

1 ARTICLE

2 **Implications of predator species richness in terms of zoonotic spillover**  
3 **transmission of filoviral hemorrhagic fevers in Africa**

4

5 **Taehee Chang<sup>1</sup> (Orcid; 0000-0001-9224-6961), Sung-il Cho<sup>1,2</sup> (Orcid; 0000-0003-4085-1494),**  
6 **Kyung-Duk Min<sup>3\*</sup> (Orcid; 0000-0002-1000-2187)**

7 <sup>1</sup>Department of Public Health Sciences, Graduate School of Public Health, Seoul National University,  
8 Seoul, Republic of Korea

9 <sup>2</sup>Institute of Health and Environment, Seoul National University, Seoul, Republic of Korea

10 <sup>3</sup>College of Veterinary Medicine, Chungbuk National University, Republic of Korea

11

12 **Abstract**

13 Previous studies found that higher species richness of predators could reduce spillover risks of  
14 rodent-borne diseases. However, the effects on bat-borne diseases remains to be investigated. To this  
15 regard, we evaluated associations between predator species richness and the spillover events of  
16 *Ebolavirus* and *Marburgvirus*, the highly pathogenic bat-borne diseases in Africa. Stacked species  
17 distribution model approach was used to estimate predator species richness and Logistic regression  
18 analyses that considered spatiotemporal autocorrelations were conducted. The results showed that the  
19 third quartile (OR = 0.02, 95% CI 0.00–0.84) and fourth quartile (0.07, 0.00–0.42) of species richness  
20 of Strigiformes and the third quartile (0.15, CI 0.01–0.73) and fourth quartile (0.53, 0.03–0.85) of  
21 Colubridae showed significantly lower risks of spillover transmission of *Ebolavirus*. However, no  
22 significant association was found between predator species richness and *Marburgvirus* spillover. The  
23 results support a possible effect of predator species diversity on spillover suppression.

24

25

26

## 27 Introduction

28 *Ebolavirus* and *Marburgvirus* are non-segmented, negative-stranded RNA viruses belonging to the  
29 family Filoviridae, a subgroup of the order Mononegavirales (1). There are six virus species in the  
30 *Ebolavirus* genus (Ebola virus, Sudan virus, Bombali virus, Tai Forest virus, Bundibugyo virus, and  
31 Reston virus) and two species in the *Marburgvirus* genus (Marburg virus and Ravn virus) (2). With  
32 the exception of the Reston virus, the viruses are considered indigenous to Africa, where multiple  
33 human outbreaks have occurred (3, 4). Filovirus epidemics cause catastrophic losses of human and  
34 animal life given the high case fatality rates, which are typically 60–70% but can reach 90% (1).

35 Significant progress in shortening the list of potential filovirus reservoir hosts has been made  
36 during the past decade. Apart from the *Rousettus aegypticus* fruit bat, which has repeatedly tested  
37 positive for *Marburgvirus*, antibodies against various Ebola species have been found in at least 14  
38 other species of bats; however, only *Epomops franqueti*, *Hypsignathus monstrosus*, and *Myonycteris*  
39 *torquata* tested positive using PCR methods (5-8). The viruses may spread to other animals, including  
40 non-human primates, duikers (antelopes), or humans, from bat species shown to be vulnerable to  
41 filoviruses. Humans might contract the virus by handling or eating so-called bushmeat, such as  
42 roosting bats close to human dwellings, or via contact with infected mammalian bodily fluids (1).

