis. *Ecology*. 2015 Aug;96(8):2056-63.

[Response 20] Thank you for your valuable feedback. In our analysis, we used the log response ratio (ln RR) as an effect size measure, and we are aware of the potential biases associated with this method, as highlighted in Lajeunesse (2015). To address these

concerns, we estimated effect sizes using the Delta method recommended by Lajeunesse (2015), which provides a more reliable and bias-correcting method of calculating $\ln RR$, particularly in ecological meta-analyses. In response to your suggestion regarding the equation, we have revised it to include parentheses for the ratio to improve clarity. The updated and additional equations are as follows...

L206: We do not think variance can be calculated using equation (3). Provide an equation for sampling variance.

[Response 21] Thank you for your comment. We have added the formula for calculating the variance. The variances (v) were calculated using the following equation...where n_t and n_c are the sample sizes of the variable in the treatment and control, respectively; s_t and s_c are the standard deviations of the variable in the treatment and control, respectively.

(14) Thank you for addressing our comments. However, it would be helpful to provide a clearer justification for using this new method. The statement “To reduce bias in the effects...” (L196) does not seem to be a sufficient explanation.

[Response 14 – Second Round] Thank you very much for your helpful suggestion. We have revised the manuscript to provide a clearer justification for using the Delta method. Specifically, we have added the following sentence (see Lines 221-227):

“To minimize bias in effect size estimation and variance calculation, we adopted the multivariate Delta method proposed by Lajeunesse (2015), which addresses the small-sample bias associated with conventional log response ratio ($\ln RR$) estimators. Specifically, we used a second-order Taylor series expansion to adjust both the mean (Equation 3) and sampling variance (Equation 4) of $\ln RR$, enhancing the robustness and accuracy of effect size estimates, particularly under the small-sample conditions common in ecological research.”

This clarification highlights the statistical rationale behind our method choice and reflects our intention to follow best practices in ecological meta-analysis, particularly under small-sample conditions where conventional estimators are prone to bias.

The authors ticked the item about Bayesian analysis in the Nature portfolio reporting summary but did not provide information on the choice of priors or Markov Chain Monte Carlo setting - they should have given the information about this in the Results section. Please also provide all packages, functions, and their versions used to run the models. - [13.2-5]

[Response 22] Thanks for your attention. This study indeed does not include the Bayesian analysis. We have revised this choice in the Nature portfolio reporting summary.

(15) The authors have not yet reported the specific functions and package versions used in their analyses. If possible, please include this information in the manuscript.

[Response 15– Second Round] Thank you for pointing this out. In response, we have added the names of the specific functions and the corresponding R package versions used in our statistical analyses. These details have been incorporated into the revised manuscript and can be found at Lines 271, 316, 322, 323, and 328.

L215-232: We guess the authors used multilevel meta-analyses from their descriptions. It would be great to see two different levels of heterogeneity (e.g. I² at the effect size level and I² at the study level).

[Response 23] Thank you for your suggestion. We have included the I² statistic in the two-level subgroup analyses in the revised manuscript.

L218: Did the authors consider cohort ID (groups of the same subjects), shared control ID (indicating effect sizes sharing control groups), as well as study ID? - [14.1-3]

Note that the related points were recently explained in this paper with detailed tutorials - see Nakagawa S, Yang Y, Macartney EL, Spake R, Lagisz M. Quantitative evidence synthesis: a practical guide on meta-analysis, meta-regression, and publication bias tests for environmental sciences. *Environmental Evidence*. 2023 Apr 24;12(1):8.

[Response 24] Thank you for your insightful suggestion. In our previous manuscript, we considered the influence of the study ID but did not fully account for the non-independence of data arising from shared control groups within the same study. In

response to your comment, we have revised our analysis accordingly. Specifically, we now incorporate both the study ID and control ID as random effects in our mixed-effects models. The control ID is nested within the study ID, meaning that each study has its own one or several sets of control groups, and the effect sizes within a study are potentially correlated due to this structure.

