

1 **Spatial scale moderates the shape of the biodiversity-disease relationship**

2 Running title: Spatial scale, biodiversity, and disease risk

3

4 Fletcher W. Halliday<sup>1,2</sup>, Jason R. Rohr<sup>1,3</sup>

5 <sup>1</sup> Department of Integrative Biology, University of South Florida, Tampa, FL 33620

6 <sup>2</sup> Department of Biology, University of North Carolina, Chapel Hill, NC 27599

7 <sup>3</sup> Department of Biological Sciences

8 Correspondence: 120 South Road, CB#3280, Chapel Hill NC, 27499; fh@live.unc.edu; 919-843-

9 8177

10

11 **Classification:** Biological Sciences; Ecology

12 **Corresponding Author:** Fletcher W. Halliday

13 **Keywords:** biodiversity; parasitism; dilution effect; amplification effect

14

15 **Abstract**

16 Diverse host communities commonly inhibit the spread of parasites in studies at small and  
17 intermediate scales, leading some to suggest that conserving biodiversity could help control  
18 infectious diseases. However, the generality of this “dilution effect” remains controversial. First,  
19 most studies assume a linear, monotonic relationship between biodiversity and disease, though  
20 the actual shape is unknown. Second, most studies are conducted at a single spatial scale, though  
21 biotic interactions are often scale-dependent, thus spatial scale might determine the direction of  
22 biodiversity-disease relationships. Third, most studies focus only on a small range of possible  
23 diversity levels, though the direction of biodiversity-disease relationships may change outside of  
24 this range. By analyzing 231 biodiversity-disease relationships on 77 parasite species, we  
25 provide broad evidence that biodiversity-disease relationships are generally non-linear and  
26 moderated by spatial scale; biodiversity generally inhibits disease at local scales (<100 km<sup>2</sup>) and  
27 amplifies disease at regional scales (>1,000,000 km<sup>2</sup>). These effects did not depend on any tested  
28 host, parasite, or study characteristics, though the spatial scale of a study was often related to  
29 study design and parasite type, highlighting the need for additional multiscale research. Few  
30 studies were missing substantial data at low diversity, but missing data at low diversity could  
31 result in underreporting of amplification. Experiments might be missing data at high diversity,  
32 which could result in underreporting of dilution. Despite context-dependence in biodiversity-  
33 disease relationships, most conservation is implemented at local scales where biodiversity  
34 appears to inhibit disease and thus these results suggest that local conservation actions could  
35 reduce disease risk.

36

37

38

39 **Significance statement**

40 It has been suggested that diverse ecological communities limit disease spread, but the generality  
41 of this pattern is contentious. Therefore, the degree to which biodiversity conservation can limit  
42 harmful epidemics remains unresolved. We address this fundamental question by analyzing 231  
43 published relationships between biodiversity and disease. We find evidence that most  
44 biodiversity-disease relationships are nonlinear and scale-dependent with biodiversity generally  
45 associated with reduced disease at small and intermediate scales, but increased disease at large  
46 scales. Moreover, these results were generally robust to missing data at low and high biodiversity  
47 levels and variation in host, parasite, and study characteristics. This suggests that conservation  
48 efforts aimed at reducing the impacts of human and wildlife diseases will be most successful at  
49 local scales.

50

51

## 52 **Introduction**

53           Understanding whether there is a general relationship between biodiversity and disease  
54 risk is critical for projecting and reducing the impacts of future disease outbreaks (1, 2, 11–15,  
55 3–10). If increasing biodiversity generally reduces disease, a phenomenon coined the dilution  
56 effect, then biodiversity loss could have negative consequences for human and wildlife  
57 populations (2, 16). Concordantly, biodiversity conservation could limit the spread of unknown  
58 infectious diseases (13, 17). However, if biodiversity-disease relationships are idiosyncratic or  
59 context dependent, then biodiversity conservation could have no effect on, or, in the case of an  
60 amplification effect, even exacerbate the risk of disease to wildlife and humans (18, 19). Such  
61 context dependence in the biodiversity-disease relationship has become a major concern among  
62 disease ecologists (8, 13, 20). Consequently, the value of conserving biodiversity to protect  
63 against disease risk has been called into question (20–22).

64           Context dependence in the biodiversity-disease relationship can arise when the shape of  
65 the biodiversity-disease relationship is nonlinear. By definition, parasites require hosts for food  
66 and habitat. Thus, all else being equal, an increase in host biodiversity from zero hosts must  
67 initially increase the risk of disease (1, 12). However, if parasites are selected to infect the most  
68 abundant and widespread hosts or there are trade-offs between defending against parasites and  
69 host growth, reproduction, and dispersal, then communities might assemble in a manner where  
70 the first species added to communities are generally competent, disease amplifying hosts and  
71 later additions might be rarer, diluting hosts (23). If so, the initial increase in disease risk when  
72 moving from zero hosts to a few might reverse at higher diversity levels (Fig. 1a), and the skew  
73 of the biodiversity-disease relationship might affect the predominance of amplification or  
74 dilution. When biodiversity-disease relationships are left-skewed or asymptotic, amplification

75 effects should predominate, because most increases in biodiversity will be associated with  
76 increased parasite abundance (12) (Fig. 1a). Alternatively, when biodiversity-disease  
77 relationships are right-skewed, dilution should predominate (12). Nevertheless, understanding  
78 the shape of nonlinear biodiversity-disease relationships remains a major research gap (1, 15, 20,  
79 23).

80         Where communities fall on nonlinear biodiversity-disease curves is also important. If  
81 changes in biodiversity all occur to the right or left of the peak of unimodal diversity-disease  
82 curves, then dilution or amplification, respectively, will be most common, regardless of the  
83 direction of the skew of that relationship (Fig 1b). This might create biases for both observational  
84 and manipulative studies. Observational studies might not capture low host diversity levels if  
85 they are rare in nature, which could lead researchers to spuriously conclude that there is a linear  
86 dilution effect, even though amplification might occur at non-sampled low levels of host  
87 diversity (15, 20) (Fig 1b). In contrast, manipulative studies might include mostly low levels of  
88 biodiversity because of the logistical challenges of collecting sufficient numbers of many host  
89 species for an experiment; this could potentially bias results towards amplification. The former  
90 bias can be easily detected by simply quantifying the minimum diversity level in observational  
91 studies, whereas the latter bias of experimental studies might be more difficult to assess because,  
92 in most systems, the distribution of natural levels of host diversity is unknown.