43 Predators impact prey density, distribution, and behavior both directly and indirectly. Theoretically,  
44 such impacts might cascade to lower trophic levels and thus reduce the risk of zoonotic spillover (9,  
45 10). Generalist predators (e.g., certain snakes, cats, owls, and raptors) that are either non-specialized  
46 in terms of prey selection and can thus move among target species, or that are highly mobile and  
47 therefore wander in search of better hunting grounds, have been suggested to chronically suppress  
48 prey numbers and thus stabilize population dynamics (9, 11). Predator non-lethal effects can influence  
49 the behavioral patterns of prey and reduce prey fitness. The predatory risk cues detected by prey,  
50 including visual, auditory, or chemical signals, allow them to identify the presence of predators and  
51 consequently alter their behavior in response to the danger of predation (12, 13). Few vertebrate  
52 predators specialize in hunting bats, and bat predation appears to be mainly opportunistic in nature.  
53 However, generalist and opportunistic predators may exert substantial effects on bat ecology,  
54 eventually reducing the rate of contact between reservoir hosts and humans and thus mitigating the  
55 risk of zoonotic spillover (9, 12, 14).

56 We hypothesized that high predator species richness will reduce the zoonotic spillover of  
57 filoviruses in Africa. We examined the associations between predator species richness and historical

58 spillovers of *Ebolavirus* and *Marburgvirus* based on distributional data from known predators of bats  
59 only, as well as satellite-derived environmental data.

60

## 61 **Methods**

### 62 **Study design and study area**

63 In this ecological study, we used stacked species distribution models and the maximum entropy  
64 method (Maxent modeling) to calculate the number of predator species. We considered potential  
65 confounding factors when conducting logistic regression analyses of the relationship between predator  
66 species richness and spillover risk. We included all African countries with at least one reported human  
67 case of Ebola or Marburg infection. We confined the study regions to areas proposed in previous  
68 studies to harbor the reservoir species *E. franqueti*, *H. monstrosus*, and *M. torquata* of *Ebolavirus* and  
69 the *R. aegypticus* fruit bat of *Marburgvirus* (5-8). The distribution ranges were constructed using the  
70 geographical database of the International Union for the Conservation of Nature (IUCN) (15). To  
71 examine the relationship between predator species richness and filovirus cases, three datasets were  
72 compiled: (i) a comprehensive list of index case locations, (ii) geographical information on the  
73 distributions of predators and reservoir hosts, and (iii) environmental factors suggested to be  
74 ecologically significant. R software (v. 4. 2. 1) (16) was used for all data processing and analyses. The  
75 “dismo” package (17) was employed to model species niches. Bayesian parameter estimations that  
76 considered spatial and spatiotemporal autocorrelations were conducted using the “CARBayes” and  
77 “CARBayesST” packages (18).

### 78 **Outcome definitions**

79 We identified index cases and rebuilt zoonotic spillover cases in both space and time. We searched  
80 the formal scientific literature using PubMed and the Web of Science for data on all historical  
81 filovirus outbreaks (3-8, 19). We sought to recreate the outbreaks in detail and locate the most likely  
82 index cases, thus infected humans who had interacted with disease-causing non-human sources. Cases  
83 reported between 2000 and 2021 were included in analysis because the environmental covariates used  
84 in the present report share their temporal ranges since that time. On the map of the study regions, we  
85 generated  $1^\circ \times 1^\circ$  grids and classified them in terms of their intersections with the point locations of  
86 index cases.

## 87 **Niche modeling and diversity maps**

88 The suitability of habitats for natural predators of bats, i.e., the order Accipitriformes, Strigiformes,  
89 and Carnivora and family Colubridae and reservoir hosts in the order Chiroptera, was predicted using  
90 a simple species distribution modeling strategy (also termed ecological niche modeling), which  
91 integrates the reported occurrences of species with local climatic and geographic information. We  
92 used a maximum entropy approach (Maxent modeling) (20); this is one of the most widely used  
93 models when identifying species distributions. The approach employs presence-only data, which are  
94 helpful when modeling small and mobile species because it is (appropriately) challenging to establish  
95 their absence.