We have also referred to the detailed guidance provided in Nakagawa et al. (2023) to implement these adjustments in the context of environmental science meta-analysis.

Below is the updated model:

$$\text{Effect Size}_{i,j} = \beta_0 + (1|\text{Study ID}) + (1|\text{Study ID} : \text{Control ID}) + \epsilon_{i,j}$$

Where $\text{Effect Size}_{i,j}$ represents the effect size for the i -th observation within the j -th study. β_0 is the fixed effect (the total effect of GCF on plant beta diversity). $(1|\text{Study ID})$ accounts for variation between studies. $(1|\text{Study ID} : \text{Control ID})$ accounts for variation due to control groups nested within each study. $\epsilon_{i,j}$ is the residual error term.

(16) We appreciate the authors' revisions. One minor suggestion is to briefly explain the meaning of I^2 and Q_i .

[Response 16 – Second Round] Thank you for your valuable suggestion. In response, we have added a brief explanation of the I^2 and Q_i statistics in the revised manuscript (see Lines 286–291). Specifically, we now clarify that the I^2 statistic describes the proportion of total variation in observed effect sizes that is due to true heterogeneity rather than sampling error, while Q_i (the total Q statistic) is used to assess whether the observed variability in effect sizes is greater than would be expected by chance alone. These additions aim to improve the clarity and interpretability of our heterogeneity analysis for readers.

L237: The authors used the fail-safe number, but it is heavily criticized. Consult a recent review of publication bias.

Nakagawa S, Lagisz M, Jennions MD, Koricheva J, Noble DW, Parker TH, Sánchez-Tójar A, Yang Y, O'Dea RE. Methods for testing publication bias in ecological and evolutionary meta-analyses. *Methods in Ecology and Evolution*. 2022 Jan;13(1):4-21.

[Response 25] Thank you for your suggestion. Considering the concerns regarding the reliability of the fail-safe number, we have removed the results based on this metric from the manuscript. Instead, we focused on more robust methods for assessing publication bias, such as the Egger regression test and the trim-and-fill method, as recommended by recent literature (Nakagawa et al., 2022). We believe these methods provide more reliable estimates of the potential effects of publication bias and offer a better evaluation of the robustness of our results.

(17) L264-273: Could the authors clarify what is meant by “marginal publication bias”? Additionally, the relevant results should be reported in the Results section.

[Response 17 – Second Round] Thank you for your helpful comment. We apologize for the confusion caused by our wording. The term “marginal publication bias” was indeed misleading—our intention was to indicate that only a small or limited degree of publication bias was detected in some cases, based on Egger’s regression test. To improve clarity, we have revised the sentence in the Methods section accordingly (see Line 338). In addition, as suggested, we have reported the corresponding statistical results—including p-values from Egger’s tests—in the Results section (see Lines 410–415 and Supplementary Tables 2 and 3).

Provide the number of effect sizes for the data included in meta-analyses (does the number of tests equal effect size?). - [20.1]

State a summary of key characteristics for reported outcomes. In particular, the authors did not provide any summary of the study design and research subject. - [20.3]

Report the results section based on statistical analysis; do not simply state that there was a significant positive or negative effect without providing statistical details or reporting only the p-value. Always describe the full statistical results along with the findings. - [21.1, 23.1-4]

[Response 27] Thank you for your valuable feedback. We have carefully revised the manuscript to address your concerns. We have now included the number of effect sizes

for the data included in the meta-analyses and clarified that the number of tests corresponds to the number of effect sizes used in the analyses. In the results section, we have ensured that all statistical analyses are presented in full, including effect sizes, confidence intervals, and relevant p values. We have avoided simply stating whether the effects were positive or negative and have provided a more comprehensive description of the statistical results. We hope these revisions address your concerns and improve the clarity of our manuscript. Thank you again for your helpful suggestions.