93         Context dependence in the biodiversity-disease relationship may also arise when the  
94 direction of the biodiversity-disease relationship depends on the spatial scale of observation (4,  
95 7, 20). Local processes influence the abundance of species at relatively small spatial scales,  
96 while regional processes influence the distributions of species across large spatial extents (24).  
97 Relying on this well-characterized ecological phenomenon, it has been proposed that

98 biodiversity-disease relationships should be strongest at local scales, where biotic interactions are  
99 most likely to occur, and should weaken or even reverse at larger scales, where abiotic factors  
100 like climate may cause the distributions of hosts and parasites to covary (23, 25). Moreover,  
101 whether hosts can dilute disease might be more observable at small scales where encounter  
102 reduction can occur, while the amplifying effect of hosts might only be observable at larger  
103 temporal and spatial scales (26). Even though theory indicates that spatial scale can moderate  
104 biodiversity-disease relationships, and biodiversity-disease studies have occurred from global to  
105 local scales (3, 27, 28), few studies have been conducted across multiple spatial scales. Thus the  
106 degree to which biodiversity-disease relationships are moderated by spatial scale remains largely  
107 untested (but see 25).

108         By quantifying the shape and direction of 231 published biodiversity-disease  
109 relationships, this study aimed at testing three contingencies in the biodiversity-disease  
110 relationship. Specifically, we tested whether biodiversity-disease relationships are generally (a)  
111 non-linear, (b) moderated by spatial scale, and (c) sensitive to missing data at low and high  
112 diversity. Our results indicate that biodiversity-disease relationships are generally non-linear, that  
113 dilution most commonly occurs at small (i.e., local) scales and amplification most commonly  
114 occurs at large (i.e., regional) scales, that few studies are missing substantial data at low  
115 diversity, but that missing data at low diversity could potentially result in the underreporting of  
116 amplification effects, and that experimental studies might be missing data at high diversity levels  
117 that could potentially result in underreporting of dilution effects.

118

## 119 **Results and Discussion**

120 *What is the shape of the biodiversity-disease relationship?*

121 First, we tested whether the published relationship between biodiversity and disease was  
122 linear or nonlinear by comparing intercept-only, linear, second-order, and third-order polynomial  
123 regression models for all biodiversity-disease relationships, selecting the best-fitting model using  
124 AIC (Table S1). Importantly, these models were only based on the data presented in each study  
125 and thus did not constrain the biodiversity-disease relationship to the origin (see “*Do missing*  
126 *data at low and high diversity bias studies to report dilution effects?*” section below). Out of the  
127 231 studies that included more than three levels of biodiversity, 63% were best fit by a linear,  
128 second-order, or third-order polynomial model (i.e., exhibited a relationship between biodiversity  
129 and disease). Of these studies, biodiversity-disease relationships were most commonly non-  
130 linear, as predicted. More specifically, 60% exhibited non-linear relationships (either second- or  
131 third-order polynomial), while 8% exhibited linear amplification effects, and 32% exhibited  
132 linear dilution effects.

133 While comparing regression models identified many non-linear biodiversity-disease  
134 relationships, this approach is constrained by the functional form of each regression model. In  
135 other words, we are only able to detect non-linear relationships where those relationships were  
136 best fit by second- or third-order polynomials. To relax this constraint, we used Spearman rank  
137 correlation tests (not constrained to pass through the origin), which make no assumption about  
138 the underlying distribution of the data nor the linearity of the relationship between variables, and  
139 are therefore not constrained by the functional form of the biodiversity-disease relationship. We  
140 quantified whether each biodiversity-disease relationship was monotonic and positive (disease  
141 increases, but may level off, as diversity increases), monotonic and negative (disease decreases  
142 but may level off as diversity increases), or non-monotonic (disease increases with diversity at  
143 low levels, but eventually decreases at high enough diversity; Fig. 1a). The estimated Spearman

144 rank correlation coefficient ( $\rho$ ) approaches one for monotonic, positive relationships, and  
145 approaches negative one for monotonic, negative relationships. We therefore used  $\rho$  to define  
146 monotonic amplification ( $\rho > 0, p < 0.05$ ), monotonic dilution ( $\rho < 0, p < 0.05$ ), and non-significant  
147 or non-monotonic relationships ( $p > 0.05$ ). Consistent with the previous analysis, 15% of the 231  
148 relationships exhibited monotonic amplification effects, 31% exhibited monotonic dilution  
149 effects, and 54% exhibited non-significant or non-monotonic relationships.

150       Given that non-linear and non-monotonic biodiversity-disease relationships are most  
151 common and that amplification effects might predominate when these relationships are left-  
152 skewed or asymptotic, whereas dilution might predominate when they are right-skewed (12), we  
153 next assessed the skew of each biodiversity-disease relationship. To do so, we fit a smoothing  
154 spline to each published biodiversity-disease relationship, that was not constrained to pass  
155 through the origin, and then calculated Pearson's skewness from the shape of the estimated  
156 curve, excluding studies where there was no relationship (i.e., where the slope of the curve was  
157 not significantly different from zero;  $n=29$ ). As expected, Pearson's skewness and Spearman  
158 rank correlation were in agreement when studies exhibited monotonic biodiversity-disease  
159 relationships. Specifically, studies exhibiting monotonic dilution effects were significantly right-  
160 skewed ( $p < 0.001$ ), and studies exhibiting monotonic amplification effects were significantly left-  
161 skewed ( $p < 0.001$ ; Fig. 2). Studies exhibiting non-significant or non-monotonic relationships  
162 based on Spearman rank correlation were not significantly skewed ( $p=0.59$ ), indicating that non-  
163 monotonic biodiversity-disease relationships, on average, showed similar levels of amplification  
164 or dilution. These results were qualitatively similar for the analysis comparing intercept-only,  
165 linear, second-order, and third-order polynomial regression models (Fig. S1, Fig. S2).



166        These results support the hypothesis that the shape of biodiversity-disease relationships  
167 might be nonlinear (12, 20) and might therefore have implications for biodiversity conservation.  
168 Specifically, because biodiversity-disease relationships are often nonlinear, where an individual  
169 system falls along a biodiversity gradient might influence whether that system experiences  
170 amplification or dilution. Thus, understanding how an individual conservation action will alter  
171 biodiversity can have a large impact on whether that action is expected to increase or decrease  
172 wildlife and human disease risk.

173

174 *Is the biodiversity-disease relationship moderated by spatial scale?*

175        Second, we tested whether the shape and direction of the biodiversity-disease relationship  
176 was moderated by spatial scale, measured as the spatial extent of each study. Spatial extent  
177 represents the total area over which a study is conducted, including all measures of biodiversity  
178 and disease for a given study. Spearman's  $\rho$ , which measures the monotonicity and direction of  
179 association between biodiversity and disease, was positively associated with spatial extent  
180 ( $p < 0.001$ ; marginal  $R^2 = 0.34$ ; Fig 3a), with monotonic dilution effects most commonly occurring  
181 at small to intermediate spatial scales and monotonic amplification effects most commonly  
182 occurring at the largest spatial scales. Incorporating the shape of non-monotonic relationships did  
183 not alter this result; Pearson's skewness was significantly associated with spatial extent  
184 ( $p < 0.001$ ; marginal  $R^2 = 0.22$ ; Fig 3b), with right-skewed relationships (indicating more dilution)  
185 occurring at small to intermediate spatial scales and left-skewed relationships (indicating more  
186 amplification) occurring at large spatial scales.