96 The occurrence data of included species within the study area (thus the African continent: 12.69 to –  
97 22.42 N and 41.57 to –14.97 E) were those of the Global Biodiversity Information Facility. Species  
98 that occurred at more than 10 points were included in the models. Climatic and geographical data  
99 served as predictive variables when simulating the distributions of the species. The bioclimatic  
100 variables were derived from WorldClim ver. 2.0 (21) and the elevation data from the Shuttle Radar  
101 Topography Mission (ver. 4) (22) with a spatial resolution of 2.5 minutes (~20 km<sup>2</sup>). The variables  
102 included in the modeling process were considered ecologically crucial in terms of species distribution,  
103 and they evidenced one-to-one intercorrelations < 0.7. These variables were the mean diurnal range  
104 (Bio02), temperature seasonality (Bio04), maximum temperature in the warmest month (Bio05),  
105 precipitation in the wettest quarter (Bio16), precipitation in the warmest quarter (Bio18), precipitation  
106 in the coldest quarter (Bio19), and the elevation. When fitting the models for each species, all  
107 variables were verified using the Jackknife test (20).

108 We used k-fold cross-validation to assess the models. Next, we converted the habitat suitability into  
109 a binary value (suitable habitat 1; unsuitable habitat 0). The threshold was the modeled prevalence  
110 closest to the observed prevalence. The numbers of species for which suitable habitat pixels in each  
111 grid exceeded 50% of the total grid areas were counted.

## 112 **Data acquisition and preprocessing**

113 Global climatic data from 1970 to 2000 were collected from WorldClim ver. 2.0 (21), which features  
114 a spatial resolution of 2.5 minutes (~ 20 km<sup>2</sup>) and offers monthly average precipitation and  
115 temperature data in raster format. The average annual precipitation and temperature for each grid  
116 region were computed.

117 The geographical confounding factors collected included elevation land cover, agricultural land use,  
118 and forest cover data. We obtained elevation data from the Shuttle Radar Topography Mission (22),  
119 which offers 90-m-scale worldwide elevation data in raster format. The values for each grid were  
120 averaged. Data on agricultural land use during 2000–2021 were gathered in raster format (23). The  
121 dataset contains the most likely International Geosphere-Biosphere Programme class for each  
122 0.05°pixel, and we calculated the proportion of each International Geosphere-Biosphere Programme  
123 class for all grids of interest. The Global Forest Change (GFC) (24) data yielded forest cover  
124 information. In terms of tree canopy cover, the likelihood of a tree canopy is presented in raster  
125 format and ranges from 0 to 100. We used 75 as the cutoff when determining whether a raster cell  
126 included a forest. We collected data on forest loss, defined as a change from a forest to a non-forest  
127 state, during 2000–2021. We then computed the proportion of forest coverage in each grid by  
128 subtracting the area of forest loss from that of the tree canopy cover.

129 The sociodemographic factors analyzed were the gross domestic product, human development index,  
130 population density, and human footprint score. Gross domestic product and human development index  
131 data from 1990–2015 were gathered, and the average values for each grid computed (25). The impact  
132 of human activity on the environment during 2000–2018 was measured using the human footprint  
133 score, which presents more significant anthropogenic pressures as higher scores (26). The values for  
134 each grid were averaged. Population density data were acquired from the WorldPop website (27). A  
135 population count dataset of the unconstrained global mosaics from 2000–2020 at a resolution of 1 km  
136 was used to calculate the population density for each grid.

137 Some variables did not cover the entire period from 2000 to 2021; in such cases, data from previous  
138 years were used to fill in for missing data. Detailed descriptions of each variable, including the  
139 temporal range and spatial resolution, can be found in (Supplementary Table 1).

## 140 **Statistical analysis**

141 We developed logistic regression models for the study grids to determine odds ratios (ORs) with 95%  
142 confidence intervals (95% CIs) for the relationship between predator species richness and filovirus  
143 cases. We adjusted for all possible confounders except for variables with a Pearson correlation  
144 coefficient  $> 0.7$  and variance inflation factor (VIF)  $> 10$ . To handle potential species richness  
145 overestimation, the indicators of predator species richness were entered into saturated models as  
146 categorical variables. We defined the Ebolavirus categories by quartiles and the Marburgvirus

147 categories by medians. The categories with the lowest number of predator species served as the base  
148 categories.

