How much biodiversity is concealed in the word "biodiversity"?

Stefano Mammola[*,§], Caroline S. Fukushima, Girolama Biondo, Lucia Bongiorni, Fabio Cianferoni, Paolo Domenici, Carmelo Fruciano, Angelina Lo Giudice, Nuria Macías-Hernández, Jagoba Malumbres-Olarte, Marija Miličić, Michelangelo Morganti, Emiliano Mori, Ana Munévar, Paola Pollegioni, Ilaria Rosati, Simone Tenan, Fernando Urbano-Tenorio, Diego Fontaneto[†], Pedro Cardoso[†]

*Correspondence: stefano.mammola@helsinki.fi; stefano.mammola@cnr.it;

MAIN TEXT

Amidst a global biodiversity crisis[1], the word "biodiversity" has become indispensable for conservation and management[2]. Yet, biodiversity is often used as a buzzword in scientific literature. Resonant titles of papers claiming to have studied "global biodiversity" may be used to promote research focused on a few taxonomic groups, regions, habitats, or facets of biodiversity [taxonomic, (phylo)genetic, or functional]. This usage may lead to extrapolating results outside the target systems of these studies with direct consequences for our understanding of life on Earth and its practical conservation. Here, we used a random sample of papers with the word "biodiversity" in their title to take a long view of the use of this term. We analyzed the degree to which these studies consider different taxonomic groups and biodiversity facets and how this affects the impact of a paper. Despite improvements in analytical tools, monitoring technologies,

and data availability[3,4], we found that the taxonomic scope of research articles has not increased in recent years. We also show that studies with a wider taxonomic scope attract more citations and online attention. Our results have broad ramifications for understanding how extrapolating from studies with narrow taxonomic scope affects our view of global biodiversity and conservation.

We gathered all the articles listed in the *Web of Science* with the word "Biodiversity" in their title (N = 10,170; Supplemental Experimental Procedures). We randomly sampled ~10% of these papers and extracted detailed information on geographical focus, methodologies, and biodiversity facets considered. Furthermore, we counted the number of unique Phyla/Divisions (or higher taxonomic ranks for microorganisms) considered in each study (hereinafter "phyla"). We then computed for each study the sampled number of phyla out of the total possible phyla ("proportion of biodiversity"; Figure S1).

We found that as many as 22% of the papers using the word "biodiversity" in the title did not measure biodiversity at any level. This suggests that biodiversity is often used as a theoretical concept rather than a measurable phenomenon[2].

Across the remaining 661 papers, the proportion of biodiversity investigated by each study showed a highly skewed distribution, with most studies sampling a small proportion of biodiversity and a long tail of comparatively few studies sampling higher proportions (mean ± S.E.: 3.86% ± 0.15%; mode: 1.78%; range: 1.78–44.64%) (Figure S2). The taxonomic scope of papers has not increased in recent years either (Figure 1A).

Next, we investigated the role of 11 factors in explaining the biodiversity sampled by each paper (Figure 1B). Sampled biodiversity was lower in studies set in the Antarctic, Afrotropical, Indomalayan, and Nearctic regions (Figure S2A) and those focusing on the

2

terrestrial realm (Figure S2B). Low sampled biodiversity was associated with studies based on big data (Figure S2C) or focusing on phylogenetic diversity. The most sampled taxa were vertebrates (Chordata) followed by arthropods, whereas microorganisms and fungi were the least studied (Figure S2D).

A possible explanation for these patterns is that certain taxa and regions are more likely to receive research funds and attention[5]. Some taxa are easier to study due to their characteristics (e.g., macroscopic size, large geographic range, and ease of sampling) and greater availability of data. The finding that research using big data have narrower taxonomic scopes was unexpected. The availability of big data could potentially allow the study of an increasing number of organisms, but in fact, increasing biases in existing databases6 might not enable short-term data synthesis[6]. Much needed data for biodiversity studies await to be collected from the field, existing collections, or even "grey" literature, all requiring massive human effort. Ultimately, it seems that we are flooded by data and analyses on few taxa (e.g., in 2020, vertebrates accounted for 68% of GBIF-available data[7]), increasing biases in inadvertent ways.

Finally, we derived two measures of article impact—number of citations and Altmetric score—and tested how sampled biodiversity and the use of descriptors (mention of taxa, habitat, or locations) in the title affect impact, while accounting for the number of countries of the coauthors and the Impact Factor as confounding factors. In general, not mentioning descriptors led to more citations (Figure 1C) and societal attention (Figure 1D). All else being equal, proportion of biodiversity in interaction with the use of descriptors had a positive effect on impact. Whereas the impact of articles with more than one descriptor in the title was generally low, articles with one or no descriptor in the title attained greater impact when they sampled more biodiversity (Figure 1C, D).