(18) L311-326: Could the authors specify where in Figure 3A we should look? Furthermore, it would be helpful to provide the values before converting them to percentages, even in areas where percentages are shown.

[Response 18 – Second Round] We sincerely thank the reviewer for this helpful suggestion.

In response, we have added the original lnRR values along with their corresponding 95% confidence intervals throughout the relevant parts of the text to enhance transparency.

Additionally, we have clarified the specific subpanels and categories in Figure 3A where each result is displayed, enabling readers to easily locate the relevant data. These revisions are reflected in Lines 386–394 of the revised manuscript. To further improve clarity, we have also added the formula used for converting lnRR values to percentage change in the Methods section (see Lines 259–260).

L274: Please explain the meta-regression analysis in the methods section (e.g. how many and what moderators are used? How select the final model?) - [15.1-3; 20.2,4]

[Response 30] Thank you for your insightful comment. We have made the necessary revisions to the manuscript to clarify the meta-regression analysis, as follows:

First, we conducted a meta-regression analysis to examine the relationship between the effect sizes (ln RR) and the following moderators: latitude, longitude, elevation, mean annual temperature (MAT), mean annual precipitation (MAP), spatial extent, temporal extent.

We performed the analysis using the metafor package in R (Viechtbauer, 2010). The analysis was weighted by the variance of each observation and utilized the Restricted Maximum Likelihood (REML) method to estimate the model parameters. Regarding missing data, the temporal extent variable had over 50% missing values, so it was removed from the analysis. For the remaining variables with less than 50% missing data, missing values were imputed using the Random Forest algorithm with default settings, implemented in the missForest package (Stekhoven et al., 2012).

We hope this clarifies the approach taken in the meta-regression analysis. Thank you again for your helpful feedback.

(19) Were there any correlations among these moderators? Additionally, how was the validation of the imputed data evaluated?

[Response 19 – Second Round] Thank you for your thoughtful follow-up question. We are happy to provide further clarification regarding the treatment of moderators, missing data, and collinearity.

First, regarding the meta-regression analysis, we emphasize that imputed data were not used. Because each moderator was analyzed independently in separate univariate models, only observations with non-missing values for the specific moderator were included. This approach avoids potential biases introduced by imputation and eliminates the need to account for multicollinearity among moderators.

Second, missing data imputation was applied only in the context of the random forest analysis, which required a complete dataset including all moderators simultaneously. We used the missForest package in R, a non-parametric imputation method, and assessed imputation quality using its internal validation metrics: the normalized root mean squared error (NRMSE) for continuous variables and the proportion of falsely classified entries (PFC) for categorical variables. The imputation accuracy was satisfactory (NRMSE < 0.2; PFC < 0.1).

Third, to assess collinearity among moderators for the random forest analysis, we evaluated pairwise relationships among both continuous and categorical variables. No pair of continuous moderators exhibited a Pearson correlation coefficient greater than 0.7. However, among categorical moderators, we observed high associations: Cramér's

$V > 0.7$ between growth form and taxa type, between growth form and realm, as well as between realm and ecosystem. Furthermore, region showed moderate associations with latitude and longitude ($\eta^2 > 0.3$), while climate zone showed moderate associations with MAT ($\eta^2 > 0.3$). Despite these associations, we retained all moderators in the random forest model for two reasons. First, random forest is a tree-based, non-parametric algorithm that is relatively robust to multicollinearity and does not rely on assumptions of linear independence among predictors. Second, excluding variables based on pairwise associations may result in the loss of ecologically meaningful information, especially when correlated variables may still provide distinct contributions to the overall predictive performance. Retaining all moderators allowed the algorithm to assess their relative importance in a multivariate ecological context.

We hope this explanation addresses your concerns. Thank you again for your constructive feedback, which has helped improve the clarity and transparency of our manuscript.

L276: What is this percentage value? The back-transformed values of $\ln RR$?