149 We used Bayesian spatiotemporal models to derive the spatial and temporal patterns over 100,000  
150 iterations with a burn-in of 95,000 when the model residuals were autocorrelated as revealed by the  
151 Moran I test and Durbin–Watson test. The following are the mathematical expressions of the models:

152 Model 1:  $\ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta X + v$

153 Model 2:  $\ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta X + v + u$

154 Model 3:  $\ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta X + v + u + a_1 t$

155 Model 4:  $\ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta X + v + u + Y$

156 Model 5:  $\ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta X + v + u + Y + \varphi$

157 Where  $p$  is the probability of filovirus emergence; the constant  $\beta_0$  is the intercept;  $\beta_n$  is the regression  
158 coefficient;  $X_n$  is the set of predictive variables;  $v_n$  is the non-spatial random component for grid  $n$ ;  $u_n$   
159 is the structured spatial random component for grid  $n$ ;  $t$  is the temporal trend of the data with a  
160 constant term  $a_1$ ;  $Y_m$  is the temporal random-walk component; and the random effect  $\varphi_{mn}$  is the  
161 space–time interaction term. To choose the model affording the best performance in terms of the  
162 Bayesian framework, we compared Models 1–5 using the deviance information criterion (DIC) and  
163 the Watanabe–Akaike information criterion (WAIC).

## 164 **Sensitivity analysis**

165 We performed sensitivity analysis using the “R-INLA” package for Bayesian parameter estimation  
166 (28). We estimated species richness using a geographical database from the IUCN website (15). Only  
167 species categorized as ‘extant’ that overlapped the species included in Maxent modeling were used.  
168 The numbers of species, the distribution ranges of which spanned more than 50% of a specific grid,  
169 were counted after intersecting the polygons representing the range data for the species. Finally, we  
170 constructed models using the species richness variables calculated via Maxent modeling and included  
171 only species reported to prey on bats.

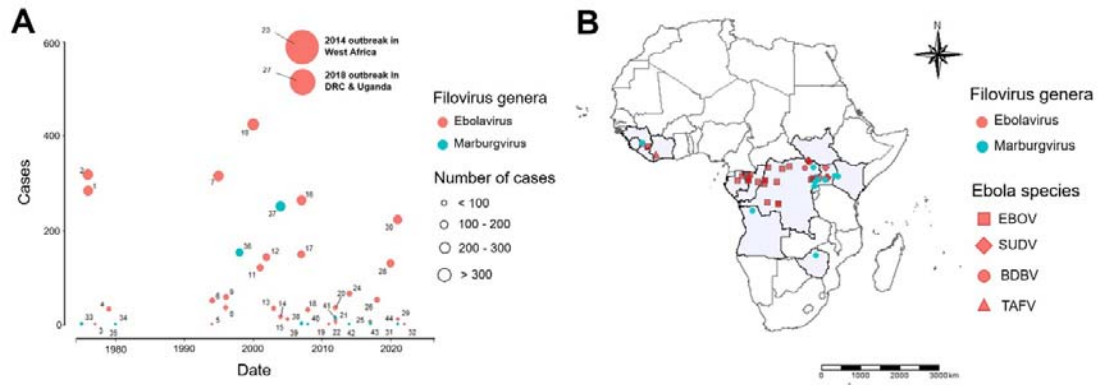
172

## 173 **Results**

### 174 **Processing datasets**

175 In total, 32 index cases of *Ebolavirus* and 12 of *Marburgvirus* were identified across the African  
176 continent. The times of disease occurrence span the last four decades, starting with the first  
177 *Marburgvirus* case in 1976 (Figure 1A). The locations of the outbreaks spanned from Guinea in West  
178 Africa to Uganda and Kenya in East Africa (Figure 1B). We classified study grids by their  
179 intersections with the point locations of index cases reported between 2000 and 2021 (Figure 2A, B).

