

Meta-analysis reveals widespread negative associations between species richness and ecological uniqueness

Corresponding Author: Professor Xugao Wang

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

Attachments originally included by the reviewers as part of their assessment can be found at the end of this file.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

I have reviewed the MS entitled Meta-analysis reveals widespread negative associations between species richness and ecological uniqueness across the globe.

The MS is generally well written and easy to follow. However, it lacks some important methodological details and coverage of the relevant literature.

General comments

On one hand, I applaud the exercise of evaluating the richness-uniqueness relationship (RUR) on many datasets across ecosystems and taxonomic groups. The main result that the RUR is generally negative is an important one. On the other hand, I believe that the paper makes a poor job of testing the many different processes that could affect the direction and strength of the RUR. Instead, the paper focusses on species replacement (LCBD_repl) as the core process. In particular, the paper fails to revisit the results of Dansereau, Legendre and Poisot (2022), who extensively tested the properties of the LCBD metric on a large dataset of bird communities.

Dansereau, G., Legendre, P. and Poisot, T., 2022. Evaluating ecological uniqueness over broad spatial extents using species distribution modelling. *Oikos*, 2022(5), p.e09063.

It is clear from Dansereau et al.'s results that the strength and direction of RUR depends on the richness of the species pool, the breadth of the richness gradient across site, the spatial extent of the sampling, as well as the proportion of rare species in the pool (Dansereau et al., 2022). Furthermore, their results clearly show that the canonical RUR is not linear (rather strongly concave), and therefore Pearson's correlation is perhaps not a proper approximation of its strength and direction.

Specific comments

It is not clear if the meta regression aims at testing the "strength" OR "direction" of the RUR as both terms are used in the introduction (line 39 and line 71 of the introduction) as well as in the results. There may be different processes affecting the strength or the direction. These differences are not well explained (but see Dansereau et al., 2022).

Judging from Figure 4, there would be no association between Rate of species replacement and RUR if only strength was tested (absolute Fisher Z). But then, Fisher Z is a metric mostly related to strength (Pearson's r) rather than direction. This is a problem.

It is also not clear why should we expect changes in the strength or direction of the RUR across taxa and ecosystems. Are factors such as the size of the species pool, the breadth of the richness gradient, the spatial extent, or the proportion of rare species, expected to vary across taxa and ecosystems, thus explaining some of the differences? Probably. These are not detailed.

More fundamentally, the authors partitioned the LCBD metric (beta diversity) into its components of species replacement (LCBD_repl) and richness difference (LCBD_rich).

They then tested if the strength and direction of the RUR is affected by LCBD_repl. First, it is mentioned in the MS that the effect of LCBD_rich was also evaluated (line 437), but I could not find these results. Second, since both the slope and Pearson's correlation of the RUR depend on LCBD values, and that LCBD_repl also depends on LCBD values ($LCBD_{tot} = LCBD_{repl} + LCBD_{rich}$), there is a high risk of spurious association here. A full simulation should be conducted to avoid any artifactual findings.

The above concerns are major, and therefore I cannot recommend the paper for publications.

(Remarks on code availability)

Reviewer #2

(Remarks to the Author)

MAJOR CONCERNS

In its current state, this is not a high-quality meta-analysis and does something I've never seen before which is only use snowballing methods to find studies for the synthesis (e.g., screening and including only studies that cite a key paper, here Legendre & Ce Caceres 2013). Typically, this approach is meant to supplement an exhaustive keyword search of multiple databases -- here they skipped the exhaustive search and went straight to screening and including only studies citing Legendre & Ce Caceres (2013). They also supplement sources but do not justify why these sources were chosen or whether they are representative of what's available. How many studies were missed with these unusual decisions? There is also no information on who screened the studies, or whether there were any practices to validate screening decisions. I suggest the authors have a look at PRISMA guidelines to develop practices that generate high quality syntheses.

They also apply a custom weighting that will have unknown consequences for the meta-analysis and publication bias tests because the variance of Fisher's Z is meant to strictly model sampling error among effect sizes (and the variability among these weights later, in turn, impact the random-effects estimated by the meta-analysis model), but here they've confounded this weight and its purpose with aspects of spatial extent without justification. The meta-analysis model should not be modified like this, and if you believe spatial extent impacts effect sizes then treat it as a continuous predictor in the model, and then explore pooled effects as marginal means corrected by spatial extent. Also please report your full meta-analysis model; it's unclear what random effects were included, what predictors are combined to achieve synthesis, or whether these were implemented correctly with the metafor R package. Results are also incomplete as they emphasize P values but do not report the statistic of these P values. On Line 131, you also cannot do a meta-analysis of absolute values of effect sizes because this moves the null from zero to an unknown null value -- making significant tests meaningless.

A lot of this work (some of it is quite amazing and remarkable) is simply waste because odd decisions were made that make it impossible to assess the robustness of what was achieved.

MINOR NOTES

- 1) Lines 37-38: do you mean "have yet to be synthesized"
- 2) k instead of N should be used to describe number of effect sizes (N is reserved for number of data points used to estimate Z)
- 3) report the degrees of freedom for all Q tests

(Remarks on code availability)

I could not review the code as it was bundled as a .rar file which my institution does not permit opening or extracting files from.

Reviewer #3

(Remarks to the Author)

In this article, the authors carried out a global meta-analysis to investigate the relationships between species richness and ecological uniqueness (RUR), defined as the local contribution to beta diversity (LCBD), and to study the role of species replacement and environmental covariates in determining the direction and strength of RUR across taxa and ecosystems.

The methodology of the meta-analysis is well explained and summarized in Figure 2.

The resulting dataset is very interesting for studying patterns of community composition across taxa and ecosystems.

The authors refer to the local beta diversity contribution (LCBD) as defined by Legendre and De Cáceres (2015).

It can be calculated using the beta.div function in the R adespatial package, which I will do to illustrate my issues below.

I'm not sure that the questions and hypotheses (as illustrated in Figure 1) that are addressed here are sufficiently relevant, and that the way of analyzing and interpreting RUR patterns can support the authors' conclusions.

More specifically, my main concern is that the authors have not taken into account the influence of dispersal limitation on RUR.

This influence is only mentioned late in the Discussion (L227): "Moreover, habitat isolation can also have non-negligible effect on the RURs since it limits the capacity of species to disperse between local communities, which are highly correlated with species replacement".

It is known that dispersal limitation leads to a decrease in beta diversity, e.g., as a function of distance (Condit et al. 2002), which in turn should influence LCBD variation in assemblages.

Here's a specific example:

Consider a neutral case such as the Hubbell's (2001) mainland-island metacommunity model. The following R code simulates many neutral communities connected to a common mainland, with varying migration rates.

```
#devtools::install_github("frmunoz/ecolottery/pkg", build_vignettes = TRUE)

require(ecolottery)

# Building a neutral mainland assemblage, as in Hubbell (2001)
J <- 500;
pool <- coalesc(100*J, theta = theta)$com

# Simulate a number of neutral communities of same size J=500, for varying migration rate from the mainland, between
0.002 and 1, with 1/500 increment
m <- 1/500
comm1 <- list()
for(i in 1:500)
comm1[[i]] <- coalesc(J, m*i, pool=pool)

# Let us calculate the richness of simulated communities
S1 <- unlist(lapply(comm1, function(x) nlevels(as.factor(x$com$sp))))
plot(S1,(1:500)/500, ylab="m")
# We note here that richness increases with decreasing dispersal limitation, as expected
# See Reviewer #3 Review Attachment #1

# Making a community x species table
tab1 <- matrix(0,nrow=500,ncol=Spool)
colnames(tab1) <- levels(as.factor(pool$sp))
for(i in 1:500) {
sad.loc1 <- table(comm1[[i]]$com$sp)
tab1[i, names(sad.loc1)] <- sad.loc1
}

# Calculating the Local Contributions to Beta Diversity of the simulated communities
contrib1 <- adespatial::beta.div(tab1)$LCBD

plot(contrib1, S1, log="xy")
# See Reviewer #3 Review Attachment #2

cor.test(contrib1, S1, method="spearman")
# Spearman's rank correlation rho
#
#data: contrib1 and S1
#S = 39535164, p-value < 2.2e-16
#alternative hypothesis: true rho is not equal to 0
#sample estimates:
# rho
#-0.8976955
```

Consequently, there may be a strongly negative RUR (shown in Reviewer #3 Review Attachment #2) for purely neutral communities that differ only in immigration intensity.

To provide a convincing analysis of the drivers of RUR in taxa and ecosystems, the study should thus acknowledge the influence of dispersal limitation and determine/discuss the extent to which dispersal limitation and environmental variation may influence RUR.

References

Condit, Richard, Nigel Pitman, Egbert G. Leigh Jr., Jérôme Chave, John Terborgh, Robin B. Foster, Percy Núñez, et al. « Beta-Diversity in Tropical Forest Trees ». *Science* 295 (janvier 2002): 666-69.

Hubbell, S. P. The Unified Neutral Theory of Biodiversity and Biogeography. 1 vol. Princeton and Oxford: Princeton University Press, 2001.

Legendre, P. & De Cáceres, M. Beta diversity as the variance of community data: dissimilarity coefficients and partitioning. *Ecol. Lett.* 16, 951-963 (2013).

(Remarks on code availability)

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

Reviewer #2

(Remarks to the Author)

This new version significantly improves clarity on what was achieved, and because the approach applies more established methodology (e.g., broad search and inclusion of the primary literature, conventional weighting, PRISMA methodology for reporting), it is also easier to assess the robustness of findings. That said, there are still ways to make things more repeatable and clear.

More work needs to be done on how studies were found and selected for synthesis. The goal is to be clear so that the work presented is repeatable; but also that the correlations analyzed is representative of what's been done. In particular that null and positive correlations were extracted neutrally without bias for negative correlations (publication bias tests help with this, but should be performed among the three databases to assess individual retrieval bias). It is great that the synthesis now includes correlations from primary literature; however, there are gaps on how these new data were retrieved. Given the many authors, how many people were involved in screening studies for inclusion? Was work divided among members, or was there a dual screening design used to improve repeatability of screening decisions? Surely some of the new studies did not explicitly report raw correlations, but presented results as regression coefficients, or slope t-tests, or maybe data were only available in figures. What methods were used to retrieve these more difficult, yet highly relevant correlations? There must have also been some overlap between the different approaches to find correlations—were studies identified in the web of science search that also were found in the figshare database? How were duplicate studies handled? All three PRISMA plots should be merged as a single plot; since the end-product is a combined synthesis of all these sources. Further, given that correlations can originate from three sources (e.g., primary literary, custom databases, etc.), a sensitivity analysis is necessary to assess whether they represent a homogeneous group and that the data can be combined under a single analysis.

The first section of the results is essentially a type of vote-counting where they enumerate the number of significant and non-significant studies (e.g., lines 120 to 122, or parts of Fig. 2). Don't do this, this undermines the whole purpose of doing a meta-analysis—which carefully does not use significance of each study but rather the magnitude, direction, and sampling error to make inferences. It's ok to report the composition of the dataset (e.g., number of studies per taxa group, or among categories of responses relative to richness), but not differences in significance. The rank order of effects to emphasize representation of negative effects is ok, but discussing their prevalence without the lens of meta-analysis is dangerous as many studies may be underpowered and visually may not be representative of the underlying (aggregate) effect. Having correlations as Fisher-Z transformed effects is better for clarity since results are analyzed with the metric (e.g., main inferences are derived from these) and consistence helps with understanding what was done in the paper. The results also need to define the acronyms (PA, AB) and more descriptive words should be used for CV_richness, Clim_PC1, and Clim_PC2. Since some details about the methodology are embedded outside the main manuscript, it's important that they are defined so that the reader knows right-way on how to interpret what was reported. These include the above acronyms, but also measure of % explained (line 176 but explained using different definitions in line 568). Many of the results reported are described as "affects", for example this "affected" that. If a moderator is being used to explain variation among effects, maybe "predict" is a better inferential description, since correlations are just associations and not causes, and similarly, moderators are predictors and not causes of variability. Finally, lines 517-518 state that significance was tested by 95% overlap, yet there are only p-values reported throughout the results; what is the test statistics these p-values, and their values with degrees of freedom need to be reported along with p-values.

Minor comments:

- 1) What is the scale of the taxonomic group? Are they Linnean rankings, or homologous groups?
- 2) Line 194, also 403: "literature studies" should be "primary studies"
- 3) Line 409: Define what a "unit" is.
- 4) Line 439: What do you mean by "previous researches"?

Reviewer #4

(Remarks to the Author)

This manuscript aims to describe the relationship between, and drivers of, species richness and ecological uniqueness. Generally, the manuscript reads very well and the comments of reviewer 1 and 3 are well addressed. However, when the quadratic models are tested (as suggested by reviewer 1) and found more appropriate compared to the linear models in 30% of the cases, these records are not considered as the results are harder to interpret. I think that the better fit of the quadratic relationship in such a high percentage of the records show that the relationship can not always be interpreted as linear and that these cases should be considered and not disregarded as the results are harder to interpret. In that way you are only using the results which are convenient, while this is not a good representation of what the data actually shows. Additionally, different drivers of the RURs are tested, following four different hypotheses. In the methods, I recommend a more thorough description of why and how well the variables chosen reflect these particular mechanisms underlying the proposed hypotheses. Generally, I strongly recommend to first run a fully randomised null model, to check if the obtained relationships, and therefore your key results, are obtained by chance or that ecological mechanisms drive this relationship. Lastly, from a traits perspective, beta diversity can be different between sites, but some sites are more ecologically unique than others as the niche and traits of the species differ more compared to the other sites. This is a point which should be considered in the introduction and discussion.

Version 2:

Reviewer comments:

Reviewer #2

(Remarks to the Author)

In terms of what was reported and presented with this meta-analysis, I believe this new version has the clarity and detail that is more in line with current expectations in research synthesis. My comments below are very very minor.

MAIN TEXT

lines 832, 838, 851: "Fisher's Z statistics" should be "Fisher's Z-transformed correlation coefficients"

lines 864-876: definition of effect sizes (see above) is missing

SUPPLEMENTARY TEXT 1:

line 11: where was this search done? E.g., using who's institutional access to IS Web of science?

line 32: Which PRISMA statement did you use? cite source

line 3863, 3872, 3879, 3890, 3898, 3906, 3978, 3990, 4005, 4018, 4030, 4033, 4037, 4054, 4070, 4086, 4102, 4118, 4134, 4150, 4166, 4182: "Fisher's Z statistics" should be "Fisher's Z-transformed correlation coefficients"

Reviewer #4

(Remarks to the Author)

I would like to thank the authors for the thorough answers and additional analyses performed. The comments are well addressed. The only question which remains for me is about the outliers in Figure 3C. Are the datapoints at the right side of the figure indeed checked and found to be valid or do these datapoints represent less robust observations? In the latter case I would disregard these datapoints and change the text accordingly, otherwise all good.