Overall, our results suggest caution when extrapolating from a few taxa, regions, or habitats to the full spectrum of living forms. This practice can misinform and misdirect conservation policies and actions by governments, organizations, and conservation practitioners, misallocating resources[5] and perpetuating known biodiversity shortfalls[6]. In the long run, this may turn out to be detrimental for most species and even the ecosystem services on which we depend. So, we must confront the important question: What can we do to improve this situation?

First, in the current trend of increasing publication numbers, fast communication through social media, and decreasing attention span of readers[8], our results are a powerful reminder that scientists should critically read papers and their scope rather than limiting the focus on the titles and abstracts. Also, as both editors and reviewers, we should play an active role in reducing the dangers related to "overselling", e.g. by calling out manuscripts with unjustified broad titles. While "overselling" may produce short-term positive effects in terms of citations and societal attention, it will not serve the long-term goals of prestige and authoritativeness which any journal should strive for.

Given that in most biodiverse biogeographic regions the sampled proportion of biodiversity is systematically low, researchers and journals from these areas should be supported in producing primary biodiversity data and involved in international collaboration[9]. This would decrease existing taxonomic and geographical biases across all biodiversity facets.

An effort may also be made to expand broad-scale biodiversity databases with a focus on underrepresented taxa and biogeographical regions. As emphasized several times[10], increasing the number of trained taxonomists and funds dedicated to this type of activity will be instrumental to increasing the taxonomic coverage of studies. Indeed, if some parts of

biodiversity research can now largely be automated[4], others build on basic natural history and taxonomic knowledge in the most under-explored regions of the world, which often harbor the vast majority of biodiversity.

Acknowledgments

Supplemental Information

Supplemental Information includes Experimental Procedures, Statements, Data and Code availability, two Figures and can be found with this article online at *bxs.

Reference list

1. Cowie, R.H., Bouchet, P., and Fontaine, B. (2022). The Sixth Mass Extinction: fact, fiction or speculation? Biol. Rev. *97*, 640–663.

2. Díaz, S., and Malhi, Y. (2022). Biodiversity: Concepts, Patterns, Trends, and Perspectives. Annu. Rev. Environ. Resour. *47*, 31–63.

3. Feng, X., Enquist, B.J., Park, D.S., Boyle, B., Breshears, D.D., Gallagher, R. V, Lien, A., Newman, E.A., Burger, J.R., Maitner, B.S., et al. (2022). A review of the heterogeneous landscape of biodiversity databases: Opportunities and challenges for a synthesized biodiversity knowledge base. Glob. Ecol. Biogeogr. *31*, 1242–1260.

4. Tosa, M.I., Dziedzic, E.H., Appel, C.L., Urbina, J., Massey, A., Ruprecht, J., Eriksson, C.E., Dolliver, J.E., Lesmeister, D.B., Betts, M.G., et al. (2021). The rapid rise of next-generation natural history. Front. Ecol. Evol. *9*, 480.

5. Adamo, M., Sousa, R., Wipf, S., Correia, R.A., Lumia, A., Mucciarelli, M., and Mammola, S. (2022). Dimension and impact of biases in funding for species and habitat conservation. Biol. Conserv. *272*, 109636.

6.   Hortal, J., de Bello, F., Diniz-Filho, J.A.F., Lewinsohn, T.M., Lobo, J.M., and Ladle, R.J. (2015). Seven shortfalls that beset large-scale knowledge of biodiversity. Annu. Rev. Ecol. Evol. Syst. *46*, 523–549.

7.   Heberling, Mason, J., Joseph, M.T., Noesgaard, D., Scott, W.B., and Schigel, D. (2021). Data integration enables global biodiversity synthesis. Proc. Natl. Acad. Sci. *118*, e2018093118.

8.   Parolo, P.D.B., Pan, R.K., Ghosh, R., Huberman, B.A., Kaski, K., and Fortunato, S. (2015). Attention decay in science. J. Informetr. *9*, 734–745.

9.   Cardoso, P., Fukushima, C.S., and Mammola, S. (2022). Quantifying the internationalization and representativeness in research. Trends Ecol. Evol. *37*, 725–728.