180



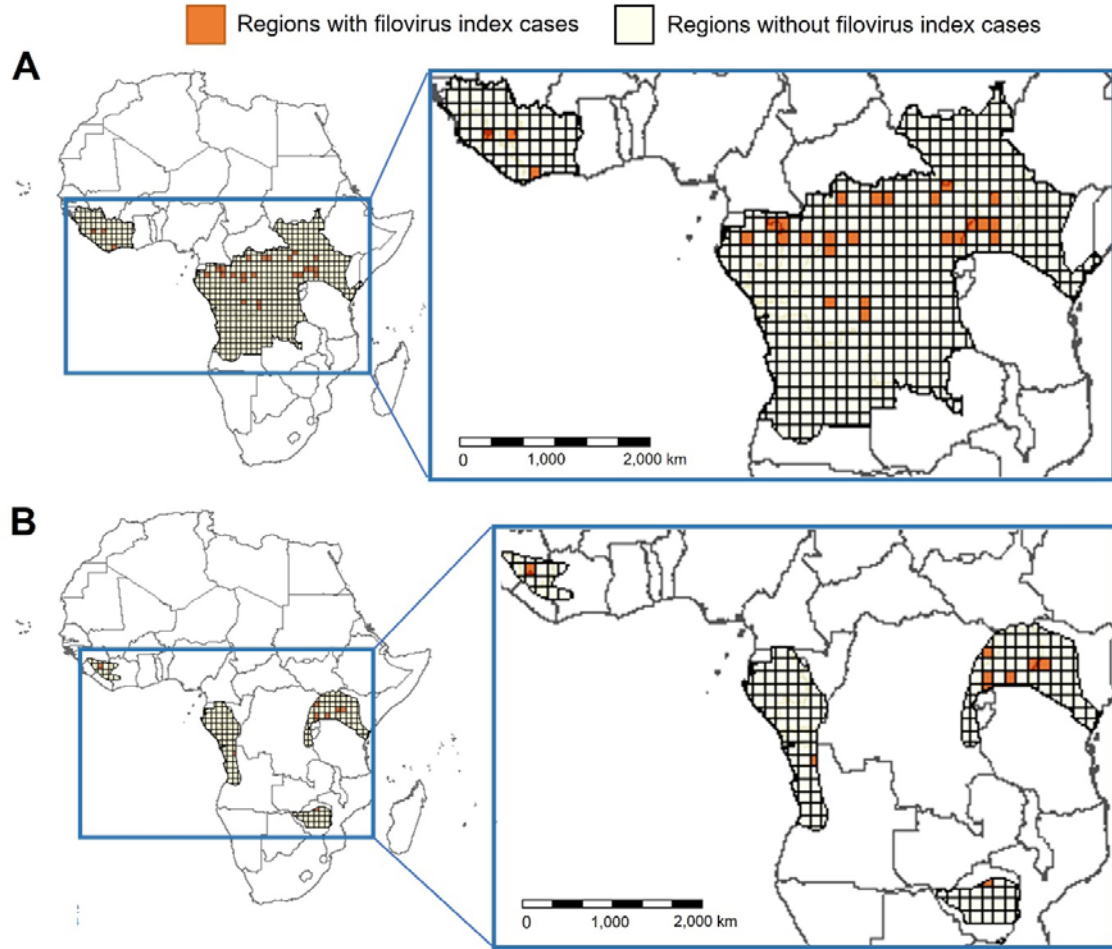
181

182 **Fig 1. The locations and points of occurrence of filovirus outbreaks in Africa.** (A) Shows the  
183 reported outbreaks of *Ebolavirus* and *Marburgvirus* through time, with its height along the y-axis  
184 reflect the number of cases. (B) Illustrates a map of the index cases for each outbreak, categorized by  
185 genus and species of the viruses. The map data of the African continent was employed to draw base  
186 maps in the figure. (Available from: <https://www.diva-gis.org/>).