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/>

Response to reviewers

Reviewer #1 (Remarks to the Author):

I have reviewed the MS entitled Meta-analysis reveals widespread negative associations between species richness and ecological uniqueness across the globe.

The MS is generally well written and easy to follow. However, it lacks some important methodological details and coverage of the relevant literature.

General comments

On one hand, I applaud the exercise of evaluating the richness-uniqueness relationship (RUR) on many datasets across ecosystems and taxonomic groups. The main result that the RUR is generally negative is an important one. On the other hand, I believe that the paper makes a poor job of testing the many different processes that could affect the direction and strength of the RUR. Instead, the paper focusses on species replacement (LCBD_repl) as the core process. In particular, the paper fails to revisit the results of Dansereau, Legendre and Poisot (2022), who extensively tested the properties of the LCBD metric on a large dataset of bird communities.

Dansereau, G., Legendre, P. and Poisot, T., 2022. Evaluating ecological uniqueness over broad spatial extents using species distribution modelling. Oikos, 2022(5), p.e09063.

Response: We sincerely appreciate your insightful comments and suggestions. Following your suggestions, we have included more detailed methodological descriptions, such as the definition of RUR forms and the selection of explanatory variables (See Methods and Supplementary Text 5). **We have also incorporated additional key literatures (including the study by Dansereau et al. 2022) by re-conducting the literature search and screening, to strengthen the contextual foundation of our work.**

Specifically, we re-conducted the literature search and screening by searching for peer-reviewed journal articles using the ISI Web of Science with predefined keyword combinations to obtain an extensive collection of studies that focused on ecological uniqueness globally (See Supplementary Figure S1 and Supplementary Text 1 for details). Following the PRISMA-guided screening process, we finally retained 100 qualified studies, which comprised 56 case studies explicitly investigating ecological uniqueness and 44 studies containing compatible observational datasets from literature studies. A PRISMA diagram summarizing the workflow of literature search is included in Supplementary Figure S1.

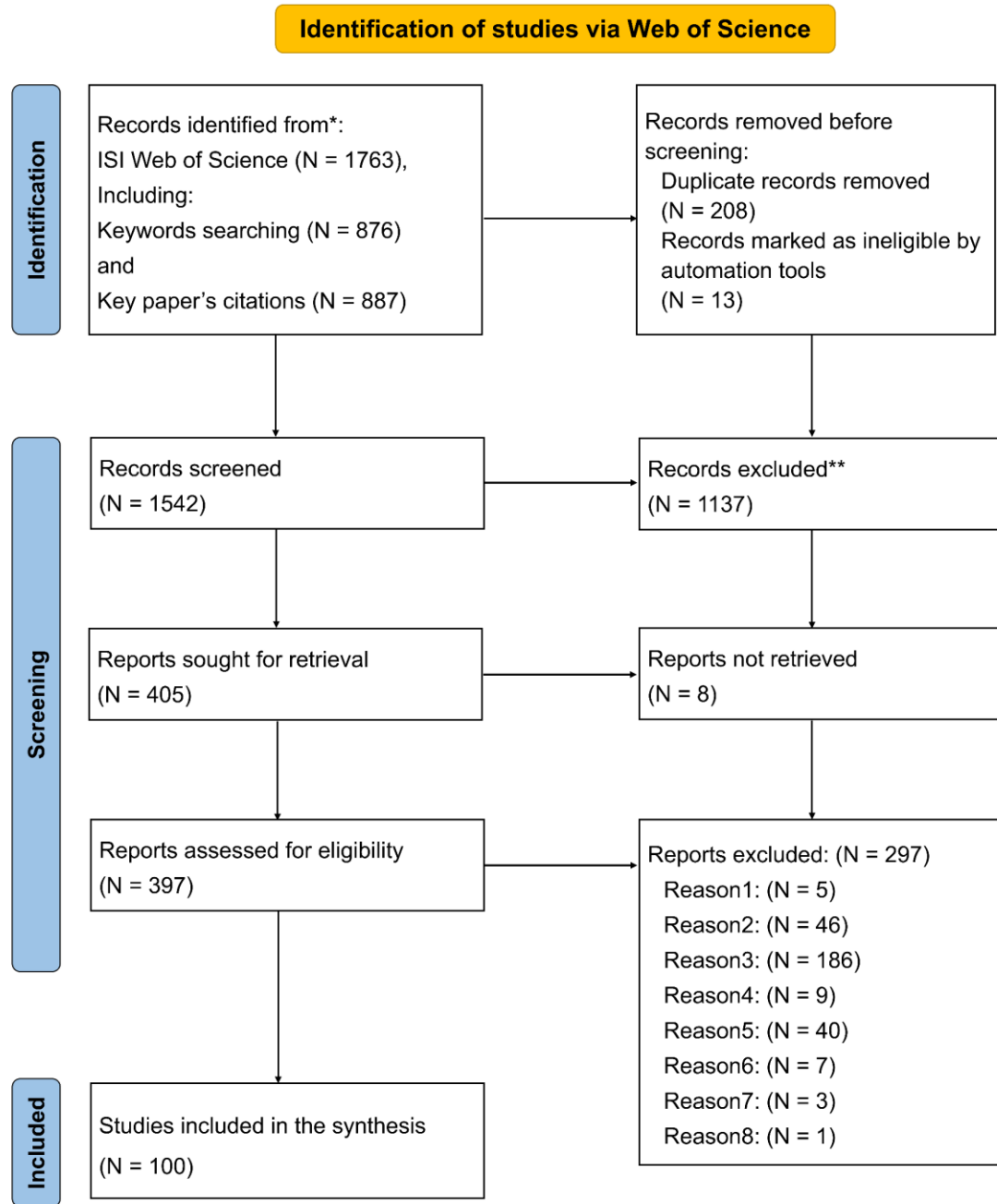


Figure S1 PRISMA diagram showing the process of locating Peer-reviewed literature studies included in this synthesis. N: number of studies.

Regarding the processes that could affect the variation of RUR, we have tested the effects of four distinct hypotheses: (1) species pool, (2) dispersal limitation, (3) environmental filtering and (4) spatial grain size and their relative importance on RUR in the current manuscript (Lines 52-100). We selected seven predictors related to these hypotheses in the analyses (See Section: *Quantifications of different hypotheses* in Supplementary Text 5 for details). Overall, we found that PA (presence-absence) -based RURs were primarily affected by species pool attributes (i.e. the proportion of rare species and gamma diversity),

whereas AB (Abundance) -based RURs were more strongly influenced by dispersal limitation (See Figures 3-5 in the main text for details).

References:

Dansereau, G., Legendre, P. & Poisot, T. Evaluating ecological uniqueness over broad spatial extents using species distribution modelling. *Oikos* **2022**, e09063 (2022).

It is clear from Dansereau et al. 's results that the strength and direction of RUR depends on the richness of the species pool, the breadth of the richness gradient across site, the spatial extent of the sampling, as well as the proportion of rare species in the pool (Dansereau et al., 2022). Furthermore, their results clearly show that the canonical RUR is not linear (rather strongly concave), and therefore Pearson's correlation is perhaps not a proper approximation of its strength and direction.

Response: Thanks very much for your constructive comments. In this revised version, we carefully reviewed the study by Dansereau et al. (2022) and incorporated it into the initial dataset of literature review.

For the first comment on the incomplete analysis in our study when comparing with Dansereau et al (2022), we re-constructed our synthesis to explore the role of four different hypotheses in driving RUR following your suggestions and others (i.e. species pool, dispersal limitation, environmental filtering, and spatial grain size; Lines 52-100). In this case, “the regional species pool size, the breadth of species richness, and the proportion of rare species” you mentioned were used to represent species pool in this new version. Specifically, the regional species pool size was represented by Gamma diversity (i.e. total number of species within each study). For breadth of species richness, we calculated coefficient of variation of species richness among different communities (CV_richness) for each observation, with larger CV_richness indicating wider breadth of species richness within each observation. The proportion of rare species was defined through dual criteria corresponding to different data types (i.e. presence-absence (PA) and abundance (AB); See Section: *Quantifications of different hypotheses* in Supplementary Text 5 for details).

As for the spatial extent of the sampling you have pointed out, we found that spatial extent data were often unavailable in most studies. Moreover, it was challenging to define the overall spatial extent of the study area for certain study groups (e.g., fish and freshwater macroinvertebrates), which limited our ability to assess its influence on RUR (Lines 343-350). **Therefore, we alternatively considered the role of spatial grain size as another important component of spatial scale (See Barton et al., 2013).** It was defined as the sampling area of each study unit within each study if reported (See Section: *Quantifications of different hypotheses* in Supplementary Text 5 for details). We also highlighted that future studies would benefit from more standardized and explicit definitions of sampling scales,

enabling systematic analyses of scale-dependent RUR patterns using more extensive and integrative datasets (Lines 353-355).

Our results revealed that *species pool* and *spatial grain size* were important drivers of RUR (See Figures 3-4 in the main text). For *species pool*, the effect size of RUR consistently declined with increasing gamma diversity (Figure 3a-b in the main text; Supplementary Table S8). By contrast, the effect size of RUR increased with CV_richness and the proportion of rare species (especially in PA data; Figure 3 in the main text; Supplementary Table S8). For *spatial grain size*, it negatively affected effect size of RUR although effects were generally weak in both single factor (Figure 4a-b in the main text, Supplementary Table S8) and full model analyses (Figure 5a-b in the main text; Supplementary Table S16b).

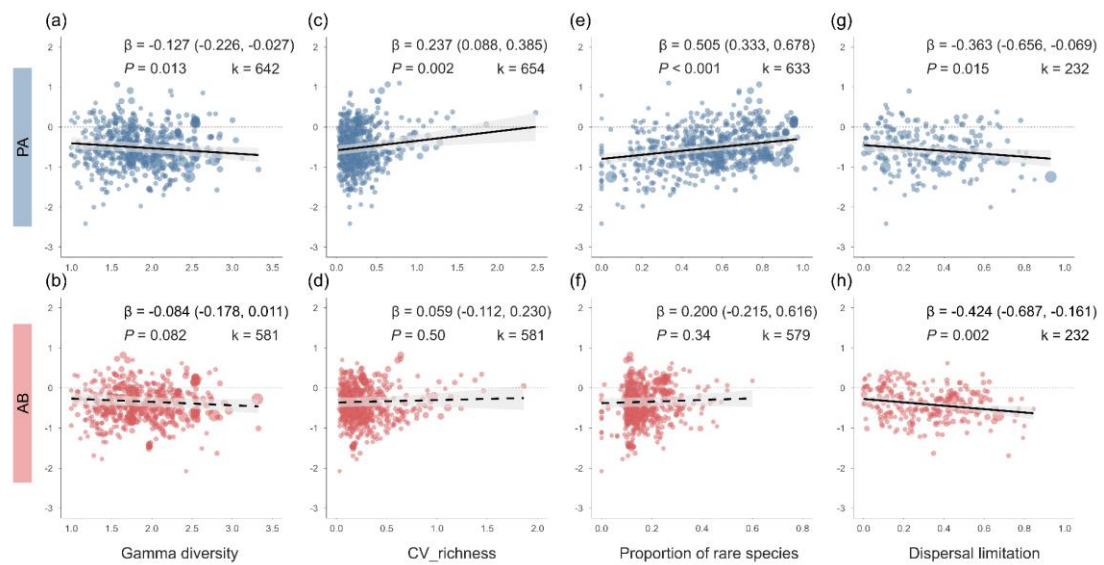


Figure 3 Variation of the effect sizes of RUR (i.e. Fisher's Z statistics) with **gamma diversity** (a-b), **CV_richness** (c-d), **proportion of rare species** (e-f), and **dispersal limitation** (g-h) based on presence absence (PA) and abundance (AB) data types. The regression lines represent the response to the above factors with the 95% confidence intervals shaded from hierarchical mixed-effects meta-regressions, and the values of the estimated coefficients (β) and 95% confidence intervals are shown in each panel. The solid and dashed lines represent significant ($P < 0.05$) and non-significant ($P > 0.05$) relationships. The points have radii relative to the inverse variance of each observation (i.e., observations with larger points have lower variance in effect size). Gamma diversity was log-transformed before modelling. CV_richness: the breadth of species richness. See Table S8 for details of model statistics. k: Number of effect sizes.

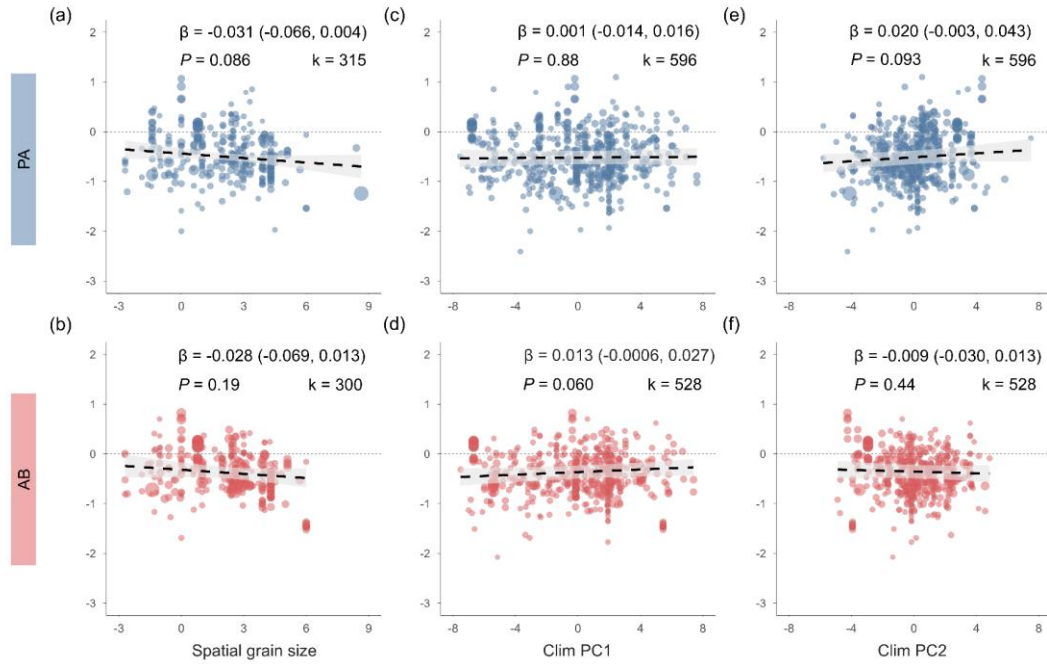


Figure 4 Variation of the effect sizes of RUR (i.e. Fisher's Z statistics) with **spatial grain size** (a-b), Clim PC1 (c-d) and Clim PC2 (e-f) based on presence absence (PA) and abundance (AB) data types. The regression lines represent the response to the above factors with the 95% confidence intervals shaded from hierarchical mixed-effects meta-regressions, and the values of the estimated coefficients (β) and 95% confidence intervals are shown in each panel. The solid and dashed lines represent significant ($P < 0.05$) and non-significant ($P > 0.05$) relationships. The points have radii relative to the inverse variance of each observation (i.e., observations with larger points have lower variance in effect size). Spatial grain size was log-transformed before modelling. Clim PC1 and Clim PC2: The first two principal components of bioclimatic variables from the WorldClim database. See Table S8 for details of model statistics. k: Number of effect sizes.