10.   Engel, M.S., Ceríaco, L.M.P., Daniel, G.M., Dellapé, P.M., Löbl, I., Marinov, M., Reis, R.E., Young, M.T., Dubois, A., Agarwal, I., et al. (2021). The taxonomic impediment: a shortage of taxonomists, not the lack of technical approaches. Zool. J. Linn. Soc. *193*, 381–387.
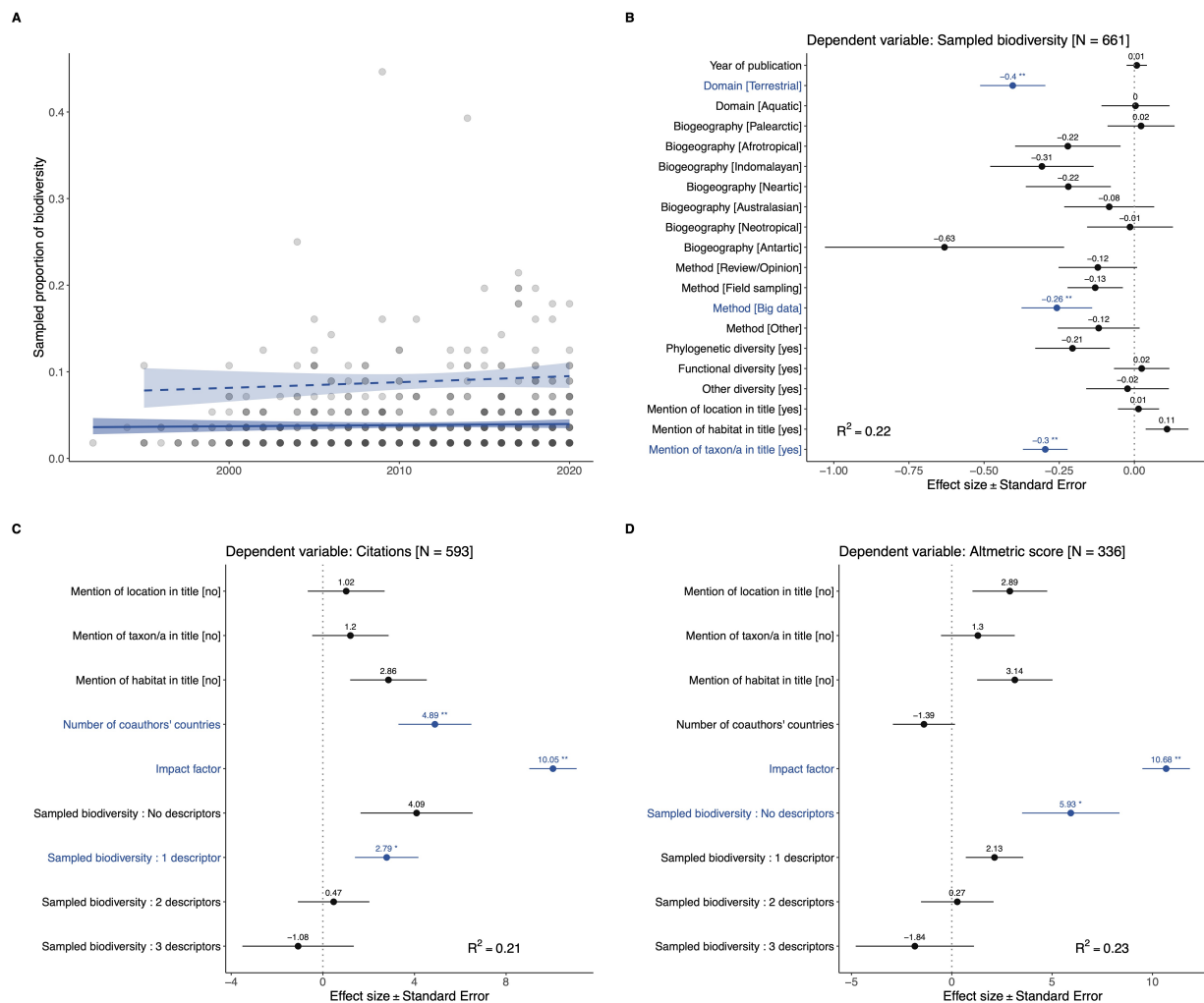
Figure 1. Sampled biodiversity across studies and its relation with their impact. A) Annual variations in the proportion of biodiversity considered in each study. Regression lines: filled, full data (quasibinomial GLM; estimated β ± SE: 0.004 ± 0.007, p = 0.587); dashed, only data in the 75–100th percentile (quasibinomial GLM; estimated β ± SE: 0.008 ± 0.009, p = 0.339). **B)** Estimated parameters for a negative binomial generalized linear model testing the relationship

between sampled biodiversity and different article-level predictors. The model is based on studies with sampled biodiversity > 0. Reference categories: Domain [Multiple]; Biogeography [Global]; Method [Multiple]. **C**) Estimated parameters for a linear model testing the relationship between citations and different article-level predictors, including the interaction between sampled biodiversity and the number of descriptors (i.e. mention of taxa, habitat, or locations). **D**) Visualization of the influence of the interaction between the number of descriptors and sampled biodiversity on Altmetric scores, including the same interaction as in C. In **B–D**, error bars indicate standard errors. Significant values (*: < 0.05; **: < 0.01) are highlighted in blue.