[Response 31] Thank you for your valuable comment. To clarify, the 12% reduction in the β -diversity of tree initially reported refers to the back-transformed values of the natural log response ratio ($\ln RR$). This represents the effect size in the original scale, after converting the $\ln RR$ estimates into percentage values. Additionally, we would like to correct this figure: based on our reanalysis, GCFs actually reduced the β -diversity of trees by 15% ($p < 0.05$). We will revise the manuscript to reflect this updated result and provide a clearer explanation of the statistical approach.

(20) The authors have not provided the explanation in their manuscript.

[Response 20 – Second Round] Thank you very much for your helpful suggestion. To avoid potential confusion for readers, we have revised this section to add the original $\ln RR$ values along with their corresponding 95% confidence intervals (see Lines 365-385). We believe this improves clarity and ensures consistency in the presentation of our results.

L301-306: Please explain the analyses regarding this paragraph's results in the Methods section.

[Response 32] Thank you for your constructive comment. In response, we have revised the Methods section to provide a more detailed explanation of the statistical analyses used for the results presented in this paragraph. The updated text is as follows: “Additionally, to investigate the consistency of the three dimensions of plant β -diversity (taxonomic, functional, and phylogenetic β -diversity) in response to GCFs, we conducted correlation analyses on the $\ln RR$ values of these dimensions using bivariate meta-analysis. ”

(21) L344: What does binary meta-analysis mean in this context? The authors mention conducting bivariate (or binary?) meta-analyses, but after reviewing the R codes (file: “9.Figure 6-bivariate meta-analysis.R”), we could not determine how such meta-analyses were performed. Could the authors clarify this point?

Also...

L378-381/ L417-422: We may have overlooked something, but we are unsure which parts of the figures correspond to these findings. Could the authors clarify?

[Response 21 – Second Round] Thank you very much for your insightful questions.

First, we apologize for the confusion caused by the terminology. The correct term should be "bivariate meta-analysis" rather than "binary meta-analysis." We have now revised the manuscript text at Line 445 to reflect this correction. The bivariate meta-analysis refers to the modeling of effect sizes for two β -diversity dimensions simultaneously, allowing us to estimate their correlation while accounting for within-study dependencies and between-dimension correlations.

In the revised analysis (as described in [Response 25 – Second Round]), we used the `rma.mv` function from the `metafor` R package with an unstructured variance-covariance matrix (`struct = "UN"`) for the two dimensions, and included nested random effects (`random = list(~ Dimension | Study_ID, ~ 1 | Obser_ID)`). This approach enabled us to estimate the correlation between, for example, taxonomic and phylogenetic β -diversity changes based on studies that reported both metrics. We have now clarified

this in both the main text and the figure legends.

Second, regarding your comment on Lines 378–381 and 417–422 in previous revised manuscript: we appreciate your careful reading and agree that the correspondence between the findings and the figures was not sufficiently clear. In the revised manuscript, we have updated the relevant figure captions and the Results section (see Lines 445–455) to clearly indicate that these correlation estimates are presented in Figure 4.

Thank you again for helping us improve the clarity and rigor of our work.

- Did they consider limitations and their influence on the generality of their conclusions?
- Although we do not know the number of data points for each variable, should not limitations such as gaps in the available evidence (e.g. geographical research biases or researched plant type bias) be considered? - [25.6]
- Please refer to the comments in the Result section as well.

[Response 33] Thank you for your insightful comments. We have added a discussion on the limitations of the study towards the end of our revised manuscript, as follows: “However, we also note that well-known biases, such as greater sampling effort in the global north compared to the Southern Hemisphere (especially in Australia), persist in the biodiversity data compiled for this study (Figure 1). This significantly limits the generalizability of our findings to well-studied regions of the planet.”

(22) The authors addressed our comments on the manuscript. However, we were wondering whether there are any citations here.