187        Dilution generally occurred *within* an ecosystem, at spatial extents  $< 100 \text{ km}^2$  (roughly the  
188 size of a small city), whereas amplification generally occurred *across* ecosystems, in studies

189 occupying >1,000,000 km<sup>2</sup> (roughly the size of France and Spain combined). These results  
190 therefore indicate that the overall disease burden in one ecosystem can be higher than another  
191 because its native biodiversity is higher, but if this ecosystem has its biodiversity lowered,  
192 disease could still worsen. Consequently, these results indicate that within individual countries,  
193 conserving biodiversity might improve human, wildlife, and ecosystem health.

194 This dependence of biodiversity-disease relationships on spatial scale may be an indicator of  
195 a more general mechanism of disease amplification. Notably, comparison of biodiversity-disease  
196 relationships *within* an ecosystem often include many of the same host species (e.g., 29), whereas  
197 comparisons of biodiversity *across* ecosystems tend to include distinct sets of host species (e.g.,  
198 30). Thus, measuring the degree to which host-species turnover ( $\beta$  diversity) drives disease  
199 amplification could help clarify why amplification occurs at large spatial extents, and possibly  
200 help predict when amplification will be more common, in general.

201 These results reveal a strong, widespread association between the shape of biodiversity-  
202 disease relationships and the spatial scale of observations, supporting the hypothesis that  
203 biodiversity-disease relationships are scale-dependent (7, 23). Importantly, however, not every  
204 small-scale study exhibited dilution, nor did every large-scale study exhibit amplification. As an  
205 example, in a global survey of human disease burden, disease generally increased with  
206 increasing diversity, but human schistosomiasis was negatively correlated with diversity (30)  
207 and, at small spatial scales, biodiversity can amplify disease via a sampling effect, if species are  
208 added randomly with respect to host competence and transmission is frequency-dependent (31).

209 We tested whether several ecological factors could explain variation in the effect of spatial  
210 scale on the shape of biodiversity-disease relationships. Specifically, we tested whether the effect  
211 of spatial scale on biodiversity-disease relationships differed between (i) parasites that infect

212 humans vs. wildlife, (ii) macro- vs. microparasites, (iii) parasites with complex vs. direct  
213 lifecycles, and (iv) observational vs. manipulative studies. We found no evidence that the effect  
214 of spatial scale on biodiversity-disease relationships depended on any of these factors (Table 1).  
215 Thus, the effect of spatial scale on biodiversity-disease relationships was generally robust across  
216 all ecological contexts examined. However, we encourage caution in interpreting these results, as  
217 there was multicollinearity in these analyses. Specifically, observational studies and studies of  
218 human pathogens both tended to occur at larger spatial scales than manipulative studies and  
219 studies of wildlife pathogens (Fig. S3). Consequently, we cannot rule out the possibility that  
220 these results could change if future studies filled these research gaps, allowing tests of these  
221 context dependencies to be less collinear.

222

223 *Do missing data at low and high diversity bias biodiversity-disease studies?*

224 Finally, we tested the hypothesis that missing data at the highest and lowest diversity  
225 levels in experimental and observational studies might bias studies to more commonly report  
226 amplification and dilution effects, respectively. Experimental studies had a lower mean  
227 maximum diversity level than observational studies (experimental mean  $\pm$  sd:  $25 \pm 10$ ,  
228 observational mean  $\pm$  sd:  $76 \pm 89$ ). Thus, it appears that experimental studies are missing data at  
229 the highest diversity levels, which could bias experimental studies towards amplification effects.  
230 This result could emerge from two key differences between experiments and observational  
231 studies. First, experimentally manipulating many species is logistically challenging at  
232 high richness, potentially biasing experimental studies to include fewer total species than  
233 observational studies of equivalent size. Second, the number of species in an area is highly  
234 sensitive to the area surveyed (32), and observational studies were, on average, five orders of

235 magnitude larger than manipulative experiments (Fig S3). Focusing on studies of comparable  
236 extent (1-10 km<sup>2</sup>) eliminated the difference in mean maximum diversity between experiments  
237 ( $29 \pm 10$ ) and observational studies ( $22.0 \pm 10$ ), supporting this second mechanism.

238 We also examined the lowest diversity levels to assess whether there was missing data at  
239 low diversity. Experimental studies had lower mean minimum diversity than observational  
240 studies (experimental mean  $\pm$  sd:  $1.2 \pm 0.6$ ), observational mean  $\pm$  sd:  $4.3 \pm 7.3$ ), which could  
241 bias observational studies towards dilution effects. However, out of 231 studies, 84% of studies  
242 included an effective species richness of two or lower. Consequently, most studies ( $n=193$ ) were  
243 not missing substantial data at low host diversity. This result indicates that the potential for  
244 missing data at low diversity to bias the estimated relationship between biodiversity and disease  
245 is quite low.

246 Even though most studies were not missing substantial data at low host diversity, we still  
247 performed an additional test of the hypothesis that missing data might bias studies to more  
248 commonly report dilution effects. Here, we again quantified the skew of each biodiversity-  
249 disease relationship, this time constraining each curve to pass through the origin because if there  
250 are no hosts there cannot be any parasites. Constraining each curve to pass through the origin  
251 should reduce the estimated skew in all studies, particularly studies that found monotonic  
252 dilution effects. As predicted, constraining the curves to the origin significantly changed the  
253 shape of the average biodiversity-disease relationship ( $p = 0.001$ ), reducing the estimated  
254 frequency of dilution effects and increasing the estimated frequency of amplification effects (Fig.  
255 4). This result indicates that ignoring missing data may bias some studies to underreport  
256 amplification effects. However, even though constraining curves to fit through the origin shifted  
257 the estimated skew, on average, the constrained curves were only marginally significantly left-

258 skewed ( $p=0.051$ ; Fig. S4). Furthermore, spatial scale still significantly moderated the sign of the  
259 constrained curves, with dilution more common at small scales and amplification more common  
260 at large scales ( $p<0.0001$ ; Fig. S5), and this effect was still robust to ecological characteristics of  
261 individual study systems (Table S2).

262         These results indicate that scale-dependence of the biodiversity-disease relationship is  
263 robust to missing data at low diversity levels. The robustness of this scale-dependence may be a  
264 product of the underlying shape of biodiversity-disease relationships. Studies that found  
265 monotonic amplification effects were unlikely to be altered by missing data at low diversity (Fig  
266 1a). Conversely, studies that found monotonic dilution had higher potential to be altered by  
267 missing data at low diversity. However, 66 of the 72 studies showing monotonic dilution effects  
268 included an effective species richness of two or lower. Thus, regardless of the shape of the  
269 relationship between the origin and the point of peak parasite abundance, the area in which  
270 amplification could occur was generally quite small. We therefore conclude that although the  
271 biodiversity-disease relationship can take on many forms, and its form may depend on a  
272 nonlinearity that is driven by parasite extinction at low host diversity, such nonlinearities are  
273 unlikely to alter a general and common phenomenon: dilution effects most commonly occur at  
274 local scales and amplification effects most commonly occur at regional scales.