187

188





189

190 **Fig 2. Study area and study units.** (A) Shows the study grids with and without *Ebolavirus* index  
191 cases. (B) Shows the study grids with and without *Marburgvirus* index cases. The grids with dark  
192 orange color represent that the region contains filovirus index cases. The study area is confined to  
193 countries with at least one filovirus case. We clipped the area with the reservoir species'  
194 distribution ranges. The map data of the African continent was employed to draw base maps in the  
195 figure. (Available from: <https://www.diva-gis.org/>).

196

197

198 **Descriptive analysis**

199 A total of 524 grids were used for *Ebolavirus* analysis and 197 for *Marburgvirus* analysis. Crude  
200 univariate analysis revealed certain characteristics of the study grids (Tables 1, 2). The predator  
201 species richness of the orders Accipitriformes and Strigiformes was lower in grids with *Ebolavirus*  
202 outbreaks than in those without (Table 1). Also, grids with *Marburgvirus* outbreaks evidenced lower  
203 species richness for the predator orders Accipitriformes, Strigiformes, and Carnivora and the family  
204 Colubridae (Table 2). All possible predictive variables were included in multivariate logistic models  
205 after screening variables with a Pearson correlation coefficient  $> 0.7$  (Supplementary Fig. 1-2). The  
206 VIF values were also calculated to assess model multicollinearity. No predictive variable had a VIF  $>$   
207 10.

208

209

210 **Table 1. Summary of descriptive analysis for *Ebolavirus* index cases.**

Variables	Not occurred (n = 504)	Occurred (n = 20)	Welch's t-test	Wilcox rank sum test
Avian species richness (order Accipitriformes)	5.42 ± 7.68	3.90 ± 6.11	0.29	0.27
Avian species richness (order Strigiformes)	1.46 ± 2.24	0.70 ± 1.17	< 0.05	< 0.05
Cat species richness (order Carnivora)	5.01 ± 5.39	5.25 ± 7.04	0.88	0.27
Snake species richness (family colubridae)	6.73 ± 6.70	7.40 ± 8.93	0.74	0.68
Bat species richness (order Chiroptera)	10.6 ± 9.17	9.65 ± 10.70	0.71	0.15
Human foot print score	8.15 ± 4.09	9.57 ± 5.48	0.26	0.41
Elevation	707.00 ± 402.00	701.00 ± 333.00	0.93	0.99
Precipitation (annual average)	69.30 ± 47.60	116.00 ± 25.20	< 0.001	< 0.001
Temperature (annual average)	23.70 ± 2.43	23.90 ± 1.39	0.69	0.88
Population density (per km <sup>2</sup> )	37.10 ± 66.80	65.60 ± 89.70	0.18	0.06
Gross Domestic Product (per capita)	2784.00 ± 3400.00	1733.00 ± 2134.00	< 0.05	< 0.05
Human Development Index	0.43 ± 0.07	0.41 ± 0.06	0.44	0.26
Forest cover (%)	30.70 ± 47.40	57.60 ± 45.80	< 0.05	< 0.001
Agricultural land use class (% of evergreen broadleaf forests)	27.50 ± 39.40	55.10 ± 43.60	< 0.05	< 0.001
Agricultural land use class (% of deciduous broadleaf forests)	1.07 ± 4.19	0.04 ± 0.19	< 0.05	0.06
Agricultural land use class (% of mixed forests)	0.79 ± 3.50	0.19 ± 0.83	< 0.05	0.22
Agricultural land use class (% of woody savannas)	12.10 ± 21.30	13.30 ± 17.50	0.78	0.46
Agricultural land use class (% of savannas)	33.70 ± 34.30	19.70 ± 26.80	< 0.05	< 0.05
Agricultural land use class (% of grasslands)	20.50 ± 32.30	0.60 ± 1.59	< 0.001	< 0.001
Agricultural land use class (% of permanent wetlands)	0.35 ± 1.36	0.52 ± 1.44	0.59	0.34
Agricultural land use class (% of croplands)	1.19 ± 4.71	0.37 ± 0.87	0.51	0.59