[Response 22 – Second Round] Thank you very much for your helpful suggestion. We initially identified a relevant citation (Chambers et al., 2017) to support the statement regarding geographic biases in biodiversity data, particularly the underrepresentation of studies in Australia. However, because this bias is already clearly illustrated in our own dataset (Figure 1), we revised the sentence to refer directly to the figure rather than include an external citation (see Lines 613–625).

Please provide a reference list of all eligible studies for the meta-analysis. Also, please list included studies as referenced sources. - [27.1-2]

[Response 34] Thank you for your helpful suggestion. We have now included the reference list of all eligible studies for the meta-analysis in Supplementary Material 2, which contains the 256 studies that were included in this research.

We hope this addition meets your expectations. Thank you again for your valuable feedback.

(23) Listing only the titles of referenced papers is insufficient. Please include the authors' names, DOIs, publication years, and journal names. Additionally, could the authors provide a full list of articles excluded at the full-text screening stage, along with the reasons for exclusion? This would allow third parties to refer to the information more easily.

[Response 23 – Second Round] Thank you very much for your valuable suggestion. In response, we have added a supplementary file entitled "**Reference lists.xlsx**", which includes the following two sheet:

- 1) The first sheet provides the complete list of the 256 studies included in our meta-analysis, with full citation details, including authors' names, publication years, journal names, and DOIs.
- 2) The second sheet contains a comprehensive list of studies excluded during the full-text screening stage, along with the specific reasons for exclusion.

We believe this addition significantly improves the transparency and reproducibility of our study. Thank you again for your thoughtful feedback.

Figure 1: Figure 1 was quite challenging to understand. First, we would like clearer separations between sections A and C. Second, in A, please indicate what the insets' X-axis (top) and Y-axis (right) represent. The circles on the world map overlap too much and are hard to interpret. For example, what does this overlap mean? Third, in B, what does each plot represent? (we assumed it was the number of tests, but if so, should not taxonomic β -diversity have the most?) Finally, in C, this might be because it is outside

our area of expertise, but in what situation is elevation over 3000 m in freshwater ecosystems?

[Response 35] Thank you for your valuable feedback on Figure 1. We have made several revisions to improve clarity and address your concerns:

4) Interpretation of Plots in Section B: Section B displays the distribution of diversity dimensions with respect to mean annual temperature (MAT) and mean annual precipitation (MAP). We have adjusted the sizes of the circles for each dimension to prevent overlap and enhance visibility. Additionally, we reviewed the representation of taxonomic β -diversity to ensure accuracy, as it is expected to have the largest number of studies. This has been corrected in the figure.

(24) It would be helpful to include an explanation of the circle sizes in the figure legend.

[Response 24 – Second Round] Thank you very much for your valuable comments. As the three dimensions of β -diversity are already differentiated by color in Figure 1A, B, and C, we used different circle sizes only in Figure 1B to reduce visual overlap and improve readability. In response to your suggestion, we have added an explanation of the circle sizes to the figure legend of Figure 1B to enhance clarity. Specifically, circle sizes were manually adjusted to represent the three dimensions of β -diversity: Taxonomic (small), Functional (medium), and Phylogenetic (large). Please refer to the updated figure legend on Lines 149–151 of the revised manuscript.

Figure 4: We could not find the method related to this. One cannot just use the correlation to analyse the relationships between 2 sets of effect sizes (e.g. pseudo-replication and also there are two correlations that can be calculated at the between- and within-study levels). One should use bivariate meta-analysis. See this link (but your model will be more complex): <https://www.metafor-project.org/doku.php/analyses:vanhouwelingen2002>

[Response 37] Thank you for your valuable feedback. You are correct that simply using correlation analysis may not be appropriate for analyzing the relationships between two sets of effect sizes, as it could lead to pseudo-replication issues. In response to your

comment, we have updated the methods section of the manuscript to include the following clarification:

"Additionally, to investigate the consistency of the three dimensions of plant β -diversity (taxonomic, functional, and phylogenetic β -diversity) in response to GCFs, we conducted Pearson correlation analyses on the $\ln RR$ values of these dimensions. This analysis was performed only for studies that included data from at least two dimensions of β -diversity."