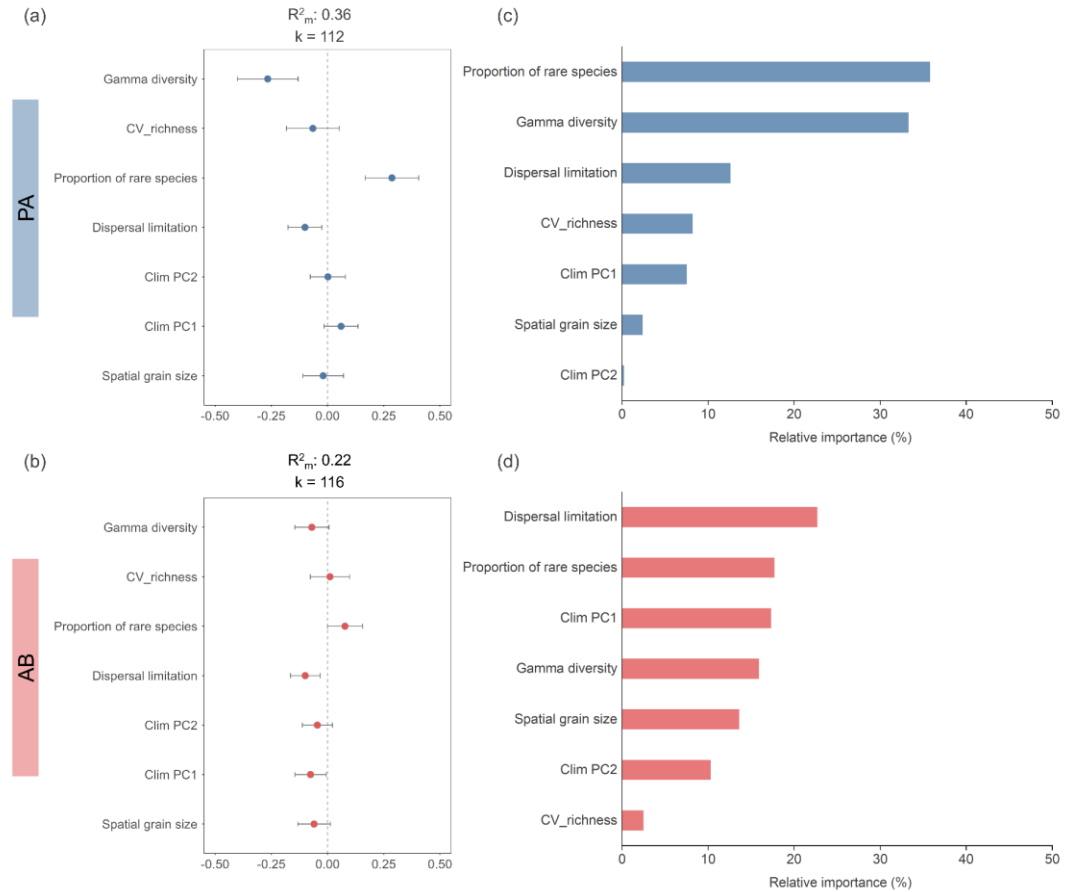


Figure 5 (a-b) Explanatory power of predictor variables on effect sizes of RUR from multiple hierarchical mixed-effects meta-regressions, and (c-d) the relative importance of multiple factors in determining the effect size of RUR across four major taxa groups (i.e. Terrestrial plants, freshwater macroinvertebrates, birds and terrestrial insects) based on presence absence (PA) and abundance (AB) data types. The effect of each factor is considered to be significant if the 95% confidence intervals did not overlap zero. Each factor is standardized before modelling and the relative importance of each factor is expressed as the percentage of variance explained. Gamma diversity and spatial grain size were log-transformed before modelling. CV_richness: the breadth of species richness. Clim PC1 and Clim PC2: The first two principal components of bioclimatic variables from the WorldClim database. See Table S16a, S16b for details of model statistics. R^2_m : the amount of variance explained by multiple fixed factors. k: Number of effect sizes.

For the second comment regarding different formats (i.e. linear or nonlinear) of RUR, we identified linear (first-order) and quadratic (second-order) models as two predominant forms of RUR characterization through a systematic review of previous researches (See Appendix 1) and compared these two formulations across all studies (See section: *Effect size of RUR* in Methods; Lines 438-457). Our analysis revealed that 32.1% of records (616 out of 1918) exhibited significantly better fit with quadratic model than linear model ($P < 0.05$; See Appendix 2), while 67.9% (1302 records) showed no statistical difference between linear and quadratic formulations (i.e. $P > 0.05$). For these cases, the

differences in AIC or AICc between linear and quadratic models were also less than 2 units, further supporting their statistical equivalence (Burnham & Anderson 2004; Appendix 2). These results suggested that the linear model provided a more parsimonious representation of RUR patterns for these 1302 observations (Lines 457-464). Given the challenges in deriving standardized effect sizes for non-linear (quadratic) relationships of RUR, we restricted subsequent analyses to the 1302 records best represented by linear models to ensure methodological consistency and comparability across studies (Lines 468-472).

References:

- Barton, P. S. et al. The spatial scaling of beta diversity. *Glob. Ecol. Biogeogr.* **22**, 639–647 (2013).
- Burnham, K. P. & Anderson, D. R. Multimodel inference: Understanding AIC and BIC in model selection. *Sociol. Methods Res.* **33**, 261–304 (2004).

Specific comments

It is not clear if the meta regression aims at testing the “strength” OR “direction” of the RUR as both terms are used in the introduction (line 39 and line 71 of the introduction) as well as in the results. There may be different processes affecting the strength or the direction. These differences are not well explained (but see Dansereau et al., 2022).

Judging from Figure 4, there would be no association between Rate of species replacement and RUR if only strength was tested (absolute Fisher Z). But then, Fisher Z is a metric mostly related to strength (Pearson’s r) rather than direction. This is a problem.

Response: Thank you for these thoughtful and constructive comments. In the revised manuscript, we focused on characterizing the overall distribution of RUR and identifying their ecological drivers. Specifically, we aimed to show the patterns that significant negative RURs are common while positive RURs are relatively rare (i.e. we mainly focused on the direction in this study). To avoid ambiguity, we no longer frame our analyses in terms of testing the “strength” of RUR.

In addition, we agree with your concerns related to the use of species replacement in explaining RUR in Figure 4 of former manuscript. Therefore, **we have removed this part of the analysis in the manuscript**. We alternatively incorporated four different hypotheses: species pool (i.e. regional species pool size; breadth of species richness and proportion of rare species), dispersal limitation, environmental filtering, and spatial grain size in the current manuscript (Lines 52-100). Then, we mainly explored the effects of these hypotheses in determining RUR and used Fisher's Z statistics as the effect size of RUR rather than to directly infer strength *per se*.

We find that variations of RUR effect sizes are associated with several factors, such

as species pool size (i.e. gamma diversity), breadth of species richness and the proportion of rare species. The effect sizes of RUR tend to decrease with increasing gamma diversity, indicating a stronger tendency toward negative associations (Figure 3a-b in the main text). In contrast, the effect sizes of RUR tend to increase with higher proportion of rare species, suggesting a potential shift toward more positive associations (Figure 3e-f in the main text). These clarifications are now reflected particularly in the discussions (Lines 244-247; Lines 257-259).

It is also not clear why should we expect changes in the strength or direction of the RUR across taxa and ecosystems. Are factors such as the size of the species pool, the breadth of the richness gradient, the spatial extent, or the proportion of rare species, expected to vary across taxa and ecosystems, thus explaining some of the differences? Probably. These are not detailed.

Response: Thank you for pointing this out. **We did not pursue comparisons of RUR across ecosystems and taxonomic groups, as RUR is, as you rightly pointed out, more strongly influenced by factors such as species pool size, the breadth of species richness, the proportion of rare species, and spatial grain size.** Therefore, we have revised the manuscript to focus on how these factors influence the overall distribution and variation of RUR effect sizes, both in general and across different taxonomic groups (Lines 520-564), and have largely removed comparisons across taxa and ecosystems.

We found significant negative RUR for both PA and AB data (Figure 2 in the main text). **Taxon-specific analyses demonstrated significantly negative RURs across different taxonomic groups in both PA and AB datasets, such as birds, terrestrial plants, fishes and terrestrial insects** (Supplementary Figure S3; Table S7). We also found that the drivers of RUR differed by data type: PA-based RURs were primarily shaped by attributes of *species pool* such as gamma diversity and proportion of rare species, while AB-based RURs were more strongly influenced by *dispersal limitation*. *Environmental filtering* and *spatial grain size* exhibited relatively weak effects and low contributions (Figures 3-5 in the main text).

More fundamentally, the authors partitioned the LCBD metric (beta diversity) into its components of species replacement (LCBD_repl) and richness difference (LCBD_rich). They then tested if the strength and direction of the RUR is affected by LCBD_repl. First, it is mentioned in the MS that the effect of LCBD_rich was also evaluated (line 437), but I could not find these results. Second, since both the slope and Pearson's correlation of the RUR depend on LCBD values, and that LCBD_repl also depends on LCBD values ($LCBD_{tot} = LCBD_{repl} + LCBD_{rich}$), there is a high risk of spurious association here. A full simulation should be conducted to avoid any artifactual findings.

Response: **For the first comment on the LCBD_rich result,** we partitioned ecological uniqueness (LCBD metric) into the species replacement component (the local contributions of species replacement to beta diversity; $LCBD_{repl}$) and richness difference (the local contributions of richness difference to beta diversity; $LCBD_{rich}$)

(Legendre 2014), and the proportion explained by richness difference is equal to one minus the proportion explained by species replacement (i.e., the sum of the proportions explained by species replacement and richness difference equals one). Therefore, we only provided the proportion explained by species replacement in the previous manuscript. We no longer considered this factor in the revised manuscript (See response below).

For the second comment, we agree with the point that there is a high risk of spurious association between LCBD and its components, $LCBD_{repl}$ and $LCBD_{rich}$. **Therefore, we have removed this part of the analysis in the revised manuscript.** Instead, we now focused on the effects and relative importance of four different hypotheses: species pool, dispersal limitation, environmental filtering, and spatial grain size on the effect size of RUR (See Figures 3-4 in the main text). **Because the general hypotheses in this study have been changed in this version following your comments and the comments of other reviewers (Lines 52-100), we believe that the full simulation you mentioned is not necessary here.** Nevertheless, we sincerely appreciate your careful suggestions.

References:

Legendre, P. Interpreting the replacement and richness difference components of beta diversity. *Glob. Ecol. Biogeogr.* **23**, 1324-1334 (2014).

The above concerns are major, and therefore I cannot recommend the paper for publications.

Response: We sincerely thank you for your constructive comments and valuable suggestions. In response, we have carefully and thoroughly revised the manuscript. We believe that the changes made in this version adequately address both the major and minor concerns raised.

Reviewer #2 (Remarks to the Author):

MAJOR CONCERNS

In its current state, this is not a high-quality meta-analysis and does something I've never seen before which is only use snowballing methods to find studies for the synthesis (e.g., screening and including only studies that cite a key paper, here Legendre & Ce Caceres 2013). Typically, this approach is meant to supplement an exhaustive keyword search of multiple databases -- here they skipped the exhaustive search and went straight to screening and including only studies citing Legendre & Ce Caceres (2013). They also supplement sources but do not justify why these sources were chosen or whether they are representative of what's available. How many studies were missed with these unusual decisions? There is also no information on who screened the studies, or whether there were any practices to validate screening decisions. I suggest the authors have a look at PRISMA guidelines to develop practices that generate high quality syntheses.

Response: Thank you very much for this helpful comment. Following your suggestions, **we have improved the synthesis by incorporating three distinct approaches to allow for a broad and representative collection of studies or datasets** (See Supplementary Text 1-4 for details).

(1) Peer-reviewed literature studies: First, we conducted a systematic search of peer-reviewed journal articles in the ISI Web of Science using predefined keyword combinations to compile a comprehensive set of studies focused on ecological uniqueness globally. The keyword structure and methodological framework, which follow the PRISMA guidelines, are now described in detail in Supplementary Text 1-2 and illustrated in Supplementary Figure S1.

(2) Public data repositories: Second, we compiled additional raw datasets across different taxonomic groups from public data repositories, as both ecological uniqueness and species richness can be calculated from species composition datasets (Legendre & De Cáceres, 2013). The search terms and methods used for dataset collection are now detailed in Supplementary Text 1, Text 3 and Supplementary Figure S2.

(3) Compiled datasets: Third, we compiled a series of additional raw datasets from our own data. These datasets were mainly acquired through: (1) direct correspondence with principal investigators maintaining unpublished datasets, or (2) using the existing data sources including: GrassPlot (Biurrun et al., 2021; Dengler et al., 2018), sPlot (Bruehlheide et al., 2019) and CESTES (Jeliazkov et al., 2020) by ourselves. The detailed procedures are now described in Supplementary Text 1 and Supplementary Text 4.

In summary, our final dataset incorporated a total of 1918 data records from 451

literature studies or datasets (Figure 1 in the main text). The final dataset included three types of data sources: Peer-reviewed literatures studies (100 studies, 288 data records), public data repositories (243 datasets, 1415 data records) and compiled datasets (108 datasets, 215 data records; See Figure 1 in the main text and Supplementary Table S1). However, for consistency in description, we uniformly referred to each dataset or literature study as a "study" in the subsequent text (Lines 402-407).