Supplemental Information: How much biodiversity is concealed in the word "biodiversity"?

Stefano Mammola, Caroline S. Fukushima, Girolama Biondo, Lucia Bongiorni, Fabio Cianferoni, Paolo Domenici, Carmelo Fruciano, Angelina Lo Giudice, Nuria Macías-Hernández, Jagoba Malumbres-Olarte, Marija Miličić, Michelangelo Morganti, Emiliano Mori, Ana Munévar, Paola Pollegioni, Ilaria Rosati, Simone Tenan, Fernando Urbano-Tenorio, Diego Fontaneto, Pedro Cardoso
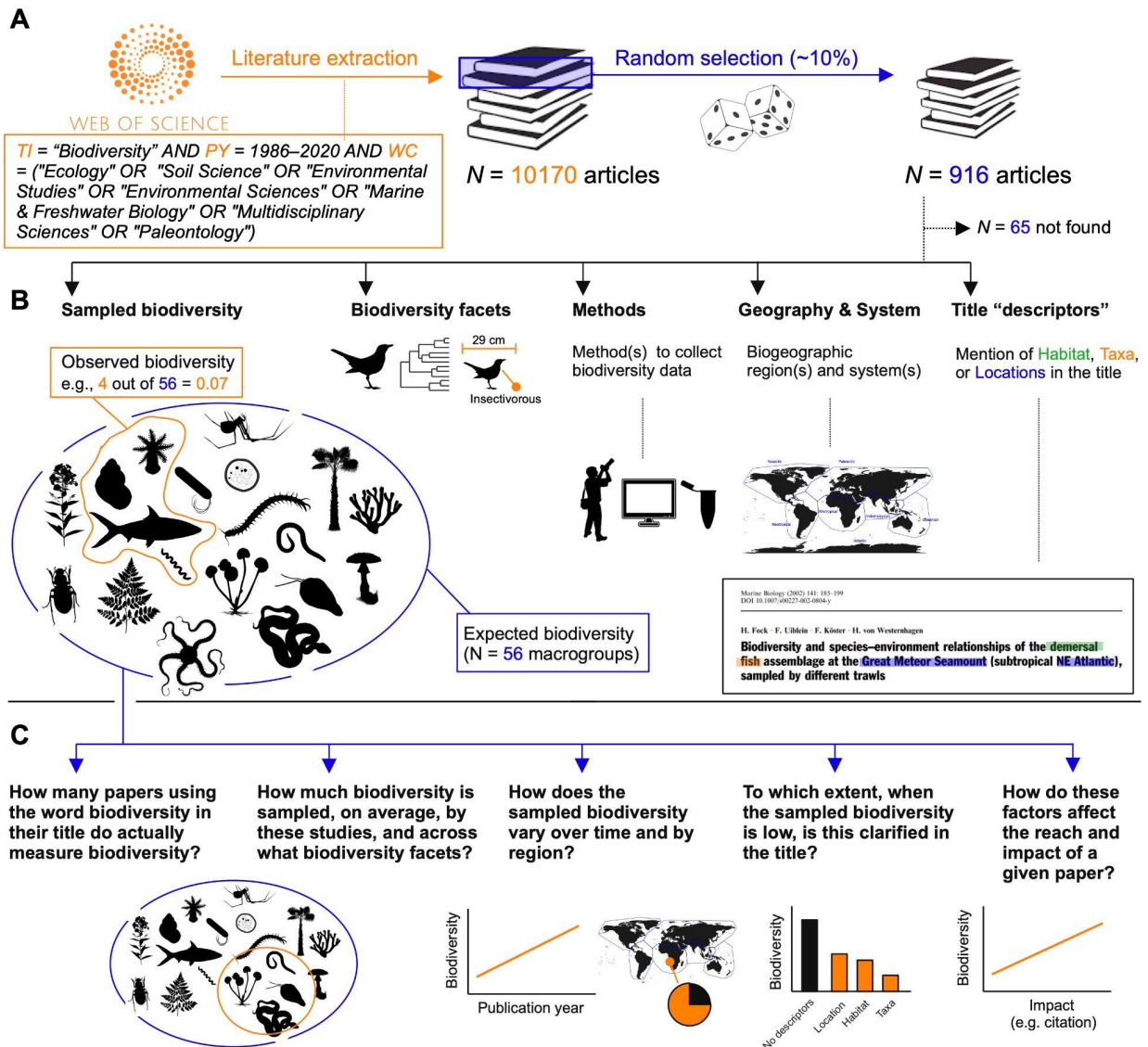
Figure S1. Infographic summarizing the study design. A) Literature sampling; **B)** Summary of the main variables extracted from each paper; **C)** Summary of the research questions and hypotheses.
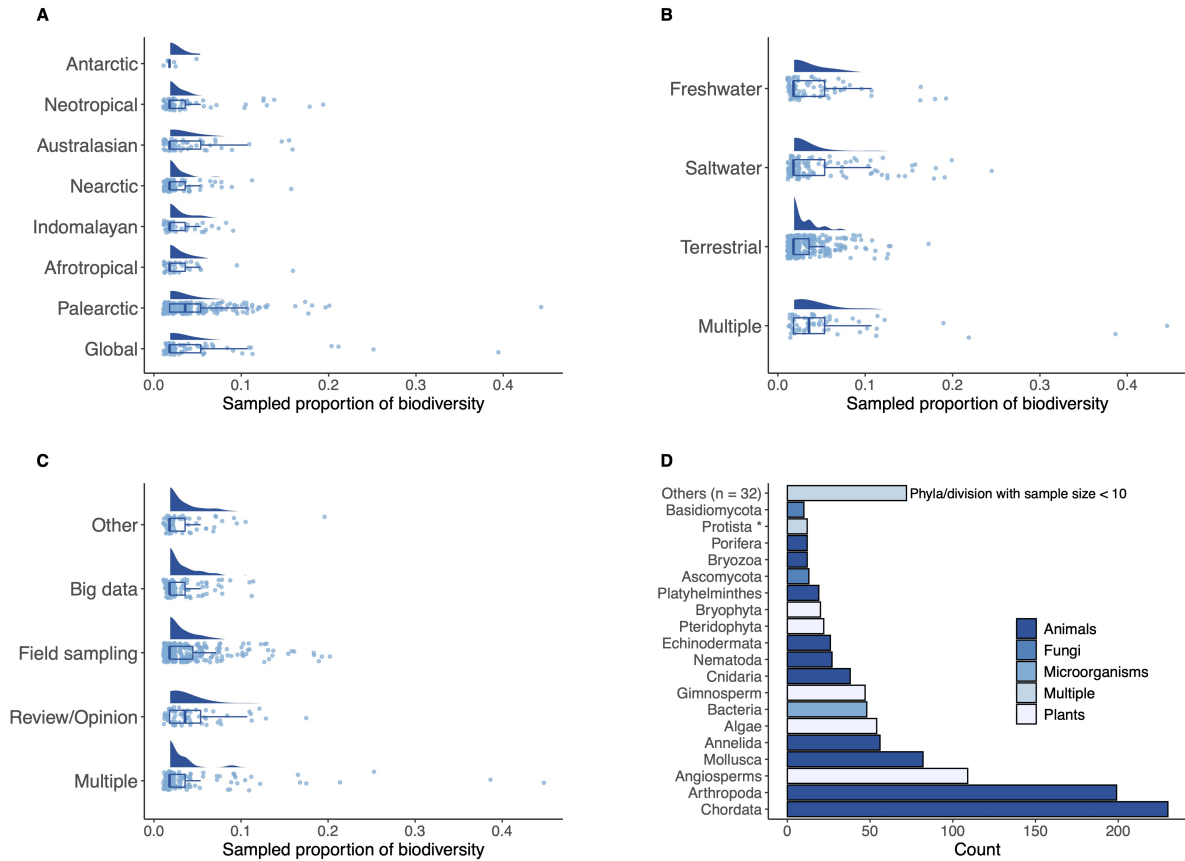
Figure S2. Change in biodiversity proportion across regions, systems, and methodologies. A–C) Breakdown of biodiversity proportion by biogeographic regions, systems, and research methods. Jittered points are the actual values, boxplots summarize median and quantiles, and density plots summarize data distribution. **D)** Most commonly investigated biodiversity groups across studies.

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Research questions and hypotheses

In this study, we took a long view of how researchers have used the term "biodiversity" in scientific literature to understand its meaning and the consequences of its use. Using regression-like analyses, we explored a number of interrelated questions:

i) How many papers using the word biodiversity in their title do actually measure biodiversity?

ii) How much biodiversity is sampled, on average, by these studies?

iii) How does the sampled biodiversity vary over time and by regions?

iv) To which extent, when the sampled biodiversity is low, is this clarified in the title?

v) How do these factors affect the reach and impact of a given paper?