275         Together, these results indicate that the scale-dependence of biodiversity-disease  
276 relationships might be robust to missing data at low diversity levels and to ecological  
277 characteristics of individual studies. However, we are unable to test whether missing data at high  
278 diversity might bias experimental studies to more commonly observe amplification. This bias of  
279 experimental studies is difficult to assess because, in most systems, the distribution of natural  
280 levels of host diversity is unknown. Furthermore, a major limitation to this analysis is the lack of

281 empirical studies conducted across spatial scales and ecological conditions using the same  
282 methodologies. We are hopeful that as large-scale replicated studies, such as the Nutrient  
283 Network (33) and National Environmental Observatory Network (34), become more widespread,  
284 the quality of data at the largest spatial scales will improve. These results also highlight the need  
285 to consider whether the scales of conservation actions and public health interventions are  
286 appropriate to influence biodiversity-disease relationships in a way that will benefit humans (3,  
287 27, 30). In general, these results suggest that biodiversity conservation can be beneficial to  
288 human health when conducted at small or intermediate scales, which are the scales at which they  
289 are most commonly implemented (35, 36).

290 Understanding how biodiversity alters infectious diseases remains a critical frontier in  
291 disease ecology (17, 23, 37) as human activities continue to alter global biodiversity (38, 39),  
292 and disease outbreaks continue to increase (40–42). This study provides quantitative evidence  
293 that the relationship between biodiversity and disease is non-linear and scale-dependent. This  
294 general pattern indicates that biodiversity loss could exacerbate disease outbreaks at the scales in  
295 which humans are most likely to encounter disease, and highlights important scales in which  
296 biodiversity conservation might be most useful for minimizing and mitigating these  
297 consequences.

298

## 299 **Materials and Methods**

### 300 *Data compilation*

301 This study aimed to analyze the shape of every published relationship between host  
302 diversity and the abundance of parasites. We updated the list of studies from Civitello et al. (8) to  
303 include studies published between 2014 and 2018, by repeating their original search criteria.

304 Specifically, we searched the Web of Science for several combinations of search terms: parasite,  
305 pathogen, diversity, richness, evenness, dilution effect, and amplification effect (the final search  
306 was conducted in June 2018). We identified additional papers by searching the literature cited  
307 sections of these articles and by searching Web of Science for all papers citing Civitello et al (8),  
308 including those critical of the dilution effect hypothesis. We included observational and  
309 experimental studies in lab and field environments.

310

### 311 *Selection criteria and data collection*

312 We only included studies that measured parasite abundance or prevalence at more than  
313 two host diversity levels. We included studies that reported infection prevalence, mean parasite  
314 load, density of infected vectors, or percent diseased tissue, because these quantities are the most  
315 relevant metrics of disease risk for microparasites, macroparasites, vector-borne parasites, and  
316 plant parasites, respectively. We did not standardize parasite abundance, and therefore did not  
317 compare parasite abundance among studies. Host biodiversity was reported as species richness,  
318 Simpson's diversity index (J), or Shannon's diversity index (H). We standardize across these  
319 measures to facilitate comparisons across studies by transforming diversity into the effective  
320 number of species, following Jost (43). In experiments, estimated diversity included all taxa  
321 added by the experimenters, while the diversity estimate in observational studies was limited to a  
322 focal taxonomic or functional group of host species, defined in the primary study (e.g.,  
323 herbaceous plants, trees, birds, or small mammals).

324 We extracted data from text and tables manually and from figures using  
325 WebPlotDigitizer version 4.1 (44), and recorded other data relating to the biology or  
326 methodology of each study. For all studies, we recorded parasite and host taxa, type of parasite

327 (infecting only wildlife or also infecting humans), focal host species, associated species (i.e.,  
328 additional species whose presence may dilute or amplify parasite abundance, operationally  
329 defined as “potential diluters”), the diversity (e.g., richness) in the treatments (or in the field  
330 survey), parasite functional group (macroparasite vs. microparasite), parasite lifecycle (complex  
331 vs. direct), and study design (manipulative vs. observational). Spatial extent was quantified as  
332 the area (expressed in square kilometers) over which all biodiversity estimates were compared in  
333 a given study. Studies rarely provided an exact value for spatial extent. Because a value for  
334 spatial extent was rarely provided, and spatial extents varied by six orders of magnitude, we  
335 estimated the extent of each survey to the nearest order of magnitude rather than attempting to  
336 assign a specific spatial extent for each study. For example, we assigned studies a value of 0.1 if  
337 the extent was less than 1 km<sup>2</sup>, and a value of 1 if the extent was greater than 1 km<sup>2</sup>, but less than  
338 10 km<sup>2</sup>, etc.

339

#### 340 *Assessing the shape of the biodiversity-disease relationship*

341 We first quantified whether each biodiversity-disease relationship was linear or nonlinear  
342 by comparing a series of regression models using the lm and AIC functions in R version 3.5.0  
343 (45) (see main text for methods). Four studies included fewer than five host diversity levels and  
344 were therefore not tested using a third-order polynomial. Next, we quantified the monotonicity  
345 and direction of each biodiversity-disease relationship using Spearman rank correlations (see  
346 main text for methods). We then assessed the skew of each biodiversity-disease relationship  
347 using R package cobs (46) to fit an unconstrained spline to the biodiversity-disease relationship,  
348 limited to a maximum of four knots to prevent overfitting. This approach to fitting an  
349 unconstrained curve makes no assumptions about the underlying shape of the relationship



350 between biodiversity and disease. We transformed the predicted curve into a frequency  
351 distribution, assigning any negative value (occurring in 19 regressions) to zero, and then  
352 calculated Pearson's skewness. A right-skewed relationship (Pearson skewness  $> 0.25$ ) indicates  
353 that most of the data falls in the area where dilution is observed, while a left-skewed relationship  
354 (Pearson skewness  $< -0.25$ ) indicates the possibility for measured or unmeasured amplification  
355 effects. To assess whether missing data at low diversity could bias the estimated shape of the  
356 biodiversity-disease relationship, we constrained curves to pass through the origin and again  
357 calculated the skewness of each curve. Specifically, to fit qualitatively constrained quantile (CQ)  
358 smoothing splines (47), we added a value at the origin for each data set, corresponding to a  
359 situation in which there is no host diversity, generated a constraint matrix to force the line  
360 through the origin, and then fit the curve, limiting the maximum number of knots in the curve to  
361 three to prevent overfitting.