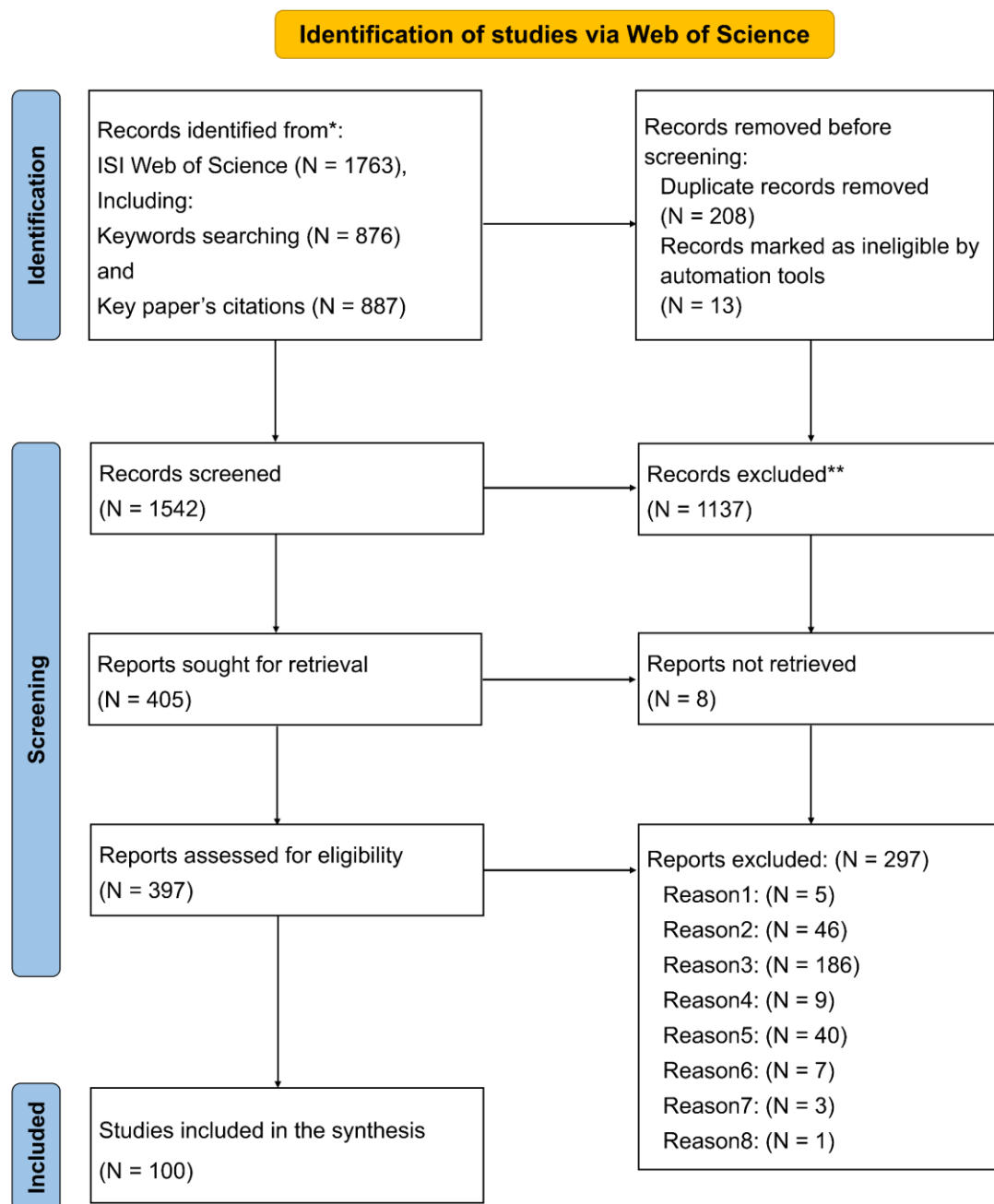


Figure S1 PRISMA diagram showing the process of locating Peer-reviewed literature studies included in this synthesis. N: number of studies.

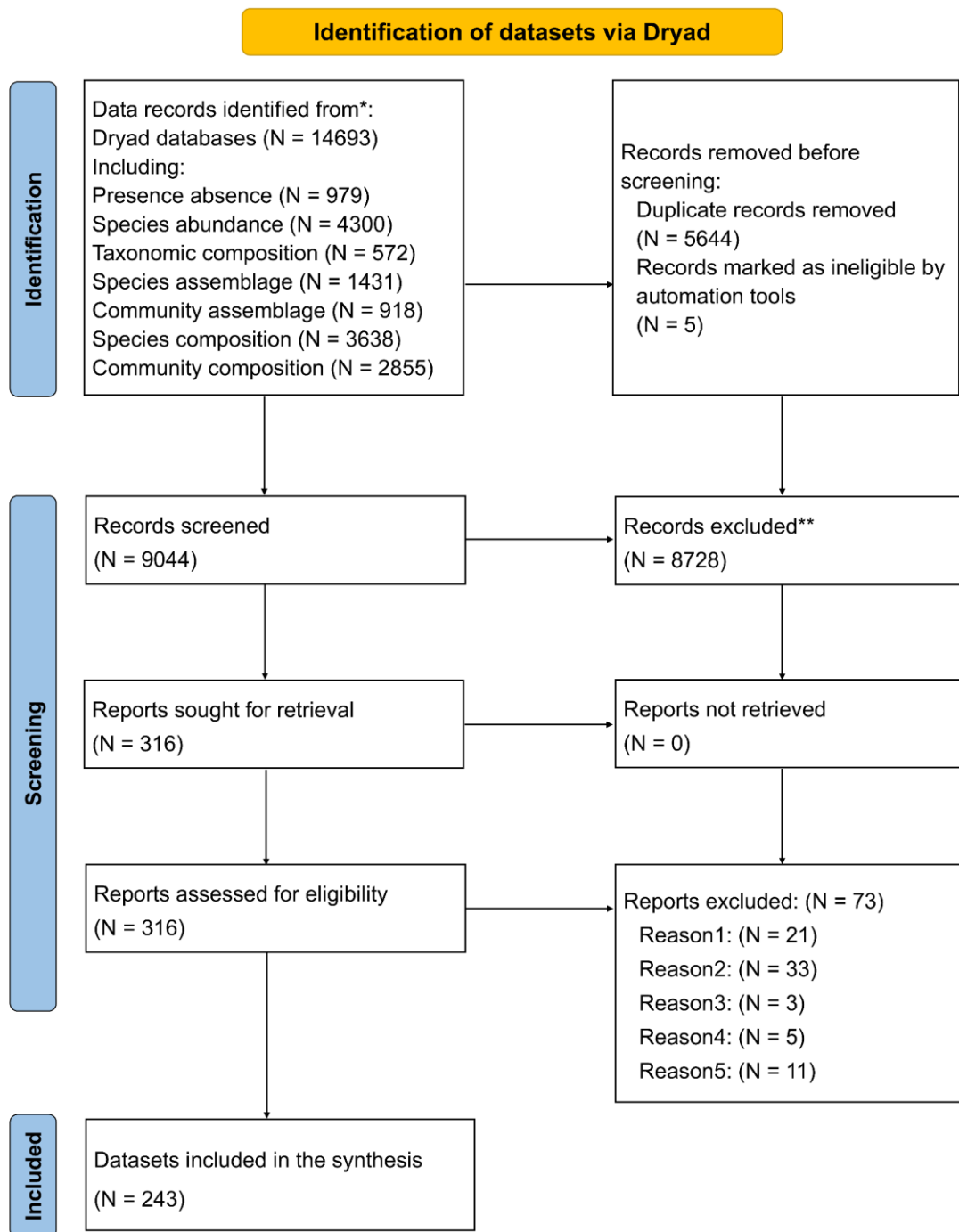


Figure S2 PRISMA diagram showing the process of locating datasets included in this synthesis from Dryad repository. N: number of datasets.

References:

- Biurrun, I. et al. Benchmarking plant diversity of Palaearctic grasslands and other open habitats. *J. Veg. Sci.* **32**, e13050 (2021).
- Bruehlheide, H. et al. sPlot – A new tool for global vegetation analyses. *J. Veg. Sci.* **30**, 161–186 (2019).
- Dengler, J. et al. GrassPlot – A database of multi-scale plant diversity in Palaearctic grasslands. *Phytocoenologia* **48**, 331–347 (2018).
- Jeliazkov, A. et al. A global database for metacommunity ecology, integrating species, traits, environment and space. *Scientific Data* **7**, 6 (2020).
- Legendre, P. & De Cáceres, M. Beta diversity as the variance of community data: dissimilarity coefficients and partitioning. *Ecol. Lett.* **16**, 951-963 (2013).

They also apply a custom weighting that will have unknown consequences for the meta-analysis and publication bias tests because the variance of Fisher's Z is meant to strictly model sampling error among effect sizes (and the variability among these weights later, in turn, impact the random-effects estimated by the meta-analysis model), but here they've confounded this weight and its purpose with aspects of spatial extent without justification. The meta-analysis model should not be modified like this, and if you believe spatial extent impacts effect sizes then treat it as a continuous predictor in the model, and then explore pooled effects as marginal means corrected by spatial extent.

Also please report your full meta-analysis model; it's unclear what random effects were included, what predictors are combined to achieve synthesis, or whether these were implemented correctly with the metafor R package. Results are also incomplete as they emphasize P values but do not report the statistic of these P values.

Response: Thank you for your suggestions. We agree that including additional variables in the calculation of corresponding weights is inappropriate. Therefore, in this revised version, **we retained Fisher's Z transformation as the effect size of RUR but removed other factors from the definition of weights to ensure consistency and clarity in our analysis.**

We weighted each effect size based on the sample size (i.e., the number of sites for which of ecological uniqueness and species richness were calculated) using the formula:

$$v_i = \frac{1}{N_s - 3} \quad (2)$$

Where N_s is the sample size (i.e., number of sampling units) of each estimate within each study. This indicates that studies had larger sample sizes would be given higher weights (Lines 495-501).

In accordance with your suggestion, we included spatial grain size as an explanatory variable in our analysis to account for the variation in effect sizes

of RUR. Spatial grain size was defined as the sampling area of each study unit within each study if reported (Barton et al., 2013; See Section: *Quantifications of different hypotheses* in Supplementary Text 5).

We agree that the previous version may not provide detailed descriptions of the statistical analysis. To address this issue, we have included additional details in the revised version. For **the inclusion of predictors**, we explored the effect and relative importance of four distinct hypotheses on the effect size of RUR. They are: (1) species pool, (2) dispersal limitation, (3) environmental filtering and (4) spatial grain size (Lines 52-100). We examined the potential influence of seven predictors belong to four hypotheses including: (1) gamma diversity, (2) CV_richness, (3) proportion of rare species, (4) dispersal limitation, (5) Clim PC1, (6) Clim PC2 and (7) spatial grain size on the effect sizes of RUR across different taxa groups (See Section: *Quantifications of different hypotheses* in Supplementary Text 5 for details). We first conducted single-factor analyses for each predictor. We then examined relative importance of these hypotheses by conducting a full-model analyses incorporating different predictors (See *Statistical analysis* for details, Lines 520-564).

For **the inclusion of random effects**, we implemented an inverse-variance-weighted hierarchical mixed-effects meta-regression model incorporating a multi-level random effects structure when assessing the overall mean effect size across data types (PA and AB data). **The model specified two independent random effects to account for potential sources of heterogeneity: (1) nested random effects of repeated individual RUR estimates within each study and (2) an independent random effect for taxonomic group to address between-taxa heterogeneity** (Lines 504-509). We also conducted subgroup analyses to examine RURs across different taxonomic groups. we retained only the nested random effect (i.e., estimates nested within studies), as taxonomic group was the grouping variable being analyzed (Lines 510-514).

It should be noted that we conducted single-factor analyses separately for each taxa group. In these analyses, **the nested random effect of estimate within study was retained whenever applicable**. However, for taxa groups where most studies include only one estimate, we specified random effects at the study level instead (Lines 535-540).

Regarding the reporting of results, we agree that presenting only P -value is insufficient. **Therefore, in the revised manuscript, we have included the regression coefficients along with their 95% confidence intervals (CIs), variance estimates for random effects, marginal R^2 values (explained variance by fixed effects), overall heterogeneity (I^2) and the I^2 values corresponding to each level of random effects** (See *Statistical analysis* for details). These results are presented in Figures 3–5 in the main text and Supplementary Tables S5–S17. The data and R code for analysis are available in Figshare at

<https://figshare.com/s/a974caae4f643878afc7>

References:

Barton, P. S. et al. The spatial scaling of beta diversity. *Glob. Ecol. Biogeogr.* **22**, 639–647 (2013).

On Line 131, you also cannot do a meta-analysis of absolute values of effect sizes because this moves the null from zero to an unknown null value -- making significant tests meaningless.

Response: We agree that analyzing the absolute values of Fisher's Z statistics was not appropriate in the previous version of the manuscript. In the revised manuscript, we have removed the corresponding results and method descriptions. **To assess the effect of environmental filtering, we alternatively extracted 19 bioclimatic variables (Bio1-Bio19) from the WorldClim database (Hijmans et al., 2005) with 30 arc seconds spatial resolution using each study's mean geographic coordinates.** Principal Component Analysis (PCA) was applied to mitigate multicollinearity among these variables. The first two principal components (Clim PC1 and Clim PC2) were retained as synthetic climatic predictors for explaining the effect size of RUR in subsequent analyses. (See Section: *Quantifications of different hypotheses* in Supplementary Text 5 for details).

References:

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G. & Jarvis, A. Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.* **25**, 1965–1978 (2005).

A lot of this work (some of it is quite amazing and remarkable) is simply waste because odd decisions were made that make it impossible to assess the robustness of what was achieved.

Response: In this revised manuscript, we have carefully updated the data sources and checked the statistical methods for the synthesis as well as the corresponding descriptions in the Methods section. Specifically, we (1) updated the PRISMA flowchart and data sources (Supplementary Text 1-4; Supplementary Figures 1-2), (2) revised the definition of variance (Lines 495-501), and (3) provided more detailed model parameter settings (especially for the specification of random effects; Lines 504-514) as well as model statistics and results (Supplementary Tables S5-S17). We believe that these issues have now been properly addressed and that the revised manuscript provides an accurate and appropriate description of the statistical methods used (Lines 504-582).

MINOR NOTES

1) Lines 37-38: do you mean "have yet to be synthesized"

Response: We have clarified this point in the text. It now reads: “Despite perceived importance, the patterns of this relationship and underlying mechanisms have yet to be synthesized at the global scale.” (Lines 36-37).

2) k instead of N should be used to describe number of effect sizes (N is reserved for number of data points used to estimate Z)

Response: Thank you for pointing this out. We have made this correction and used k to represent the number of effect sizes in analyses (Figures 3-5 in the main text, Supplementary Tables S6-S17).

3) report the degrees of freedom for all Q tests

Response: We have reported the degrees of freedom corresponding to the Q statistic in the supplementary table (See Supplementary Tables S5-S17 for details).

Reviewer #2 (Remarks on code availability):

I could not review the code as it was bundled as a .rar file which my institution does not permit opening or extracting files from.

Response: Thank you for pointing this out. In this revised manuscript, the R code and the associated dataset supporting the findings of this study are available in Figshare at <https://figshare.com/s/a974caae4f643878afc7> Meanwhile, we have also clarified this issue in the Code Availability statement of the revised manuscript (Lines 588-590).

Reviewer #3 (Remarks to the Author):

In this article, the authors carried out a global meta-analysis to investigate the relationships between species richness and ecological uniqueness (RUR), defined as the local contribution to beta diversity (LCBD), and to study the role of species replacement and environmental covariates in determining the direction and strength of RUR across taxa and ecosystems.