Owing to recent gigantic advances in analytical tools, monitoring technologies, and the availability of biodiversity variables and data[S1–S5], we hypothesized that the biodiversity scope of papers should increase over time and should be higher in studies focusing on more biodiverse regions. If authors are not overselling their results, we also expect that papers with a narrow biodiversity scope should make it explicit in their title by descriptor terms specifying taxa, geographic regions or habitats of focus. Finally, we expected to observe a direct relationship between biodiversity scope and article impact (Figure S1C).

Data collection

On 22 May 2021, we queried the *Web of Science* core collection database for articles using the word "biodiversity" in their titles. We restricted the search to titles only given that they are the

4

"hook" to readers[S6,S7], the element of a paper that is most commonly assessed by scientists while screening for relevant papers. Indeed, it is estimated that a researcher, on average, skims 1,100 titles a year but will then go on to read 97 full texts only[8].

We used the search string *TI = "Biodiversity" AND DT = "Article" AND WC = "Ecology" OR "Soil Science" OR "Environmental Studies" OR "Environmental Sciences" OR "Marine & Freshwater Biology" OR "Multidisciplinary Sciences" OR "Paleontology" AND PY = 1986–2020*. Note that we restricted the search to general Web of Science categories (WC) pertaining to biodiversity, avoiding taxon-specific categories (e.g., "Entomology", "Fisheries", "Ornithology") which would have biased the search toward articles dealing with restricted samples of organisms. We selected the year 1986 as a lower boundary for the search because the term "BioDiversity" was coined in 1986 by Walter G. Rosen during the organization of the "*National Forum on BioDiversity*" (Washington, D.C).

The initial search yielded 10,170 hits. From this database, we randomly sampled 916 articles to be analyzed. Note, however, that we managed to extract metadata for 851 full texts—we could not find and/or download 65 articles. Hence, in all analyses, the final sample size is 851.

Metadata extraction

We inspected the full text of the sampled articles to extract the relevant data for our analyses. Note that we could not access the full-text for 65 articles, resulting in a final sample size of 851.

For each study, we first scored the year of publication, the year(s) the study occurred, and the biogeographic region ("Global", "Nearctic", "Neotropical", "Afrotropical", "Palearctic", "Indomalayan", "Oceanian", "Australasian", and/or "Antarctic") and ecological domain ("Terrestrial", "Saltwater", and/or "Freshwater") of focus. We also noted the approach(es) taken by the authors to study biodiversity ["Field sampling" (data collected in the field), "Big data" (use of pre-collected data, e.g., from online databases such as GBIF), "Review/Opinion" (theoretical studies or reviews), and "Other" (none of the previous)]. A single study may include multiple biogeographic regions, domains, and methods.

For each study's title, we marked ("yes" or "no") whether, alongside the word "biodiversity", it mentioned: i) taxa or organisms (e.g., "biodiversity of dragonflies", "biodiversity of wildflowers", "biodiversity of zooplankton"); ii) locality or geographic regions (e.g., "Indo-Pacific biodiversity", "tropical biodiversity"); and iii) habitats (e.g., "biodiversity of deserts", "biodiversity of coral reefs", "benthic biodiversity"). We interpreted these variables as the "descriptors" of the title (Figure S1B, S1C). Therefore, for each study, the number of descriptors varied from 0 to 3.

Concerning the facets of biodiversity, we marked ("yes" or "no") whether a study considered: i) taxonomic diversity; ii) (phylo)genetic diversity; iii) functional diversity; and iv) other forms of diversity (e.g., cultural diversity). Next, we noted the different organisms considered in the study at the Phylum (for animals, plants, and fungi) or higher-order (for microorganisms) level. Given the frequently changing taxonomy of microorganisms, we simply scored whether a study considered "Protista" (an artificial category used in several studies), "Bacteria", "Archaea", and/or "Viruses", showing that we are already biased towards larger

organisms even when trying to disentangle such biases. As a backbone taxonomy for Phyla/Divisions, we followed ref. [S9] for Metazoa and ref. [S10] for Fungi. Regarding land plants, we adopted the traditional division into Bryophyta, Pteridophyta, Gymnospermae, and Angiospermae. Regarding algae, given the number of classifications adopted by different authors and the number of *incertae sedis* taxa, we decided to group it in just one category ("Algae"). We also included three generic categories for animals, fungi, and plants ("Animal_generic", "Plant_generic", "Fungi_generic") to be used for general studies when taxa were not explicitly named.