As you suggested, we acknowledge that bivariate meta-analysis would be more appropriate for this type of analysis. However, due to the limitations of the available data, we opted to conduct the correlation analysis and present the results graphically. We used this approach to display the consistency of the direction of change in β -diversity dimensions due to GCFs, based on data where at least two dimensions were reported.

We have updated the figure legend and manuscript accordingly to reflect these changes. (25) The legend of Figure 4 and the main text state that a bivariate meta-analysis was conducted. However, the authors mention that Pearson correlation analysis was used instead due to data limitations. Could the authors clarify whether a bivariate meta-analysis was performed? If so, please provide the corresponding results.

[Response 25 – Second Round] Thank you for your valuable comment. In our previous revision, we used a multivariate meta-analysis model (`rma.mv` in the `metafor` R package) with an unstructured variance-covariance matrix (`struct = "UN"`) to estimate the correlations among the three β -diversity dimensions (taxonomic, functional, and phylogenetic). This model is flexible in handling missing dimensions and does not require each study to report all three dimensions diversity metrics simultaneously. The r -values presented in the previous version of the figure were derived from this model.

However, based on your helpful suggestion, we further considered the alignment between the analysis and the scatterplot data, and realized that it would be more appropriate to restrict the analysis to studies that reported effect sizes ($\ln RRs$ and their variances) for at least two of the three β -diversity dimensions. This ensures that the estimated correlations are based on paired observations, thereby more accurately

reflecting the structure of the data presented in the figure.

Therefore, we re-ran the analysis using a bivariate meta-analysis approach via a multilevel mixed-effects model (`rma.mv` in `metafor`). We again specified an unstructured variance-covariance matrix (`struct = "UN"`) for the two effect size dimensions and included nested random effects (`random = list(~ Dimension | Study_ID, ~ 1 | Obser_ID)`) to account for both within-study dependency and cross-dimensional correlations. This structure allowed us to accurately estimate the correlation between taxonomic and phylogenetic diversity changes while accounting for non-independence of effect sizes within studies.

The updated results are now presented in Figure 4 and Lines 445-455 in the revised manuscript, and the full analysis code is provided in the revised supplementary file (*R code for Figure 5-Bivariate meta-analysis.R*).

References:

- Chambers, L. E., Barnard, P., Poloczanska, E. S., Hobday, A. J., Keatley, M. R., Allsopp, N., and Underhill, L. G. (2017) Southern Hemisphere biodiversity and global change: Data gaps and strategies. *Austral Ecology*, 42: 20–30.
- Hedges, L.V., Tipton, E. and Johnson, M.C. (2010), Robust variance estimation in meta-regression with dependent effect size estimates. *Res. Synth. Method*, 1: 39-65. <https://doi.org/10.1002/jrsm.5>
- Huynh, L.T.M., Su, J., Wang, Q. et al. Meta-analysis indicates better climate adaptation and mitigation performance of hybrid engineering-natural coastal defence measures. *Nat Commun* 15, 2870 (2024). <https://doi.org/10.1038/s41467-024-46970-w>
- Hua, F., Wang, W., Nakagawa, S. et al. Ecological filtering shapes the impacts of agricultural deforestation on biodiversity. *Nat Ecol Evol* 8, 251–266 (2024). <https://doi.org/10.1038/s41559-023-02280-w>
- Mupepele, A.-C., Walsh, J.C., Sutherland, W.J. and Dormann, C.F. (2016), An evidence assessment tool for ecosystem services and conservation studies. *Ecol Appl*, 26: 1295-1301. <https://doi.org/10.1890/15-0595>
- Su, J., Friess, D.A. & Gasparatos, A. A meta-analysis of the ecological and economic outcomes of mangrove restoration. *Nat Commun* 12, 5050 (2021). <https://doi.org/10.1038/s41467-021-25349-1>

Reviewer #1 (Remarks to the Author):

(1) We appreciate the author's work in revising this manuscript. While the majority of the changes are acceptable, the figures are all very difficult to read. Figures should all be revised to increase text size for readability when printed on a single page.