The methodology of the meta-analysis is well explained and summarized in Figure 2. The resulting dataset is very interesting for studying patterns of community composition across taxa and ecosystems.

Response: Thank you for this positive assessment of our work.

The authors refer to the local beta diversity contribution (LCBD) as defined by Legendre and De Cáceres (2015).

It can be calculated using the beta.div function in the R adespatial package, which I will do to illustrate my issues below.

Response: Thank you for this suggestion. We applied two alternative approaches to calculate ecological uniqueness and found these approaches are highly correlated (Lines 420-435). Therefore, we prioritized the method based on the *beta.div* function in the current synthesis, as it is the most commonly used in studies of ecological uniqueness (Chen et al., 2024; da Silva et al., 2018; Vilmi et al., 2017; Yao et al., 2021).

References:

- Chen, Y. et al. Multiple processes jointly determine ecological uniqueness across forest plant life-forms in Northeast China. *J. Biogeogr.* **51**, 1133-1147 (2024).
- da Silva, P. G., Medina Hernandez, M. I. & Heino, J. Disentangling the correlates of species and site contributions to beta diversity in dung beetle assemblages. *Divers. Distrib.* **24**, 1674-1686 (2018).
- Vilmi, A., Karjalainen, S. M. & Heino, J. Ecological uniqueness of stream and lake diatom communities shows different macroecological patterns. *Divers. Distrib.* **23**, 1042-1053 (2017).
- Yao, J., Huang, J., Ding, Y., Xu, Y., Xu, H. & Zang R. Ecological uniqueness of species assemblages and their determinants in forest communities. *Divers. Distrib.* **27**, 454-462 (2021).

I'm not sure that the questions and hypotheses (as illustrated in Figure 1) that are addressed here are sufficiently relevant, and that the way of analyzing and interpreting RUR patterns can support the authors' conclusions.

Response: We have conducted a thorough revision on the manuscript based on your suggestions and comments from other reviewers. We carefully checked our

hypotheses and analytical methods. We have removed the original hypotheses and the corresponding data analysis from the previous version of the manuscript. In the current version, we focus primarily on examining the effects and relative importance of four distinct hypotheses: (1) species pool; (2) dispersal limitation; (3) environmental filtering and (4) spatial grain on RUR (Lines 52-100). Accordingly, we have also revised the analytical methods to align with the updated hypotheses (See *Statistical analysis* in the main text for details). After this revision, we believe the data analysis and hypotheses about RUR are clearer and more relevant.

More specifically, my main concern is that the authors have not taken into account the influence of dispersion limitation on RUR.

This influence is only mentioned late in the Discussion (L227): "Moreover, habitat isolation can also have non-negligible effect on the RURs since it limits the capacity of species to disperse between local communities, which are highly correlated with species replacement".

It is known that dispersal limitation leads to a decrease in beta diversity, e.g., as a function of distance (Condit et al. 2002), which in turn should influence LCBD variation in assemblages.

Here's a specific example:

Consider a neutral case such as the Hubbell's (2001) mainland-island metacommunity model. The following R code simulates many neutral communities connected to a common mainly, with varying migration rates.

```
#devtools::install_github("frmunoz/ecolottery/pkg", build_vignettes = TRUE)
```

```
require(ecolottery)
```

```
# Building a neutral mainland assemblage, as in Hubbell (2001)
```

```
J <- 500;
```

```
pool <- coalesc(100*J, theta = theta)$com
```

```
# Simulate a number of neutral communities of same size J=500, for varying migration rate from the mainland, between 0.002 and 1, with 1/500 increment
```

```
m <- 1/500
```

```
comm1 <- list()
```

```
for(i in 1:500)
```

```
comm1[[i]] <- coalesc(J, m*i, pool=pool)
```

```
# Let us calculate the richness of simulated communities
```

```
S1 <- unlist(lapply(comm1, function(x) nlevels(as.factor(x$com$sp))))
```

```
plot(S1,(1:500)/500, ylab="m")
```

```
# We note here that richness increases with decreasing dispersal limitation, as expected
```

See Reviewer #3 Review Attachment #1

```
# Making a community x species table  
tab1 <- matrix(0,nrow=500,ncol=Spool)  
colnames(tab1) <- levels(as.factor(pool$sp))  
for(i in 1:500) {  
  sad.loc1 <- table(comm1[[i]]$com$sp)  
  tab1[i, names(sad.loc1)] <- sad.loc1  
}
```

Calculating the Local Contributions to Beta Diversity of the simulated communities
contrib1 <- adespatial::beta.div(tab1)\$LCBD

```
plot(contrib1, S1, log="xy")  
# See Reviewer #3 Review Attachment #2
```

```
cor.test(contrib1, S1, method="spearman")  
# Spearman's rank correlation rho  
#  
#data: contrib1 and S1  
#S = 39535164, p-value < 2.2e-16  
#alternative hypothesis: true rho is not equal to 0  
#sample estimates:  
# rho  
#-0.8976955
```

Consequently, there may be a strongly negative RUR (shown in Reviewer #3 Review Attachment #2) for purely neutral communities that differ only in immigration intensity.

To provide a convincing analysis of the drivers of RUR in taxa and ecosystems, the study should thus acknowledge the influence of dispersal limitation and determine/discuss the extent to which dispersal limitation and environmental variation may influence RUR.

References

Condit, Richard, Nigel Pitman, Egbert G. Leigh Jr., Jérôme Chave, John Terborgh, Robin B. Foster, Percy Núñez, et al. « Beta-Diversity in Tropical Forest Trees ». Science 295 (janvier 2002): 666-69.

Hubbell, S. P. The Unified Neutral Theory of Biodiversity and Biogeography. 1 vol. Princeton and Oxford: Princeton University Press, 2001.

Legendre, P. & De Cáceres, M. Beta diversity as the variance of community data:

Response: Thank you for this insightful suggestion. Following the suggestions from you and other reviewers, we have incorporated four distinct hypotheses into this synthesis: (1) species pool, (2) dispersal limitation (as you specifically suggested), (3) environmental filtering, and (4) spatial grain size (Lines 52-100). We explored these hypotheses and their corresponding factors to assess their potential impacts and relative importance on the effect sizes of RUR (See Section: *Quantifications of different hypotheses* in Supplementary Text 5 and *Statistical analysis* for details).

To assess the potential influence of dispersal limitation on RUR, we employed the method of distance-based Moran's Eigenvector Maps (dbMEM) to model spatial structure across study sites (See Section: *Quantifications of different hypotheses* in Supplementary Text 5 for details). **Our results showed that dispersal limitation exerted significantly negative effects on RUR** (Figure 3g-h in the main text; Supplementary Table S8) especially for freshwater macroinvertebrates and terrestrial insects (Supplementary Table S12). The same trend persisted after accounting for other variables in the full multivariate models (Figure 5a-b in the main text; Supplementary Table S16b), highlighting the robustness of this negative effect. For relative contributions, we found that dispersal limitation (22.7%) was the primary determinant of RUR based on AB data while it was also an important driver of PA-based RUR (Figure 5c-d in the main text; Supplementary Table S16b).

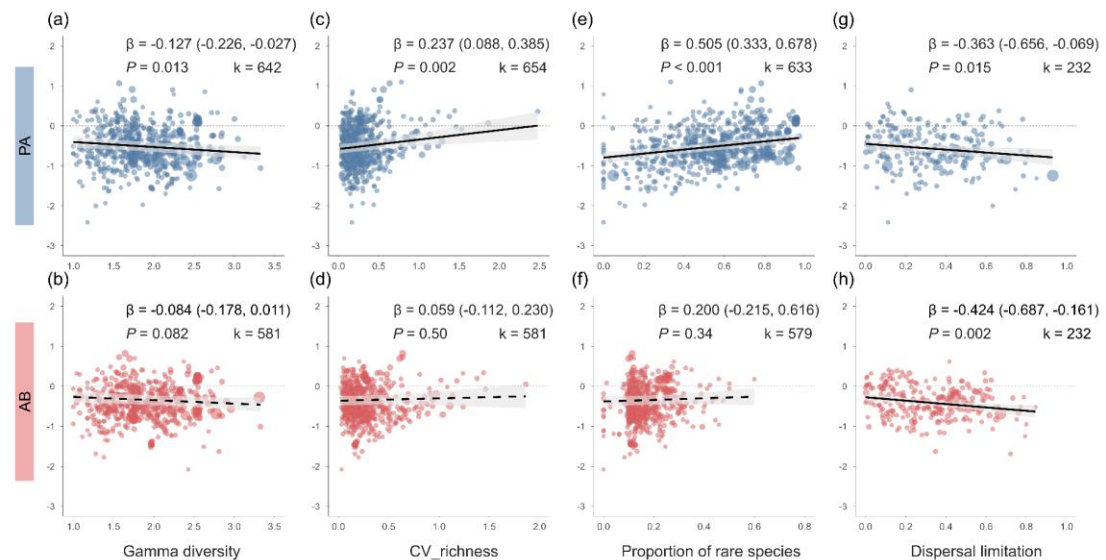


Figure 3 Variation of the effect sizes of RUR (i.e. Fisher's Z statistics) with gamma diversity (a-b), CV_richness (c-d), proportion of rare species (e-f), and **dispersal limitation** (g-h) based on presence absence (PA) and abundance (AB) data types. The regression lines represent the response to the above factors with the 95% confidence intervals shaded from hierarchical mixed-effects meta-regressions, and the values of the estimated coefficients (β) and 95% confidence intervals are shown in each panel. The solid and dashed lines represent significant ($P < 0.05$) and non-significant ($P > 0.05$) relationships. The points have radii relative to the inverse variance of each observation (i.e., observations with larger points have lower variance).

in effect size). Gamma diversity was log-transformed before modelling. CV_richness: the breadth of species richness. See Table S8 for details of model statistics. k: Number of effect sizes.

Importantly, dispersal limitation had greater explanatory power in AB models (Figure 5c-d in the main text; Supplementary Table S16b), likely because dispersal limitation governs both range sizes and abundance gradients (Condit et al., 2000). For instance, poor dispersers (e.g., wind-dispersed pioneer plant) may form dense, localized populations, with sharp abundance gradients along sites (Condit et al., 2000; Wang et al., 2010). while superior dispersers (e.g., large seeds for animal-dispersed trees) may achieve broad but shallow abundance distributions, thereby homogenizing community abundance patterns (Wang et al., 2010). AB metrics may capture these patterns more effectively than PA data (Lines 292-299).

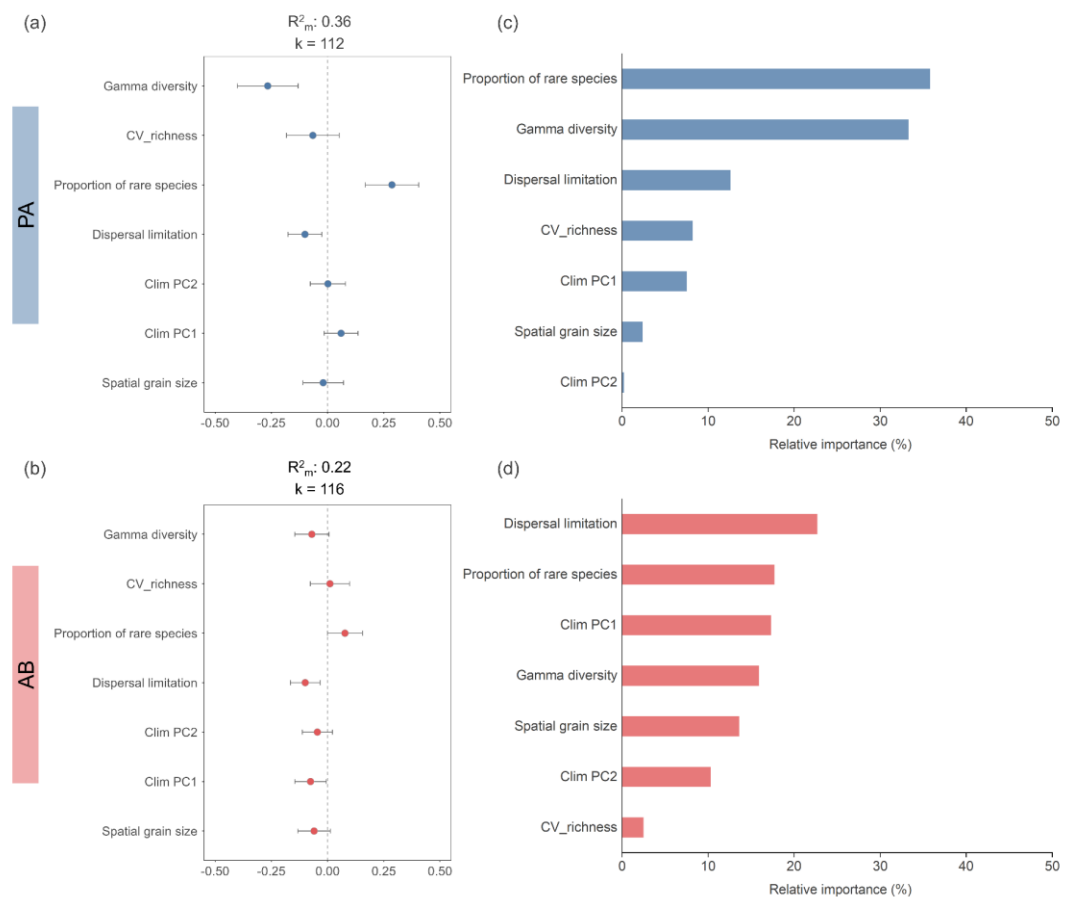


Figure 5 (a-b) Explanatory power of predictor variables on effect sizes of RUR from multiple hierarchical mixed-effects meta-regressions, and (c-d) the relative importance of multiple factors in determining the effect size of RUR across four major taxa groups (i.e. Terrestrial plants, freshwater macroinvertebrates, birds and terrestrial insects) based on presence absence (PA) and abundance (AB) data types. The effect of each factor is considered to be significant if the 95% confidence intervals did not overlap zero. Each factor is standardized before modelling and the relative importance of each factor is expressed as the percentage of variance explained. Gamma diversity and spatial grain size were log-transformed before modelling. CV_richness: the breadth of species richness. Clim PC1 and Clim PC2: The first two principal

components of bioclimatic variables from the WorldClim database. See Table S16a, S16b for details of model statistics. R^2_m : the amount of variance explained by multiple fixed factors. k: Number of effect sizes.

In addition, we have added text to the Discussion to highlight how the effect of dispersal limitation on RUR may also depend on several dispersal-related traits (e.g. physiological or morphological traits) of different taxa groups (Alzate & Onstein, 2022) (Lines 356-362)

References:

- Alzate, A. & Onstein, R. E. Understanding the relationship between dispersal and range size. *Ecol. Lett.* **25**, 2303–2323 (2022).
- Condit, R. et al. Spatial patterns in the distribution of common and rare tropical tree species: a test from large plots in six different forests. *Science* **288**, 1414–1418 (2000).
- Wang, X., Ye, J., Li, B., Zhang, J., Lin, F. & Hao, Z. Spatial distributions of species in an old-growth temperate forest, northeastern China. *Can. J. For. Res.* **40**, 1011–1019 (2010).