362 We omitted studies with fewer than four unique measures of host diversity for Spearman  
363 rank correlations and unconstrained splines and fewer than three unique measures of host  
364 diversity for CQ splines. Twenty-nine of the unconstrained splines ( $n=231$ ) and 39 of the CQ  
365 splines ( $n=243$ ) showed no relationship between biodiversity and disease (e.g., a fit with a slope  
366 of zero), resulting in no estimate of Pearson's skewness. This resulted in 231 estimates of  $\rho$ , 202  
367 estimates of skew from unconstrained splines, and 204 estimates of skew from CQ splines.

368

### 369 *Data analysis*

370 All analyses were carried out in R version 3.5.0 (45). We constructed multilevel random  
371 effects models using the lmer function in R packages lme4 (48) and lmerTest (49). We accounted

372 for nonindependence arising from multiple measures from the same observational units in the  
373 same year by including such non-independent surveys as random intercepts in each model.

374 Using the model described above, we first tested whether studies exhibiting no  
375 relationship, linear amplification, linear dilution, a unimodal relationship or a third-order  
376 polynomial relationship predicted Spearman rank correlation and Pearson's skewness. We next  
377 verified that studies exhibiting monotonic dilution, monotonic amplification, and non-monotonic  
378 relationships (categorized using the Spearman rank correlation) predicted Pearson's skewness.  
379 Next, we tested whether the Spearman rank correlation coefficient between biodiversity and  
380 disease or Pearson's skewness were influenced by spatial extent by fitting two separate models,  
381 each with one response ( $\rho$  or skew) and one predictor (extent).

382 We then tested for context dependence in the spatial moderation of dilution effects. To  
383 test for context dependence, we fit the same two models, but included a two-way interaction  
384 between spatial extent and four binary factors that might explain variation in the effects of scale  
385 on the biodiversity-disease relationship: parasite functional group (macroparasite vs.  
386 microparasite), parasite lifecycle (complex vs. direct), study design (manipulative vs.  
387 observational), and parasite type (infects humans vs. infects only wildlife).

388 We next tested whether missing data at low and high diversity might bias studies to more  
389 commonly report amplification and dilution effects. We quantified the maximum and minimum  
390 diversity level of each study and compared whether the mean maximum and mean minimum  
391 diversity level differed between experiments and observational studies. Because the species-area  
392 relationship is nonlinear and sample area was highly variable across studies, we compared  
393 minimum and maximum diversity across studies qualitatively rather than quantitatively.

394 To quantitatively test whether missing data at low host diversity could bias studies to  
395 more commonly report dilution effects, we tested whether constraining the curves to pass  
396 through the origin altered the predicted skew. Specifically, we calculated the difference in skew  
397 between constrained and unconstrained curves and then performed an intercept-only model on  
398 this value, where an estimate significantly lower than zero would indicate that constraining the  
399 curve favored amplification, and an estimate significantly higher than zero would indicate that  
400 constraining the curve favored dilution. Finally, we analyzed whether spatial scale moderated the  
401 shape of the biodiversity-disease relationship when curves were constrained to pass through the  
402 origin. Here, we fit a model of Pearson's skewness and spatial extent and then performed the  
403 same test of context dependence on the model that was performed before.

404

#### 405 **Acknowledgements**

406 We are thankful to D. Civitello, R.W. Heckman, G. Legault, C.E. Mitchell, J. Umbanhowar, and  
407 members of the Rohr lab for helpful discussions on data analysis and interpretation. C.E.  
408 Mitchell, R. Poulin, H. Young, and S. Zhou provided raw data from published manuscripts. This  
409 research was supported by grants from the National Science Foundation (EF-1241889), National  
410 Institutes of Health (R01GM109499, R01TW010286), US Department of Agriculture (NRI  
411 2006-01370, 2009-35102-0543), and US Environmental Protection Agency (CAREER  
412 83518801) to J. R. Rohr.

413

#### 414 **References**

- 415 1. Ostfeld RS, Keesing F (2000) Biodiversity series: The function of biodiversity in the  
416 ecology of vector-borne zoonotic diseases. *Can J Zool* 78(12):2061–2078.

- 417 2. Keesing F, Holt RD, Ostfeld RS (2006) Effects of species diversity on disease risk. *Ecol*  
418 *Lett* 9(4):485–98.
- 419 3. Randolph SE, Dobson ADM (2012) Pangloss revisited: a critique of the dilution effect and  
420 the biodiversity-buffers-disease paradigm. *Parasitology* 139(07):847–863.
- 421 4. Lafferty KD, Wood CL (2013) It’s a myth that protection against disease is a strong and  
422 general service of biodiversity conservation: Response to Ostfeld and Keesing. *Trends*  
423 *Ecol Evol* 28(9):503–504.
- 424 5. Ostfeld RS (2013) A Candide response to Panglossian accusations by Randolph and  
425 Dobson: biodiversity buffers disease. *Parasitology* 140(10):1196–1198.
- 426 6. Ostfeld RS, Keesing F (2013) Straw men don’t get Lyme disease: response to Wood and  
427 Lafferty. *Trends Ecol Evol* 28(9):502–503.
- 428 7. Wood CL, Lafferty KD (2013) Biodiversity and disease: a synthesis of ecological  
429 perspectives on Lyme disease transmission. *Trends Ecol Evol* 28(4):239–247.
- 430 8. Civitello DJ, et al. (2015) Biodiversity inhibits parasites: Broad evidence for the dilution  
431 effect. *Proc Natl Acad Sci* 112(28):8667–8671.
- 432 9. Civitello DJ, et al. (2015) Reply to Salkeld et al.: Diversity-disease patterns are robust to  
433 study design, selection criteria, and publication bias. *Proc Natl Acad Sci* 112(46):E6262–  
434 E6262.
- 435 10. Salkeld DJ, Padgett KA, Jones JH, Antolin MF (2015) Public health perspective on  
436 patterns of biodiversity and zoonotic disease. *Proc Natl Acad Sci* 112(46):E6261–E6261.
- 437 11. Levi T, et al. (2016) Does biodiversity protect humans against infectious disease?  
438 Comment. *Ecology* 97(2):536–542.
- 439 12. Wood CL, et al. (2016) Does biodiversity protect humans against infectious disease?