We calculated the number of groups that were considered in the article as "Observed biodiversity". We considered the sum of Phyla (56 groups) to be our reference pool of biodiversity ("Expected biodiversity") and used this reference to calculate the sampled proportion of biodiversity for each article.

Scientometric factors

To explore the relationship between sampled biodiversity, use of descriptors, and impact of a given paper, we extracted two measures of article impact: i) the number of citations received by each paper on the Web of Science; and ii) the Altmetric score, a measure of the general attention that a scholarly article has received online. Furthermore, we selected three confounding factors that are well-known correlates of these measures of impact[S11–S14]: i) Journal Impact Factor at the year of publication, based on annual Journal Citation Reports by Clarivate Analytics; ii) the number of coauthors in a given paper; and iii) the diversity of countries represented in the author's list (i.e., the number of unique countries based on the author's affiliations).

7

Data analysis

We carried out all analyses in R version 4.1.0[S15] and used the package 'stats' version 4.1.0 for modelling and 'ggplot2' version 3.3.4[S16] for visualizations. In all regression-type analyses, we followed the general protocol by ref. [S17]. For data exploration, we visually inspected variable distribution and presence of outliers, multicollinearity among predictors, and balance of factor levels[18]. In regression models, we scaled continuous variables to facilitate convergence.

Predictors of sampled biodiversity

To evaluate whether studies are increasing their taxonomic scope in recent years, we modeled the relationship between the proportion of biodiversity and the year of publication with a quasibinomial regression. By visually inspecting the data, we noticed they presented a "triangular" distribution with most data concentrating around zero (i.e., low sampled biodiversity), and a minor fraction of outliers that visually seemed to increase in recent years. Thus, we repeated the model by fitting two quantile regressions, one with the full set of data and another with the data in the 75–100th percentile.

Next, we explored the role of different factors in explaining the observed biodiversity (dependent variable). As a result of data exploration, we removed two extreme outliers from the dependent variable Observed biodiversity. These were two studies with sampled biodiversity of 22 and 25, alone defining the 25–100% percentiles of the variable and thus able to strongly inflate the regression coefficient estimation. We decided to exclude these observations rather

than transforming the data because the response variable was our primary interest[18]. No collinearity was detected among predictors. Finally, in the categorical variable "Domain", we created a new level "Aquatic" to balance the factor levels, merging the levels "Saltwater" and "Freshwater".

We fitted an initial model assuming a Poisson error structure and a log link function to achieve positive fitted values. The model had the formula (in R notation):

(eq. 1) Observed biodiversity ~ Publication year + Domain + Biogeography + Method + Phylogenetic diversity + Functional diversity + Other diversities + Mention of location in title + Mention of habitat in title + Mention of taxon/a in title

The model was overdispersed (dispersion ratio = 1.611; Pearson's Chi^2 = 918.199, p < 0.001). Therefore, we fitted a new model assuming a negative binomial distribution—i.e., a generalization of Poisson distribution which loosens the assumption that the variance should be equal to the mean.

Drivers of article impact

We tested for relationships between article impact (citation or Altmetric counts) and seven article-level predictors. As a result of data exploration, we excluded the number of coauthors as this variable was correlated with the number of coauthors' countries (Pearson's r = 0.63). Furthermore, we log-transformed observed biodiversity, Impact factor, and the number of

countries of the coauthors to homogenize their distributions and deal with a few outliers. Given that old papers had more time to attract citations and Altmetric attention than recent ones, we obtained a measurement of citation and Altmetric counts unaffected by age. Following ref. [S19], we fitted two Poisson generalized additive models, exploring the relationship between the measure of article impact and the age of the paper. We then extracted the Pearson residuals from the two models, and used the age-residual values for citations and Altmetric scores as the response variables in two linear mixed models with the following formula (in R notation):

(eq. 2) Article impact ~ Impact Factor + Number of countries of coauthors + Mention of location in title + Mention of habitat in title + Mention of taxon/a in title + Observed biodiversity : N° of descriptors in title

Note that, in the model, we tested for the interaction between observed biodiversity and the number of descriptors used in the title (see Figure S1C). The Impact Factor and Number of countries of the co-authors were included as confounding factors. Specifically, by the design of the study, we assumed that articles with a greater number of coauthors and published in high-impact factor venues will, on average, achieve a greater impact[S11,S12,S14].

SUPPLEMENTARY REFERENCES

S1.   Heberling, Mason, J., Joseph, M.T., Noesgaard, D., Scott, W.B., and Schigel, D. (2021). Data integration enables global biodiversity synthesis. Proc. Natl. Acad. Sci. *118*, e2018093118.