Response to reviewers

Reviewer #2 (Remarks to the Author):

This new version significantly improves clarity on what was achieved, and because the approach applies more established methodology (e.g., broad search and inclusion of the primary literature, conventional weighting, PRISMA methodology for reporting), it is also easier to assess the robustness of findings. That said, there are still ways to make things more repeatable and clear.

Response: Thank you for this positive assessment of our work.

More work needs to be done on how studies were found and selected for synthesis. The goal is to be clear so that the work presented is repeatable; but also that the correlations analyzed is representative of what's been done.

Response: Thank you for this valuable suggestion. In this revision, we have revised the Methods section to provide a more detailed description of our study selection process. This includes specifying the number of authors involved in screening and how any discrepancies were resolved to ensure consistency (See Supplementary Text 1 for details). Furthermore, we have now listed all excluded primary studies and datasets under each of our pre-defined criteria, enhancing the transparency and reproducibility of our workflow (See Supplementary Text 5 for details). Additionally, to address the representativeness of the synthesized data, we conducted publication bias tests on the effect sizes from different data sources (i.e. primary studies, data repositories and compiled datasets). **The results confirmed that no significant publication bias was detected across the various data sources (See Supplementary Figures S5-S8),** so the correlations analyzed here are representative.

In particular that null and positive correlations were extracted neutrally without bias for negative correlations (publication bias tests help with this, but should be performed among the three databases to assess individual retrieval bias).

Response: Thank you for the suggestion. We have conducted publication bias tests on the effect sizes from different data sources (i.e. primary studies, data repositories and compiled datasets). **The results indicate that there is no significant publication bias in the effect sizes obtained from any of the data sources (See Supplementary Figures S5-S8).**

It is great that the synthesis now includes correlations from primary literature; however, there are gaps on how these new data were retrieved. Given the many authors, how many people were involved in screening studies for inclusion? Was work divided among members, or was there a dual screening design used to improve repeatability of screening decisions?

Response: Thank you for this important question regarding the screening process. To ensure the reliability and reproducibility of our study screening and selection,

we indeed adopted a dual-independent screening approach. Specifically, two authors independently screened all the identified studies and datasets against the pre-defined exclusion criteria. This process was conducted simultaneously to minimize bias. These clarifications have been added to Supplementary Text 1:

“Following the PRISMA guidelines, two authors independently conducted the literature and dataset screening based on the above exclusion criteria. The two authors examined 3 primary studies and 4 datasets that were only selected by one of the authors. Following a joint re-assessment of the studies and datasets with initial screening discrepancies between the two authors, a full consensus was reached to exclude all of them. We have listed all unselected literature and datasets with the reasons for exclusion (See Supplementary Text 5).”

We believe such a rigorous, dual-independent screening method significantly strengthens the validity of our inclusion process and the overall findings of our synthesis.

Surely some of the new studies did not explicitly report raw correlations, but presented results as regression coefficients, or slope t-tests, or maybe data were only available in figures. What methods were used to retrieve these more difficult, yet highly relevant correlations?

Response: Thank you for this comment. As you pointed out, some studies presented results as regression coefficients, t-statistics, or in figures. While it is methodologically possible to convert such statistics into Pearson's correlation coefficients (Nakagawa & Cuthill, 2007), **we made a conscious decision to exclude studies that provided only these summary statistics without the underlying raw data or values for species richness and ecological uniqueness.** This decision was driven by a critical methodological consideration raised by Reviewer #1, who noted that the RUR may not always be linear (e.g., it could be quadratic). Because 1) to empirically test the form of the relationship for each study, it was essential for us to have access to the data distributions; 2) relying solely on converted summary statistics would have presupposed a linear association and prevented us from verifying this key assumption; in this case, we prioritized data integrity and the ability to assess the relationship's form rather than simply maximize sample size. **This specific inclusion criterion is explicitly stated in our screening protocol** (See Supplementary Text 1):

“(3) We excluded studies that did not report the distributions of both species richness and ecological uniqueness values, or studies that only report related statistics (e.g. regression coefficients, slope t-tests or model estimates) without providing richness or uniqueness distributions. Although it is methodologically feasible to convert such statistics into correlation coefficients (Nakagawa & Cuthill, 2007), this exclusion was necessary because previous empirical evidence demonstrates that the RUR may exhibit either linear or nonlinear (e.g. curve-linear) relationships (Dansereau et al., 2022; Luukkonen et al., 2024; Hill et al., 2021). Without detailed distributions of species richness and ecological uniqueness, we could not reliably characterize the association form of RUR. To empirically test the

form of the relationship for each study, it was essential for us to have access to the data distributions. Moreover, relying solely on converted summary statistics would have presupposed a linear association and prevented us from verifying this key assumption. Such incomplete reporting would fundamentally limit our capacity to distinguish between alternative ecological models while introducing substantial uncertainty in the synthesis (Reason 3, 186 studies; Fig. S1)”.

References:

- Dansereau, G., Legendre, P. & Poisot, T. Evaluating ecological uniqueness over broad spatial extents using species distribution modelling. **2022**, *Oikos*, e09063 (2022).
- Hill, M. J. et al. Local contributions to beta diversity in urban pond networks: Implications for biodiversity conservation and management. *Divers. Distrib.* **27**, 887-900 (2021).
- Luukkonen, S., Heino, J., Hjort, J., Helm, A. & Alahuhta, J. A macroecological analysis of ecological uniqueness of freshwater macrophyte assemblages across Europe and North America. *J. Ecol.* **2024**, 1–15 (2024).
- Nakagawa, S. & Cuthill, I. C. Effect size, confidence interval and statistical significance: a practical guide for biologists. *Biol. Rev. Camb. Philos. Soc.* **82**, 591–605 (2007).

There must have also been some overlap between the different approaches to find correlations—were studies identified in the web of science search that also were found in the figshare database? How were duplicate studies handled? All three PRISMA plots should be merged as a single plot; since the end-product is a combined synthesis of all these sources.

Response: Thank you for this important comment. Indeed, there was an overlap of studies identified through the literature search in Web of Science and dataset search from the Dryad repositories. All duplicate records were meticulously identified and removed to ensure each unique dataset was represented only once in our synthesis.

This is now stated in exclusion criteria of public data repositories:

“(5) We also excluded studies sharing identical datasets (i.e., the same dataset used in primary studies or multiple data publications containing the same datasets) (Reason 5, 11 datasets; Figure S1).” (See Supplementary Text 1 for details)

The number of excluded duplicates is clearly documented in the merged PRISMA flow diagram (See Supplementary Figure S1), which now integrates the search and selection processes for data sources including Web of Science and Dryad repositories, into a single, cohesive figure as recommended.

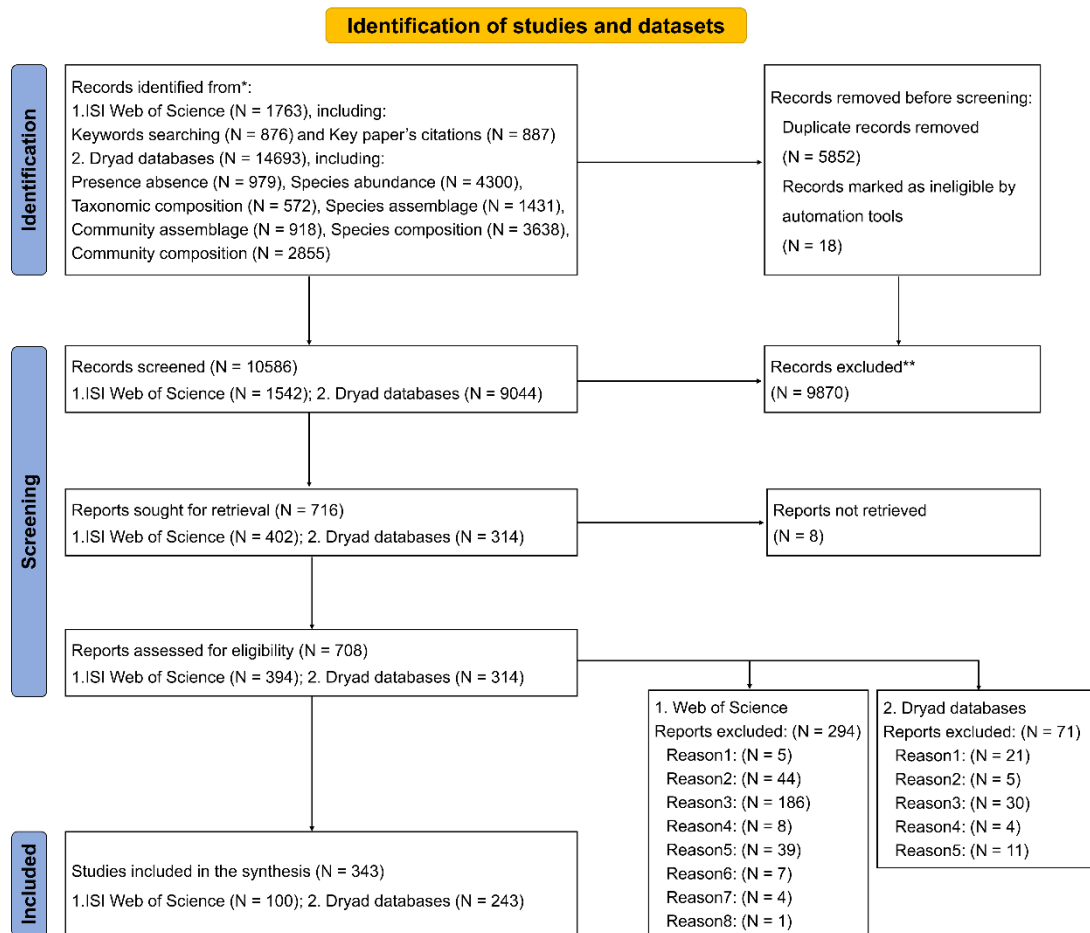


Figure S1 PRISMA diagram showing the process of locating Peer-reviewed primary studies from ISI Web of Science and datasets from Dryad repository included in this synthesis. N: number of studies.

Further, given that correlations can originate from three sources (e.g., primary literary, custom databases, etc.), a sensitivity analysis is necessary to assess whether they represent a homogeneous group and that the data can be combined under a single analysis.

Response: Thank you for this suggestion, we have assessed whether the overall effect sizes of RUR differed significantly across different data sources (i.e. primary studies, data repositories and compiled datasets). **The results demonstrate that the overall effect sizes of RUR were consistently and significantly negative across all three data sources, and no significant differences were found among them either for presence absence or abundance data types** (See Supplementary Figure S3 and Table S7a-b). This indicates that the heterogeneity originating from different data sources is low, thereby justifying their combination into a single dataset for our overall analysis. We have now highlighted this point in the Results section (Lines 127-131).

Table S7b Differences in effect sizes of RUR among data sources. Non-significant

P-values ($P > 0.05$) suggest no differences between the data sources in their RUR effect sizes.

Pair-wise comparison	<i>P</i> -value for presence absence data	<i>P</i> -value for abundance data
Primary studies v.s. Compiled datasets	0.70	0.98
Primary studies v.s. Data repositories	0.072	0.48
Compiled datasets v.s. Data repositories	0.19	0.40

The first section of the results is essentially a type of vote-counting where they enumerate the number of significant and non-significant studies (e.g., lines 120 to 122, or parts of Fig. 2). Don't do this, this undermines the whole purpose of doing a meta-analysis—which carefully does not use significance of each study but rather the magnitude, direction, and sampling error to make inferences. It's ok to report the composition of the dataset (e.g., number of studies per taxa group, or among categories of responses relative to richness), but not differences in significance.

Response: Thank you for raising this critical methodological point. We agree that vote-counting based on significance undermines the purpose of a meta-analysis. In this version, we have thoroughly revised the first section of the Results to eliminate this approach and drop the proportion of different RUR types (positive or negative). Instead, we mainly reported the overall effect size, confidence interval and significance in this revision (Lines 123-127).

Furthermore, regarding the subgroup analyses across different taxa, we only reported the effect sizes and their significance for several groups as you suggested (Lines 132-139; See also Table S9 for details). However, our primary intent is not to directly compare whether the effect size of one group is significantly higher or lower than that of another. Rather, **the key message we aim to convey is that the overall negative RUR pattern was found to be widespread and prevalent across diverse taxonomic groups.**

The rank order of effects to emphasize representation of negative effects is ok, but discussing their prevalence without the lens of meta-analysis is dangerous as many studies may be underpowered and visually may not be representative of the underlying (aggregate) effect. Having correlations as Fisher-Z transformed effects is better for clarity since results are analyzed with the metric (e.g., main inferences are derived from these) and consistence helps with understanding what was done in the paper.

Response: Thank you for this comment. We agree that presenting results through the lens of meta-analysis is crucial for accurate interpretation. Accordingly, we have revised our figures to align with this approach. Specifically, we have replaced the distribution plots of raw correlation coefficients in Figure 2b with distributions based on Fisher's Z transformed effect sizes. Furthermore, Figure 2a has also been updated to present the overall effect sizes and confidence intervals for both presence absence and abundance data.

For the supplementary figures and tables showing the distribution of RUR across different data sources and taxonomic groups (Supplementary Figure S3,

Figure S4, Table S9), we have also converted the data to Fisher's Z transformed effect sizes for consistency and clarity across all visualizations and presentations.

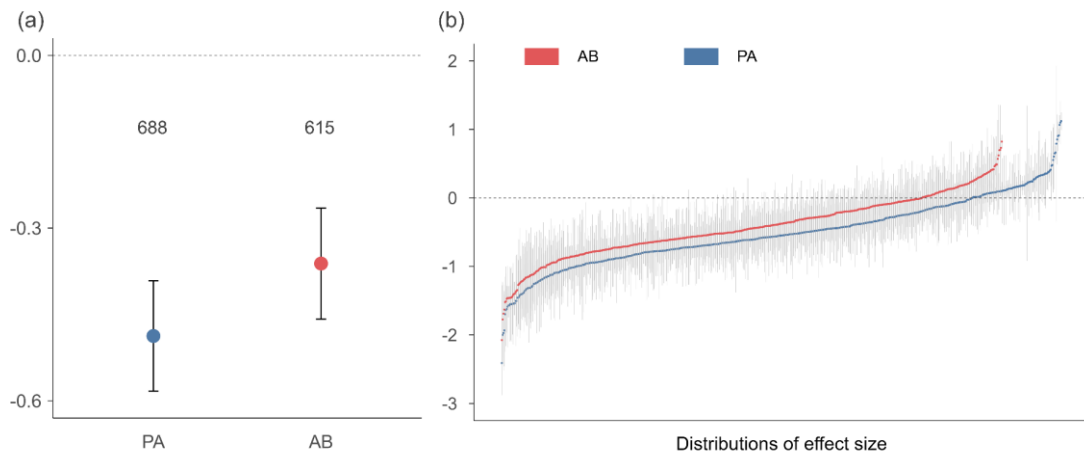


Figure 2 (a) The overall effect sizes of RUR (Fisher's Z statistics) with 95% confidence intervals across different studies and taxa groups based on presence absence and abundance data types. (b) The global variation of RUR across different studies and taxa groups based on presence absence and abundance data types. Number in the left panel indicate number of observations. Details of model output can be found in Table S5.