- 440           Reply. *Ecology* 97(2):543–546.
- 441    13.   Ostfeld RS, Keesing F (2017) Is biodiversity bad for your health? *Ecosphere* 8(3).  
442           doi:10.1002/ecs2.1676.
- 443    14.   Salkeld DJ, Padgett KA, Jones JH (2013) A meta-analysis suggesting that the relationship  
444           between biodiversity and risk of zoonotic pathogen transmission is idiosyncratic. *Ecol Lett*  
445           16(5):679–686.
- 446    15.   Wood CL, et al. (2014) Does biodiversity protect against infectious disease? *Ecology*  
447           95(4):817–832.
- 448    16.   Keesing F, et al. (2010) Impacts of biodiversity on the emergence and transmission of  
449           infectious diseases. *Nature* 468(7324):647–52.
- 450    17.   Keesing F, Ostfeld RS (2015) Is biodiversity good for your health? *Science* (80- )  
451           349(6245):235–236.
- 452    18.   Power AG, Mitchell CE (2004) Pathogen spillover in disease epidemics. *Am Nat* 164  
453           Suppl(november):S79–S89.
- 454    19.   Young H, Griffin RH, Wood CL, Nunn CL (2013) Does habitat disturbance increase  
455           infectious disease risk for primates? *Ecol Lett* 16(5):656–663.
- 456    20.   Kilpatrick AM, Salkeld DJ, Titcomb G, Hahn MB (2017) Conservation of biodiversity as  
457           a strategy for improving human health and well-being. *Philos Trans R Soc B Biol Sci*  
458           372(1722):20160131.
- 459    21.   Young HS, et al. (2017) Conservation, biodiversity and infectious disease: scientific  
460           evidence and policy implications. *Philos Trans R Soc B Biol Sci* 372(1722):20160124.
- 461    22.   Luis AD, Kuenzi AJ, Mills JN (2018) Species diversity concurrently dilutes and amplifies  
462           transmission in a zoonotic host–pathogen system through competing mechanisms. *Proc*

- 463           *Natl Acad Sci*:201807106.
- 464   23.   Johnson PTJ, Ostfeld RS, Keesing F (2015) Frontiers in research on biodiversity and  
465           disease. *Ecol Lett* 18(10):1119–1133.
- 466   24.   Ricklefs RE (1987) Community Diversity: Relative Roles of Local and Regional  
467           Processes. *Science (80- )* 235(4785):167–171.
- 468   25.   Cohen JM, et al. (2016) Spatial scale modulates the strength of ecological processes  
469           driving disease distributions. *Proc Natl Acad Sci U S A* 113(24):E3359-64.
- 470   26.   Buck JC, Perkins SE (2018) Study scale determines whether wildlife loss protects against  
471           or promotes tick-borne disease. doi:10.1098/rspb.2018.0218.
- 472   27.   Guernier V, Hochberg ME, Guégan J-F (2004) Ecology Drives the Worldwide  
473           Distribution of Human Diseases. *PLoS Biol* 2(6):e141.
- 474   28.   Hechinger RF, Lafferty KD (2005) Host diversity begets parasite diversity: bird final  
475           hosts and trematodes in snail intermediate hosts. *Proc R Soc B-Biological Sci*  
476           272(1567):1059–1066.
- 477   29.   Mitchell C, Tilman D, Groth J (2002) Effects of grassland plant species diversity,  
478           abundance, and composition on foliar fungal disease. *Ecology* 83(6):1713–1726.
- 479   30.   Wood CL, McInturff A, Young HS, Kim D, Lafferty KD (2017) Human infectious disease  
480           burdens decrease with urbanization but not with biodiversity. *Philos Trans R Soc B Biol*  
481           *Sci* 372(1722):20160122.
- 482   31.   Halliday FW, Heckman RW, Wilfahrt PA, Mitchell CE, Halliday FW (2017) A  
483           multivariate test of disease risk reveals conditions leading to disease amplification.
- 484   32.   Gotelli NJ, Colwell RK (2001) Quantifying biodiversity: procedures and pitfalls in the  
485           measurement and comparison of species richness. *Ecol Lett* 4(4):379–391.

- 486 33. Borer ET, et al. (2014) Finding generality in ecology: A model for globally distributed  
487 experiments. *Methods Ecol Evol* 5(1):65–73.
- 488 34. Keller M, Schimel DS, Hargrove WW, Hoffman FM (2008) A continental strategy for the  
489 National Ecological Observatory Network. *Front Ecol Environ* 6(5):282–284.
- 490 35. Isbell F, et al. (2017) Linking the influence and dependence of people on biodiversity  
491 across scales. *Nature* 546(7656):65–72.
- 492 36. Watson JEM, Dudley N, Segan DB, Hockings M (2014) The performance and potential of  
493 protected areas. *Nature* 515(7525):67–73.
- 494 37. Young HS, Parker IM, Gilbert GS, Sofia Guerra A, Nunn CL (2017) Introduced Species,  
495 Disease Ecology, and Biodiversity–Disease Relationships. *Trends Ecol Evol* 32(1):41–54.
- 496 38. Cardinale BJ, et al. (2012) Biodiversity loss and its impact on humanity. *Nature*  
497 486(7401):59–67.
- 498 39. Dornelas M, et al. (2014) Assemblage time series reveal biodiversity change but not  
499 systematic loss. *Science* (80- ) 344(6181):296–299.
- 500 40. Harvell D, Aronson R, Baron N (2004) The Rising Tide of Ocean Diseases : Unsolved  
501 Problems and Research Priorities. *Front Ecol Environ* 2(7):375–382.
- 502 41. Jones KE, et al. (2008) Global trends in emerging infectious diseases. *Nature*  
503 451(7181):990–993.
- 504 42. Smith KF, et al. (2014) Global rise in human infectious disease outbreaks. *J R Soc*  
505 *Interface* 11(101):20140950.
- 506 43. Jost L (2006) Entropy and diversity. *Oikos* 113(2):363–375.
- 507 44. Rohatgi A (2011) WebPlotDigitizer. Available at: <https://automeris.io/WebPlotDigitizer>.
- 508 45. R Core Team (2015) R: a language and environment for statistical computing |

- 509 GBIF.ORG. doi:10.1007/978-3-540-74686-7.
- 510 46. Ng PT, Maechler M (2017) COBS -- Constrained B-splines (Sparse matrix based).  
511 Available at: <https://cran.r-project.org/package=cobs>.
- 512 47. Ng P, Maechler M (2007) A Fast and Efficient Implementation of Qualitatively  
513 Constrained Quantile Smoothing Splines. *Stat Modelling* 7(4):315–328.
- 514 48. Bates D, Mächler M, Bolker B, Walker S (2014) Fitting Linear Mixed-Effects Models  
515 using lme4. *J Stat Softw* 67(1):1–48.
- 516 49. Kuznetsova A, Brockhoff PB, Christensen RHB (2017) lmerTest package: tests in linear  
517 mixed effects models. *J Stat Softw* 82(13).
- 518
- 519
- 520



**Table 1.** Models of ecological factors moderating the effect of spatial scale on biodiversity-disease relationships

	A) Spearman correlation coefficient			B) Pearson's skewness		
	<b>DF</b>	<b>F-value</b>	<b>p-value</b>	<b>DF</b>	<b>F-value</b>	<b>p-value</b>
Spatial extent	41.3	2.862	0.098	4.6	0.395	0.56
Human	65.0	2.171	0.145	57.6	0.061	0.81
Route	61.2	4.124	0.047	55.3	0.670	0.42
Macroparasite	30.5	1.335	0.257	12.3	0.152	0.70
Manipulative	61.2	0.664	0.418	14.2	0.020	0.89
Spatial extent × Human	42.5	2.140	0.151	15.6	0.007	0.93
Spatial extent × Route	69.9	2.464	0.121	87.7	0.523	0.47
Spatial extent × Macroparasite	34.3	0.385	0.539	23.2	1.041	0.32
Spatial extent × Manipulative	43.6	0.046	0.831	4.0	0.010	0.93