[Response 1] We appreciate the reviewer's constructive feedback regarding figure readability. In response, we have revised all figures in the revised manuscript and supplementary information by increasing the text sizes and, where appropriate, using bold text to improve clarity. We have ensured that all text is legible when printed on a single page. Thank you for pointing this out.

Reviewer #3 & #4 (Remarks to the Author):

SUMMARY

(2) Overall, the authors have responded well to the comments, and most of the points have been appropriately addressed. We appreciate the careful revisions and clarifications made throughout the manuscript. However, a few minor issues still remain, which we outline below. In addition, the figures in the supplementary materials are currently very small and difficult to read. We strongly encourage the authors to enlarge them to improve visibility and accessibility for readers.

[Response 2] We thank the reviewer for their positive overall assessment and helpful suggestions. In response to the concern about figure readability in the supplementary materials, we have carefully revised all relevant figures by enlarging them and increasing the font size of all text elements to enhance clarity and accessibility. We have ensured that they are now easy to read when viewed or printed at standard scale. We appreciate the reviewer's attention to this detail.

(3) Thank you for your detailed response and for clarifying the issues regarding the search strategy. We appreciate the added screenshots and the explanation about the quotation marks possibly causing confusion. However, we would still strongly recommend that the exact search strings used for each database be provided in full, including any field tags, Boolean operators, and filters applied. This is essential to ensure full reproducibility of your search.

For example, for Web of Science, it would be helpful to specify whether you used tags such as TS=, TI=, or others, and how the keywords were combined. Similarly, for Google Scholar and the China Knowledge Resource Integrated Database, we encourage you to include the full queries (as entered) that yielded the reported results.

[Response 3] Thank you for your thoughtful comment and for emphasizing the importance of full transparency and reproducibility in reporting the literature search strategy. In response, we have revised the manuscript to provide the exact search strings used in each database, including field tags, Boolean operators, and any filters applied

(as shown in lines 111–135 of the revised manuscript). Specifically:

- **Web of Science:** We used the following search query in the Topic field (TS=): TS=(beta diversity OR β diversity) AND (biotic homogen* OR biotic differentiat* OR biotic heterogen* OR biological homogen* OR biological differentiat* OR biological heterogen*).
- **Google Scholar:** We used the same keyword combination entered directly into the general search bar: (beta diversity OR β diversity) AND (biotic homogen* OR biotic differentiat* OR biotic heterogen* OR biological homogen* OR biological differentiat* OR biological heterogen*). As Google Scholar does not support advanced field-based syntax, we manually screened the top search results based on relevance.
- **China National Knowledge Infrastructure (CNKI):** We searched using the Chinese equivalents of the topic terms: (beta diversity OR β diversity OR biotic homogenization OR biotic heterogenization) translated into Chinese and entered in the search bar using simplified characters.

We have clarified in the revised manuscript that the total number of 35,206 studies shown in Supplementary Figure 1 represents the total number of records retrieved from each individual database after removing duplicates within each database. However, some studies were shared across databases, so cross-database duplicates were still present in this number. We then added 229 additional studies identified through backward and forward citation tracking of key reviews and foundational studies (e.g., Catano et al. 2017; Kramer et al. 2023; McKinney 2006; Petsch et al. 2022; Rolls et al. 2023; Soininen et al. 2018). After merging all records from the three databases and the citation-tracked studies, we conducted a final cross-database deduplication, resulting in 14,977 unique studies for screening.

We believe this revised description improves the clarity, transparency, and reproducibility of our literature search process. Thank you again for your valuable feedback.