The results also need to define the acronyms (PA, AB) and more descriptive words should be used for CV_richness, Clim_PC1, and Clim_PC2. Since some details about the methodology are embedded outside the main manuscript, it's important that they are defined so that the reader knows right-way on how to interpret what was reported. These include the above acronyms, but also measure of % explained (line 176 but explained using different definitions in line 568).

Response: Thank you for this suggestion. We have revised the Results section accordingly by replacing the acronyms with more descriptive terms. Specifically, we have replaced "PA" with "Presence absence" and "AB" with "Abundance" in all figures of the main text (See Figures 1-5 in the main text). Also, we have explicitly stated their definitions at the beginning of the Results section (Line 125). Moreover, "CV_richness" has been changed to "Breadth of richness" (Lines 150-154), and "Clim_PC1" and "Clim_PC2" have been replaced with "PC1 of climatic variables" and "PC2 of climatic variables," respectively (Lines 174-186). We have also added explanatory captions to the figures to clarify the meaning of PC1 and PC2 (See Figures 3-5 in the main text). These changes will help readers better understand the meaning of these factors. Additionally, in response to your point on the percentage of variance explained, we have clarified in the Results that this value represents the marginal R squared (as defined in Lines 625-626), which denotes the variance explained by the fixed effects in the mixed-effects models (Lines 188-189; Line 194).

Many of the results reported are described as “affects”, for example this “affected” that. If a moderator is being used to explain variation among effects, maybe “predict”

is a better inferential description, since correlations are just associations and not causes, and similarly, moderators are predictors and not causes of variability.

Response: Following your suggestion, we have revised the wording throughout the Results section, replacing causal terms such as "affects" with more appropriate inferential words like "predicts" or "association". (Lines 142-186)

Finally, lines 517-518 state that significance was tested by 95% overlap, yet there are only p-values reported throughout the results; what is the test statistics these p-values, and their values with degrees of freedom need to be reported along with p-values.

Response: Thank you for pointing this out. In addition to the *P*-values, we have now reported the Q_M statistic and its degrees of freedom for model regression coefficients in the Results section (Lines 146-186). Furthermore, we have included this statistical information directly in Figures 3-5 in the main text. Other statistics, such as the 95% confidence intervals and the total heterogeneity of effect sizes (i.e., I^2), can be found in the Supplementary Tables S5-S18.

Minor comments:

1) What is the scale of the taxonomic group? Are they Linnean rankings, or homologous groups?

Response: Thank you for raising this question. **Our classification system is based on Linnean rankings.** To ensure mutual exclusivity between groups, we have systematically categorized each taxon according to specific taxonomic ranks (such as Phylum, Class, or Order). The table below (Table R1) provides detailed information on the specific finer groups included within each taxa group (where applicable). If insisted, we can of course give this table also as a supplement in the manuscript.

Table R1 Details of taxonomic groups in the synthesis

Taxa group	Finer groups
Algae	Diatoms
	Benthic diatoms
	Studies only reported "Algae"
	Laminariales (Kelp)
Amphibians	Frogs
	Anurans
	Tadpoles
	Studies only reported "Amphibians"
Aquatic plants	Macrophyte
	Marine Plants
	Studies only reported "Aquatic plants"
	Wetland plants
Arthropods	Spiders
	Crabs
	Scorpions
	Decapods

	Peracarida
	Studies only reported “Arthropods”
Bats	No finer groups in the studies
Birds	No finer groups in the studies
Bivalves	Mussels
	Studies only reported “Bivalves”
Fishes	No finer groups in the studies
Freshwater macroinvertebrates	Aquatic insects
	Studies only reported “Freshwater macroinvertebrates”
	Benthic macroinvertebrates in freshwater ecosystem
Gastropods	No finer groups in the studies
Invertebrates	Intertidal invertebrates
	Marine Invertebrates
	Insular invertebrates
	Earthworms
Lichens	No finer groups in the studies
Mammals	No finer groups in the studies
Nematodes	No finer groups in the studies
Phytoplankton	No finer groups in the studies
Reptilia	Snake
	Lizard
Seaweed	No finer groups in the studies
Terrestrial insects	Bees
	Wasps
	Fireflies
	Beetles
	Bombus
	Ants
	Moths
	Staphylinidae
	Studies only reported “insects” in terrestrial ecosystem
	Orthoptera
	Grasshopper
	Hoverflies
	Flies
	Butterflies
Terrestrial plants	Terrestrial vascular plants
	Studies only reported “plants” in terrestrial ecosystem
	Mosses
	Bryophytes
Zooplankton	No finer groups in the studies

2) Line 194, also 403: “literature studies” should be “primary studies”

Response: We have revised “literature studies” into “primary studies” throughout the manuscript.

3) *Line 409: Define what a “unit” is.*

Response: Thank you for pointing this out. In this context, a "unit" refers to the number of sampling points or observation sites in each case study. We have clarified this definition in the revised manuscript. (Line 437)

4) *Line 439: What do you mean by “previous researches”?*

Response: Thank you for pointing this out. The term "previous researches" refers specifically to **all existing studies that have reported on the relationship between species richness and ecological uniqueness**. We synthesized these studies to better determine the specific form of this relationship (e.g., whether it is linear or quadratic). We have now clarified this point in the revised manuscript. (Lines 467-468)

Reviewer #4 (Remarks to the Author):

This manuscript aims to describe the relationship between, and drivers of, species richness and ecological uniqueness. Generally, the manuscript reads very well and the comments of reviewer 1 and 3 are well addressed.

Response: Thank you for this positive assessment of our work.

However, when the quadratic models are tested (as suggested by reviewer 1) and found more appropriate compared to the linear models in 30% of the cases, these records are not considered as the results are harder to interpret. I think that the better fit of the quadratic relationship in such a high percentage of the records show that the relationship can not always be interpreted as linear and that these cases should be considered and not disregarded as the results are harder to interpret. In that way you are only using the results which are convenient, while this is not a good representation of what the data actually shows.

Response: Thank you for the suggestion. As you critically pointed out, these nonlinear relationships represent distinct patterns of variation between ecological uniqueness and species richness. They are equally important and have been repeatedly observed in previous studies (See Appendix 1). In this revision, we have identified different types for the quadratic relationships via performing a qualitative analysis of the nonlinear forms of RUR (i.e., those cases in which quadratic relationships exhibited better performance than linear), following the procedure below:

First, we fitted quadratic models to describe the relationships between ecological uniqueness and species richness in these cases. Then, we calculated the axis of symmetry for each quadratic model and quantified the number of data points distributed on each side of the axis. Based on these patterns, the relationships were categorized as follows:

- (1) For models with R^2 greater than 0.1, if all data points were distributed entirely on the left or right side of the axis of symmetry, the relationship is classified as either a "negative association" or "positive association" (i.e., ecological uniqueness decreases or increases with increasing species richness; Fig. S2a-d).
- (2) For models with R^2 greater than 0.1, if more than 80% of the data points were located on one side of the axis of symmetry, the relationship is classified as a "general negative association" or "general positive association", indicating a predominant but not strictly monotonic trend (Fig. S2e-h). We also tested alternative thresholds (e.g., 70% and 90%) and found the results to be similar, and 80% was accordingly selected as the distribution threshold.
- (3) When fewer than 80% of the data points were located on one side of the axis of symmetry, the relationship was classified as "no general association", indicating that both species-rich and species-poor sites could exhibit high ecological uniqueness.
- (4) If the model R^2 is less than 0.1, regardless of the distribution of points relative to the axis of symmetry, the relationship is also classified as "no general

association". (Lines 501-521)

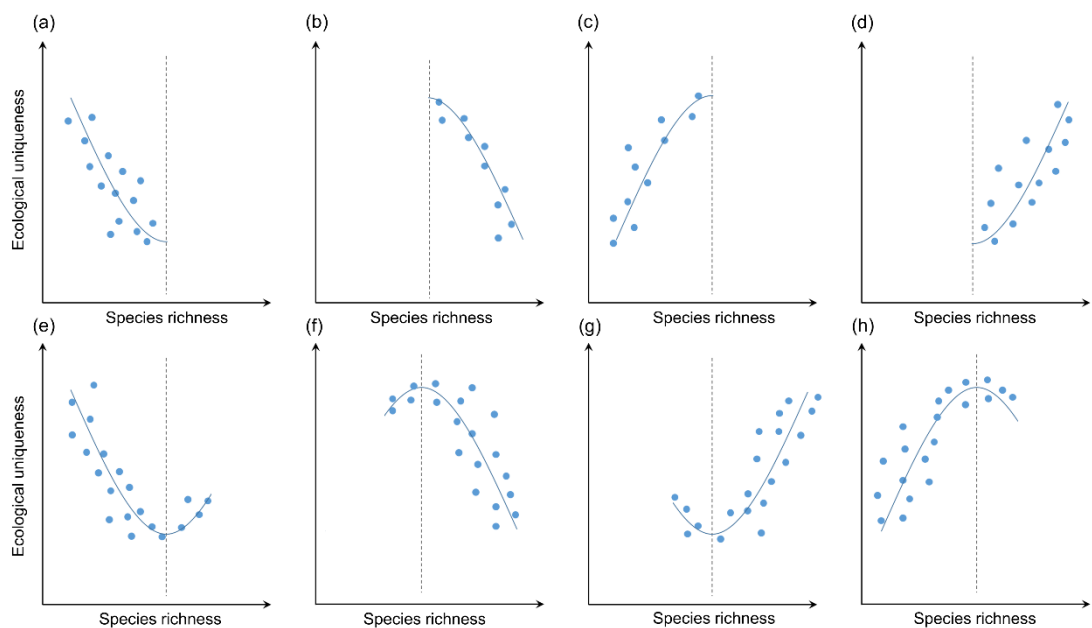


Figure S2 Schematic illustration of classification based on the quadratic richness-uniqueness relationship (RUR). (a, b) Negative association: all data points were distributed on the same side of the axis of symmetry, and ecological uniqueness decreases with increasing species richness. (c, d) Positive association: all data points were distributed on the same side of the axis of symmetry, and ecological uniqueness increases with increasing species richness. (e, f) General negative association: over 80% of data points were distributed on the same side of the axis of symmetry, and ecological uniqueness decreases with increasing species richness in the same side of the axis of symmetry. (g, h) General positive association: over 80% of data points were distributed on the same side of the axis of symmetry, and ecological uniqueness increases with increasing species richness. See Appendix 3 for details.

We found that 22.8% (140 out of 615 records) of the quadratic relationships showed "negative association", which indicates that ecological uniqueness decreased as species richness increased (Appendix 3). 20% (122 out of 615 records) of the observations showed "general negative association", which means that we tend to observe higher ecological uniqueness in species-poor sites although the relationship is not strictly monotonous. In contrast, fewer than 1% (3 out of 615) showed "positive association" and only 3% showed "general positive association", suggesting that ecological uniqueness was generally higher in species-rich sites. These results highlighted that: although RUR can be linear or quadratic depend on context, those quadratic patterns also reflect a predominant tendency for high ecological uniqueness to coincide with low species richness, reinforcing the widely observed negative association between ecological uniqueness and species richness (Lines 220-228).

The results of this qualitative analysis are presented in Appendix 3 and Supporting data, and their implications are discussed in the revised discussion (Lines 228-232). However, we have to acknowledge that our analysis integrates

datasets spanning diverse biodiversity and environmental gradients, making systematic classification challenging. Nevertheless, these cases collectively underscore the point that species richness and ecological uniqueness represent complementary aspects of biodiversity.

Additionally, different drivers of the RURs are tested, following four different hypotheses. In the methods, I recommend a more thorough description of why and how well the variables chosen reflect these particular mechanisms underlying the proposed hypotheses.

Response: Thank you for the suggestion. We have added an explanation in the Supplementary Text regarding why (and how well) these specific factors were selected as representative variables for each hypothesis (See Supplementary Text 6: *Quantifications of different hypotheses*) as follows:

Species pool attributes: These three metrics (Gamma diversity, breadth of richness and proportion of rare species) provide a multi-dimensional quantification of the *species pool hypothesis*. Gamma diversity directly measures the size of the regional species pool, representing the total source of potential colonist species available for community assembly (Lessard et al., 2012). Breadth of richness specifically reflects the variability of richness within given region - a higher value indicates greater richness variation across sampling units, thereby illustrating how the species pool maintains different community configurations across varying local habitats (Dansereau et al., 2022). The proportion of rare species reveals the pool's internal structure, with a higher proportion suggesting lower number of shared species or higher species turnover among communities (Yao et al., 2021).

Dispersal limitation: Distance-based Moran's Eigenvector Maps (dbMEM) were employed to infer *dispersal limitation* because they effectively detect spatial autocorrelation patterns characteristic of constrained species movement. Dispersal limitation produces a distance-decay relationship in community similarity, where nearby sites tend to share more species (positive spatial autocorrelation). dbMEM variables capture these underlying spatial structures at different scales, from broad-scale gradients to fine-scale patches. When these variables explain a significant portion of the variation in community composition, the most parsimonious interpretation is that dispersal limitation plays a dominant role in shaping communities.

Environmental filtering: We used 19 bioclimatic variables to represent *environmental filtering*, mainly because of their direct quantification of the critical climatic conditions that determine species' survival and distribution. The species-energy hypothesis posits that climate strongly influences primary productivity, or the total energy available for partitioning within most ecological communities, thereby affecting species richness or species composition via its impact on the number and size of organisms that can be supported (Hutchinson 1959; Currie et al., 2004; Chu et al., 2016; Storch et al., 2018). These variables such as the minimum temperature of the coldest month (Bio6) and precipitation of the driest quarter (Bio17) capture the critical extremes of temperature and water availability

that often limit species persistence.

Spatial grain size: Spatial grain size not only directly determines the number of species observed and the spatial turnover in species composition, but also potentially alters the influence of other ecological processes such as biotic interactions and environmental heterogeneity. Larger grains tend to integrate greater habitat heterogeneity and buffer fine-scale stochasticity, while smaller grains better capture local biotic and environmental filtering effects.