Type III Analysis of Variance Table with Satterthwaite's method

DF: Denominator degrees of freedom

521

522

523 **Figure legends.**

524 **Fig 1.** Hypothetical relationships between biodiversity and disease risk. A) A non-monotonic  
525 right-skewed distribution suggests that dilution might occur more frequently, but less intensely  
526 than amplification because the relationship is moderately negative over a greater portion of the  
527 biodiversity gradient than it is strongly positive. A non-monotonic left-skewed distribution  
528 suggest that amplification might occur more frequently but less intensely than dilution, because  
529 the relationship is moderately positive over a greater portion of the biodiversity gradient than it is  
530 strongly negative. A monotonic and asymptotic distribution suggests that amplification becomes  
531 increasingly moderate with biodiversity. B) In addition to the shape of biodiversity–disease  
532 relationships, the location on the curve where biodiversity levels are observed will also affect the  
533 likelihood and intensity of dilution and amplification. For example, in a left-skewed biodiversity-  
534 disease relationship, collecting measurements at biodiversity beyond the peak of parasite  
535 abundance could lead researchers to conclude that there is was a linear dilution effect, whereas  
536 measurements before the peak of parasite abundance would lead researchers to conclude that  
537 there was a linear amplification effect.

538

539 **Fig 2.** Results of the analysis comparing Spearman rank correlation to Pearson’s skewness.  
540 Points are model-estimated means and error bars are 95% confidence intervals. The colored  
541 points show the distribution of the raw data. Left-skewed relationships (Pearson’s  
542 skewness<0.25) are shown in red, right-skewed relationships (Pearson’s skewness>0.25) are  
543 shown in blue, and non-skewed relationships are shown in grey. Spearman rank correlation was  
544 strongly associated with Pearson’s skewness: monotonic amplification effects ( $\rho>0$ ,  $p<0.05$ )

545 tended to be left-skewed, monotonic dilution effects ( $\rho < 0$ ,  $p < 0.05$ ) were right skewed, and non-  
546 monotonic relationships were not significantly skewed.

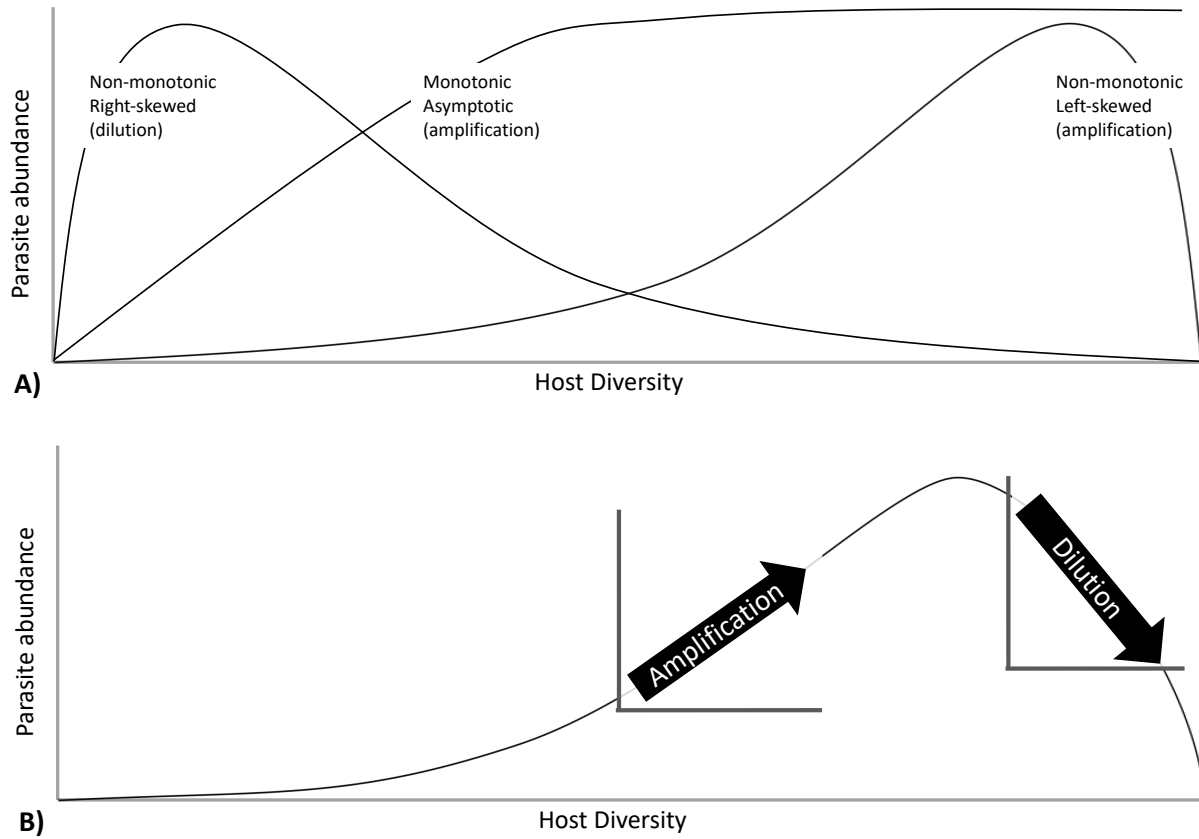
547

548 **Fig 3.** Results of the analyses relating spatial scale to the shape of the biodiversity-disease  
549 relationship. Points represent each published biodiversity-disease relationship, colored by their  
550 estimated shape (red = Monotonic amplification in panel A and left-skewed in panel B; blue =  
551 Monotonic dilution in panel A and right-skewed in panel B; grey = non-significant or non-  
552 monotonic in panel A non-skewed in panel B). Solid lines indicate the estimated fit of a  
553 multilevel random effects model, and grey ribbons indicate the 95% confidence intervals. Spatial  
554 scale moderates the relationship between biodiversity and disease: A) Spearman rank correlation  
555 between biodiversity and disease was positively associated with spatial extent, and B) Pearson's  
556 skewness was negatively associated with spatial extent.