(4) Thank you for adding a detailed quality assessment framework and providing the checklist

in Supplementary Table 1. We appreciate your effort to enhance transparency in study appraisal. However, we would kindly like to ask for further clarification on two points:

Was study quality (e.g., quality score or evidence level) used as a moderator in any sensitivity analyses or meta-regressions? If so, we would appreciate a pointer to the relevant results (main text or Supplementary).

While the checklist is detailed, would it be possible to include a brief summary in the main text (e.g., number or proportion of studies classified as weak, moderate, strong, very strong) to improve understanding for readers who may not check the supplementary files?

[Response 4] Thank you for your valuable suggestions. We would like to clarify the following:

First, study quality was not used as a moderator in sensitivity analyses or meta-regressions because all included studies achieved a quality score greater than 75%, classified as ‘very strong’ according to our framework. As such, there was insufficient variation in quality scores to warrant their use as a moderator, and we prioritized ensuring the robustness of our findings by including only studies of very strong quality.

Second, we appreciate your suggestion to summarize the study quality assessment in the main text. In the revised manuscript, we have added a brief summary noting that all included studies were categorized as very strong evidence (quality score >75%), with the exact scores for each study provided in Supplementary file - *Study quality assessment.xlsx*. We believe this addition will enhance transparency and reader understanding (as shown in lines 176–181 of the revised manuscript). For clarity and completeness, we have re-submitted the file *Study quality assessment.xlsx* along with this revised version.

(5) Thank you for providing the detailed explanation regarding collinearity and the rationale for retaining all moderators in the random forest analysis. However, we noticed that the variable “growth form”, which was mentioned in the author's response as being highly correlated with other variables, is not actually included in Figure 6.

Could the authors clarify whether growth form was excluded from the random forest models? If so, we would suggest briefly explaining why it was excluded, especially

given its reported correlation with taxa type and realm. Additionally, even if retained variables are correlated, a short discussion of how multicollinearity might affect variable importance rankings (e.g., through shared explanatory power) would be helpful for readers interpreting the results.

[Response 5] Thank you for your careful reading and for bringing this to our attention. We apologize for the confusion caused by the inconsistent terminology between our previous response and the manuscript. Specifically, what we referred to as “growth form” in our response corresponds to “life form” in the revised manuscript and in Figure 6. We confirm that life form was indeed included in the random forest models, and the apparent discrepancy was solely due to this inconsistency in terms.

Additionally, in the revised manuscript, we have included a sentence clarifying that although some variables (such as life form, taxa type, and realm) are correlated, random forest models are generally robust to multicollinearity. However, we acknowledge that multicollinearity can still influence the interpretation of variable importance due to shared explanatory power. We have briefly addressed this point in the manuscript to help readers more accurately interpret the variable importance rankings (as shown in lines 550–556 of the revised manuscript).

We appreciate your helpful suggestion, which has improved the clarity of our methods and results.

(6) Thank you for clarifying the rationale for using different circle sizes in Figure 1B. However, we would like to respectfully suggest reconsidering this design choice. In general, varying symbol size is interpreted as representing quantitative information (e.g., magnitude, frequency), and using size purely to distinguish categorical variables can be visually misleading. Since the three beta-diversity dimensions are already clearly differentiated by colour, using size as an additional encoding may add unnecessary visual complexity or confusion.

[Response 6] Thank you for this thoughtful suggestion. We fully agree that symbol size is typically interpreted as representing quantitative information, and using it to distinguish categorical variables could indeed lead to confusion. In response to your

comment, we have revised Figure 1B by replacing size-based encoding with shape-based encoding to distinguish the three β -diversity dimensions. Specifically, we now use circles for taxonomic β -diversity (TB), triangles for functional β -diversity (FB), and squares for phylogenetic β -diversity (PB). We believe this revision improves the figure's clarity and avoids any potential misinterpretation. The figure caption and legend have been updated accordingly.