References:

- Chu, C. et al. Direct effects dominate responses to climate perturbations in grassland plant communities. *Nat. Commun.* **7**, 11766 (2016).
- Currie, D. J. et al. Predictions and tests of climate-based hypotheses of broad-scale variation in taxonomic richness. *Ecol. Lett.* **7**, 1121–1134 (2004).
- Dansereau, G., Legendre, P. & Poisot, T. Evaluating ecological uniqueness over broad spatial extents using species distribution modelling. *Oikos*, e09063 (2022).
- Hutchinson, G. E. Homage to Santa Rosalia or why are there so many kinds of animals? *Am. Nat.* **93**, 145–159 (1959).
- Lessard, J-P., Belmaker, J., Myers, J. A., Chase, J. M. & Rahbek, C. Inferring local ecological processes amid species pool influences. *Trends Ecol. Evol.* **27**, 600–607 (2012).
- Storch, D., Bohdalková, E. & Okie, J. The more-individuals hypothesis revisited: the role of community abundance in species richness regulation and the productivity–diversity relationship. *Ecol. Lett.* **21**, 920–937 (2018).
- Yao, J., Huang, J., Ding, Y., Xu, Y., Xu, H. & Zang R. Ecological uniqueness of species assemblages and their determinants in forest communities. *Divers. Distrib.* **27**, 454–462 (2021).

Generally, I strongly recommend to first run a fully randomised null model, to check if the obtained relationships, and therefore your key results, are obtained by chance or that ecological mechanisms drive this relationship.

Response: Thank you for this insightful suggestion. We conducted a null model analysis to determine whether the observed RUR was statistically distinct from patterns expected by chance, thereby indicating the influence of underlying ecological processes. For presence absence data, the null model randomized species co-occurrence while preserving species incidence frequencies, site richness, and total occurrences (Gotelli, 2000). For abundance data, it maintained both site-level and species-level total abundances, randomizing only the assignment of individuals to sites (Kraft et al., 2011; Xing & He, 2019). The details of null model analysis can be found in Supplementary Text 6: *Null model analysis*.

We generated 999 randomized communities to construct the null distribution of RUR. Statistical significance was evaluated using a two-tailed test, where the observed RUR was compared against the 2.5th and 97.5th percentiles of the null distribution. **A P -value < 0.05 signified that the observed pattern was significantly different from random expectation, suggesting the influence of**

non-random ecological processes such as dispersal limitation or environmental filtering. We have provided the results (*P*-value, 2.5th and 97.5th percentiles of null distributions) of this analysis in the Supporting data.

In total, **the null model analysis revealed that a total of 645 records (290 of presence absence data and 355 of abundance data) showed significant deviations from the null expectation** (See Supporting data for details). **Therefore, our subsequent analyses examining the predictions of different hypotheses on RUR and their relative importance were conducted exclusively on this subset of records (k = 645).**

References:

- Gotelli, N. J. Null model analysis of species co-occurrence patterns. *Ecology* **81**, 2606–2621 (2000).
- Kraft, N. J. B. et al. Disentangling the drivers of β diversity along latitudinal and elevational gradients. *Science* **333**, 1755–1758 (2011).
- Xing, D. L. & He, F. L. Environmental filtering explains a U-shape latitudinal pattern in regional β -diversity for eastern North American trees. *Ecol. Lett.* **22**, 284–291 (2019).

Lastly, from a traits perspective, beta diversity can be different between sites, but some sites are more ecologically unique than others as the niche and traits of the species differ more compared to the other sites. This is a point which should be considered in the introduction and discussion.

Response: Thank you for pointing out it. Following your suggestion, we have highlighted the role of traits and niche differences in the introduction and discussions:

In the Introduction: “Moreover, sites exhibiting high ecological uniqueness may also contribute disproportionately to regional biodiversity by enhancing trait divergence and niche differentiation of specific species assemblages.” (Lines 31–34)

In the discussions: “Finally, while our findings highlight the conservation relevance of RUR, taxonomic measures alone may not fully capture the multidimensional nature of biodiversity. Variation in ecological uniqueness across sites can also be further interpreted through the lens of functional traits and niche differentiation (Violle et al., 2007; Violle et al., 2017; Heino et al., 2022; Ascanio et al., 2024). For instance, sites with higher ecological uniqueness likely harbor species assemblages that are distinct not only in taxonomic identity but also in their functional characteristics and fundamental ecological niches (Heino et al., 2022; Casanelles-Abella et al., 2023). These communities may display greater divergence in key niche-related traits and occupy underrepresented or extreme regions of the multivariate trait space. Critically, the observed uniqueness may reflect substantial differences in the ecological niches of the constituent species, which could vary greatly in resource use strategies, environmental tolerances, or other dimensions (Chesson, 2000; Ascanio et al., 2024). This perspective underscores that the

ecological uniqueness is intrinsically linked to both functional rarity and niche complementarity, highlighting the need to integrate trait-based and niche-based approaches with biodiversity partitioning to better understand the mechanisms underlying spatial heterogeneity in ecological communities.” (Lines 393-407)

References:

- Ascanio, A., Bracken, J. T., Stevens, M. H. H. & Jezkova, T. New theoretical and analytical framework for quantifying and classifying ecological niche differentiation. *Ecol. Monogr.* **94**, e1622 (2024).
- Casanelles-Abella, J., Fontana, S., Meier, E., Moretti, M. & Fournier, B. Spatial mismatch between wild bee diversity hotspots and protected areas. *Conserv. Biol.* **37**, e14082 (2023).
- Chesson, P. Mechanisms of Maintenance of Species Diversity. *Annu. Rev. Ecol. Evol. Syst.* **31**, 343-366 (2000).
- Heino, J. et al. Assessing the conservation priority of freshwater lake sites based on taxonomic, functional and environmental uniqueness. *Divers. Distrib.* **28**, 1966–1978 (2022).
- Violle, C. et al. Functional rarity: The ecology of outliers. *Trends Ecol. Evol.* **32**, 356–367 (2017).
- Violle, C. et al. Let the concept of trait be functional! *Oikos* **116**, 882–892 (2007).

Additional comment sent to the editor:

Elaborate on the relationship in Figure 3C and if they can test how this relationship is, or is not, driven by the outliers at the right side of the graph.

Response: Thank you for this careful comment, we have rigorously tested the relationship in Figure 3c. We identified and removed outliers falling outside the range of two standard deviations from the mean of breadth of richness and subsequently re-ran the meta-regression. The results indicate that a positive trend persists although the relationship is no longer statistically significant after excluding these outliers (See Figure R1 below). This indicates that the originally significant relationship was largely driven by a few outlier records. However, this finding also suggests that if future studies encompass a broader range of communities with higher breadth of richness, the relationship could become more robust and potentially reveal a clearer transition from negative to positive associations. We have incorporated this interpretation and its implications into the revised Discussions (Lines 278-286).

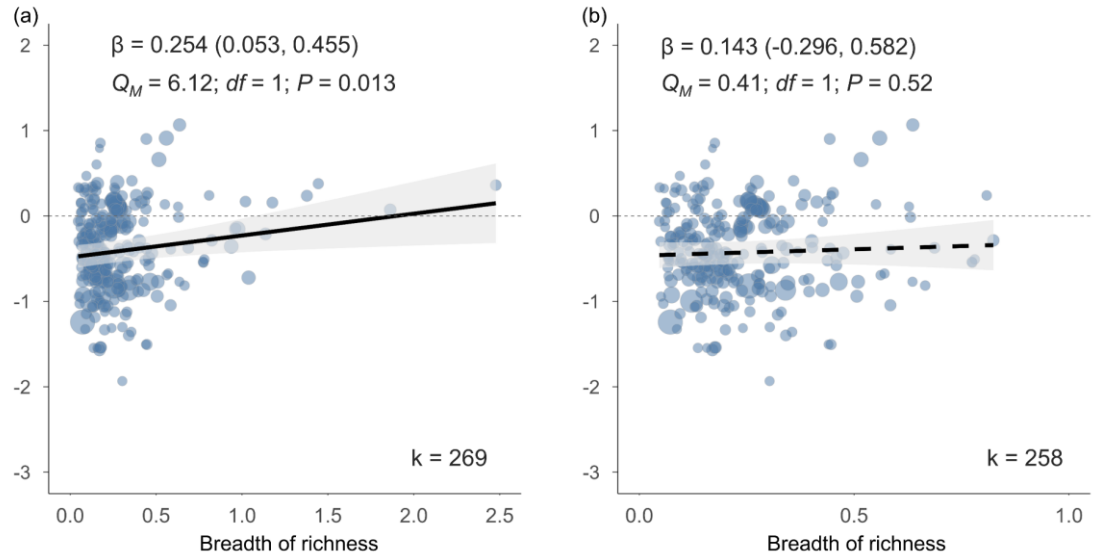


Figure R1 The association between effect sizes of RUR and breadth of richness (a) before and (b) after outlier removal. The regression lines represent the response to the above factors with the 95% confidence intervals shaded from hierarchical mixed-effects meta-regressions. The values of the estimated coefficients (β), 95% confidence intervals, Q_M statistics and degree of freedom (df) are shown in each panel. The solid and dashed lines represent significant ($P < 0.05$) and non-significant ($P > 0.05$) relationships. The points have radii relative to the inverse variance of each observation (i.e., observations with larger points have lower variance in effect size). k : Number of effect sizes.

Response to reviewers

Reviewer #2 (Remarks to the Author)

In terms of what was reported and presented with this meta-analysis, I believe this new version has the clarity and detail that is more in line with current expectations in research synthesis. My comments below are very very minor.

Response: Thank you for this positive assessment of our work.

MAIN TEXT

lines 832, 838, 851: "Fisher's Z statistics" should be "Fisher's Z-transformed correlation coefficients"

Response: As you suggested, we have revised the corresponding figure captions by replacing "Fisher's Z statistics" with "Fisher's Z-transformed correlation coefficients". (See captions of Figure 2-4 in the main text)

lines 864-876: definition of effect sizes (see above) is missing

Response: We have included the definition of effect sizes in the corresponding figure captions. (See captions of Figure 5 in the main text)

SUPPLEMENTARY TEXT 1:

line 11: where was this search done? E.g., using who's institutional access to ISI Web of science?

Response: We have clarified the search methodology by adding the following detail in the manuscript: "The search was conducted using the Web of Science Core Collection database, accessed via the institutional subscription of the University of Chinese Academy of Sciences." (See Supplementary Note 1 for details)

line 32: Which PRISMA statement did you use? cite source

Response: We have specified the version of the PRISMA statement used in our study and added the corresponding citation in the supplementary information (See Supplementary Note 1 for details). We conducted this systematic review and meta-analysis in accordance with the **Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 statement** (Page et al., 2021).

References:

Page, M. J. et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. **372**, n71 (2021).

line 3863, 3872, 3879, 3890, 3898, 3906, 3978, 3990, 4005, 4018, 4030, 4033, 4037, 4054, 4070, 4086, 4102, 4118, 4134, 4150, 4166, 4182: "Fisher's Z statistics" should be "Fisher's Z-transformed correlation coefficients"

Response: We have replaced "Fisher's Z statistics" with "Fisher's Z-transformed correlation coefficients" as you recommended. (See captions of Supplementary Figure 3-9, Supplementary Table 3-18)

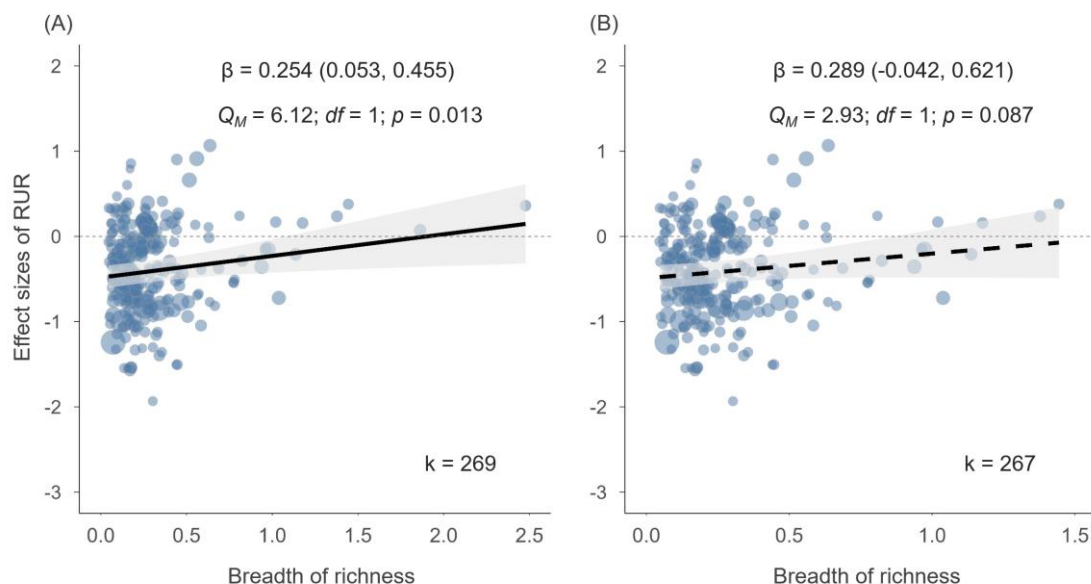
Reviewer #4 (Remarks to the Author)

I would like to thank the authors for the thorough answers and additional analyses performed. The comments are well addressed.

Response: Thank you for this positive assessment of our work.

The only question which remains for me is about the outliers in Figure 3C. Are the data points at the right side of the figure indeed checked and found to be valid or do these data points represent less robust observations? In the latter case I would disregard these data points and change the text accordingly, otherwise all good.

Response: Thank you for the thoughtful suggestion. We have carefully examined the two data points on the right side of figure 3C, particularly regarding the calculation of the coefficient of variation (CV) of species richness. After verification, we confirm that these CV values were correctly calculated. Therefore, we have retained these data points in the analysis and figure. To further assess the robustness of the relationship, we have also presented the results after removing these outliers in the supplementary information (See Supplementary Figure 9). The overall pattern remains consistent, supporting the stability of our conclusion. Furthermore, we have expanded the Discussion to emphasize the potential influence of these extreme values on the results. We note that as future studies encompass a broader range of communities with greater breadth of richness, the estimated relationship may become more robust and potentially reveal a clearer transition from negative to positive associations (Lines 301-307).



Supplementary Figure 9 The association between effect sizes of RUR (Fisher's Z-transformed correlation coefficients) and breadth of richness (A) before and (B) after outlier removal. The regression lines represent the response to the above factors, with their 95% confidence intervals shown as shaded areas from hierarchical mixed-effects meta-regressions. The values of the estimated coefficients (β), 95% confidence intervals, Q_M statistics (one-sided chi-square test)

and degree of freedom (df) are shown in each panel. The solid and dashed lines represent significant ($p < 0.05$) and non-significant ($p > 0.05$) relationships. The points have radii relative to the inverse variance of each observation (i.e., observations with larger points have lower variance in effect size). k : Number of effect sizes.

Richness