557

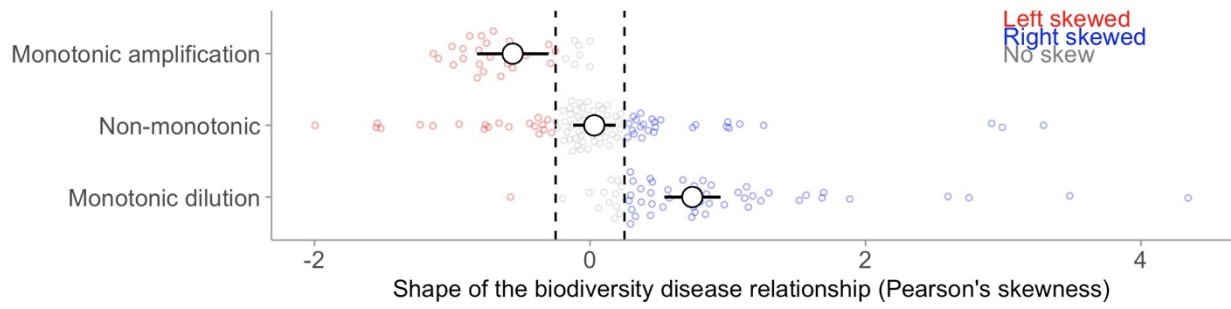
558 **Fig 4.** Results of constraining biodiversity-disease relationships to pass through the origin. The  
559 top two rows show Pearson's skewness for unconstrained curves, and curves that were  
560 constrained to pass through the origin, with each study connected by a solid line. Left-skewed  
561 relationships (Pearson's skewness  $< 0.25$ ) are shown in red, right-skewed relationships (Pearson's  
562 skewness  $> 0.25$ ) are shown in blue, and non-skewed relationships are shown in grey. The bottom  
563 row shows the model-estimated effect of constraining the curves to pass through the origin, with  
564 the point indicating the model-estimated mean, and error bars showing the 95% confidence  
565 interval. On average, constraining curves to pass through the origin results in a more left-skewed  
566 relationship between biodiversity and disease.

567



568

569 **Fig 1.**

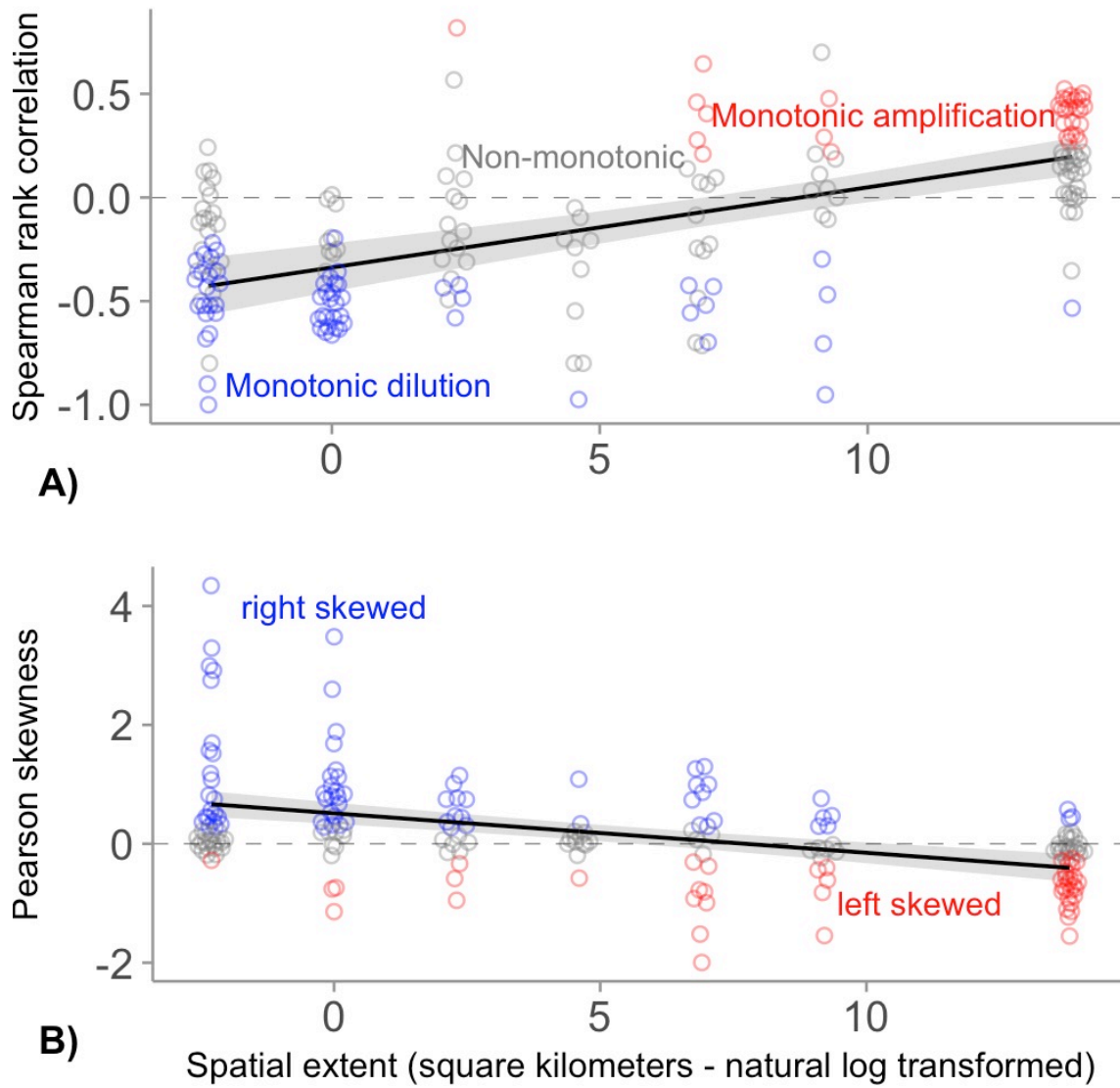


570

571 **Fig 2.**

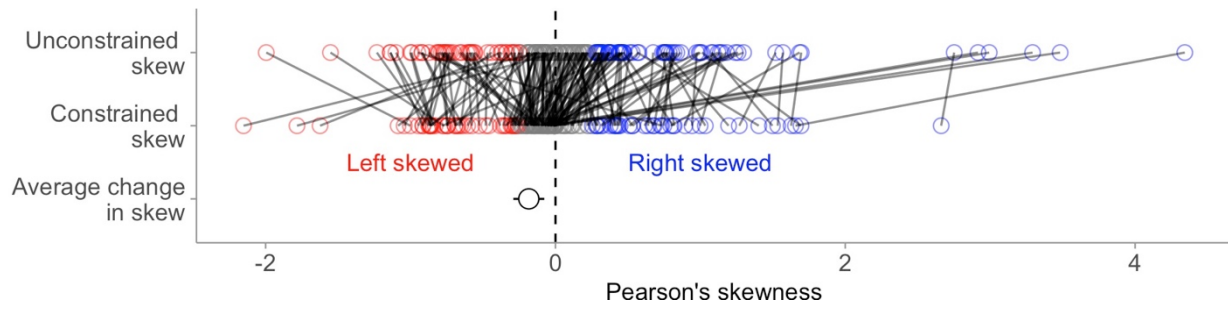
572

573



574

575 **Fig 3.**



576

577 **Fig 4.**

578