S2.   Tosa, M.I., Dziedzic, E.H., Appel, C.L., Urbina, J., Massey, A., Ruprecht, J., Eriksson, C.E., Dolliver, J.E., Lesmeister, D.B., Betts, M.G., et al. (2021). The rapid rise of next-generation natural history. Front. Ecol. Evol. *9*, 480.

S3.   Anderson, S.C., Elsen, P.R., Hughes, B.B., Tonietto, R.K., Bletz, M.C., Gill, D.A., Holgerson, M.A., Kuebbing, S.E., McDonough MacKenzie, C., Meek, M.H., et al. (2021). Trends in ecology and conservation over eight decades. Front. Ecol. Environ. *19*, 274–282.

S4.   Cavender-Bares, J., Schneider, F.D., Santos, M.J., Armstrong, A., Carnaval, A., Dahlin, K.M., Fatoyinbo, L., Hurtt, G.C., Schimel, D., Townsend, P.A., et al. (2022). Integrating remote sensing with ecology and evolution to advance biodiversity conservation. Nat. Ecol. Evol. *6*, 506–519.

S5.   Feng, X., Enquist, B.J., Park, D.S., Boyle, B., Breshears, D.D., Gallagher, R. V, Lien, A., Newman, E.A., Burger, J.R., Maitner, B.S., et al. (2022). A review of the heterogeneous landscape of biodiversity databases: Opportunities and challenges for a synthesized biodiversity knowledge base. Glob. Ecol. Biogeogr. *31*, 1242–1260.

S6.   Murphy, S.M., Vidal, M.C., Hallagan, C.J., Broder, E.D., Barnes, E.E., Horna Lowell, E.S., and Wilson, J.D. (2019). Does this title bug (Hemiptera) you? How to write a title that increases your citations. Ecol. Entomol. *44*, 593–600.

S7.   Heard, S.B., Cull, C.A., and White, E.R. (2022). If this title is funny, will you cite me? Citation impacts of humour and other features of article titles in ecology and evolution. bioRxiv, 2022.03.18.484880.

S8.   Mabe, M.A., and Amin, M. (2002). Dr Jekyll and Dr Hyde: author☐reader asymmetries in scholarly publishing. Aslib Proc. *54*, 149–157.

S9.   Nielsen, C. (2012). Animal Evolution: Interrelationships of the Living Phyla 3rd editio. (Oxford University Press, UK).

S10. Naranjo-Ortiz, M.A., and Gabaldón, T. (2019). Fungal evolution: diversity, taxonomy and phylogeny of the Fungi. Biol. Rev. *94*, 2101–2137.

S11. Tahamtan, I., and Bornmann, L. (2019). What do citation counts measure? An updated

review of studies on citations in scientific documents published between 2006 and 2018. Scientometrics *121*, 1635–1684.

S12.   Tahamtan, I., Safipour Afshar, A., and Ahamdzadeh, K. (2016). Factors affecting number of citations: a comprehensive review of the literature. Scientometrics *107*, 1195–1225.

S13.   Araujo, A.C., Vanin, A.A., Nascimento, D.P., Gonzalez, G.Z., and Costa, L.O.P. (2021). What are the variables associated with Altmetric scores? Syst. Rev. *10*, 193.

S14.   Mammola, S., Piano, E., Doretto, A., Caprio, E., and Chamberlain, D. (2022). Measuring the influence of non-scientific features on citations. Scientometrics *127*, 4123–4137.

S15.   R Core Team (2021). R: A Language and Environment for Statistical Computing.

S16.   Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis. (Springer-Verlag).

S17.   Zuur, A.F., and Ieno, E.N. (2016). A protocol for conducting and presenting results of regression-type analyses. Methods Ecol. Evol. *7*, 636–645.

S18.   Zuur, A.F., Ieno, E.N., and Elphick, C.S. (2009). A protocol for data exploration to avoid common statistical problems. Methods Ecol. Evol. *1*, 3–14.

S19.   Mammola, S., Fontaneto, D., Martínez, A., and Chichorro, F. (2021). Impact of the reference list features on the number of citations. Scientometrics *126*, 785–799.

SUPPLEMENTAL STATEMENTS

Author contribution

Conceptualization: SM, CSF, PC; Data collection: all authors except SM and PC; Analysis: SM; Writing, first draft: SM; All authors read the text, provided comments, suggestions, and corrections, and approved the final version.

SUPPLEMENTAL DATA AND CODE

The database used in the analyses is available in Figshare (doi: 10.6084/m9.figshare.21027043). The R code to generate analyses and figures is available on GitHub (https://github.com/StefanoMammola/Analysis_Mammola_et_al_Biodiversity_Misuse.git).