Agricultural land use class (% of urban and built-up lands)	$0.07 \pm 0.39$	$0.20 \pm 0.67$	0.41	0.14
Agricultural land use class (% of cropland/natural vegetation mosaics)	$1.58 \pm 7.87$	$9.84 \pm 16.80$	< 0.05	< 0.001

---

211

212

213 **Table 2. Summary of descriptive analysis for *Marburgvirus* index cases.**

Variables	Not occurred (n = 190)	Occurred (n = 7)	Welch's t-test	Wilcox rank sum test
Bat species richness (order Chiroptera)	9.09 ± 9.00	4.57 ± 3.10	< 0.05	0.11
Avian species richness (order Accipitriformes)	7.39 ± 10.10	1.14 ± 1.86	< 0.05	< 0.05
Avian species richness (order Strigiformes)	2.04 ± 3.05	0.14 ± 0.38	< 0.05	< 0.05
Cat species richness (order Carnivora)	4.82 ± 6.90	2.57 ± 2.90	0.11	0.19
Snake species richness (family colubridae)	5.27 ± 7.15	3.00 ± 2.52	< 0.05	0.47
Human foot print score	10.30 ± 4.47	13.90 ± 6.53	0.19	0.09
Elevation	778.00 ± 459.00	1084.00 ± 453.00	0.13	0.09
Precipitation (annual average)	64.70 ± 38.80	99.80 ± 32.60	< 0.05	< 0.05
Temperature (annual average)	23.00 ± 2.48	22.00 ± 2.02	0.24	0.21
Population density (per km <sup>2</sup> )	56.30 ± 83.90	157.00 ± 135.00	< 0.05	< 0.05
Gross Domestic Product (per capita)	4085.00 ± 4863.00	1657.00 ± 1536.00	< 0.05	< 0.05
Human Development Index	0.47 ± 0.08	0.42 ± 0.04	< 0.05	0.19
Forest cover (%)	27.10 ± 45.90	27.00 ± 37.00	0.92	0.75
Agricultural land use class 2 (% of evergreen broadleaf forests)	22.50 ± 34.80	24.40 ± 32.40	0.82	0.67
Agricultural land use class 4 (% of deciduous broadleaf forests)	0.38 ± 2.68	0.00 ± 0.00	< 0.05	0.23
Agricultural land use class 5 (% of mixed forests)	0.19 ± 1.52	0.00 ± 0.00	0.09	0.41
Agricultural land use class 8 (% of woody savannas)	8.71 ± 15.70	14.40 ± 19.20	0.46	0.57
Agricultural land use class 9 (% of savannas)	28.30 ± 31.50	28.40 ± 30.10	0.89	0.58
Agricultural land use class 10 (% of grasslands)	30.90 ± 32.10	4.50 ± 8.41	< 0.001	< 0.05

Agricultural land use class 11 (% of permanent wetlands)	0.26 ± 1.42	0.65 ± 0.93	0.32	0.29
Agricultural land use class 12 (% of croplands)	2.09 ± 6.79	3.61 ± 8.43	0.65	0.23
Agricultural land use class 13 (% of urban and built-up lands)	0.10 ± 0.42	0.46 ± 1.12	0.43	0.15
Agricultural land use class 14 (% of cropland/natural vegetation mosaics)	3.99 ± 9.20	23.20 ± 24.90	< 0.05	< 0.001

---

214

## 215 **Model selection and validation**

216 We constructed saturated models using all available predictive variables and evaluated those  
217 variables from an epidemiological perspective (Supplementary Fig. 3). To derive the association  
218 between predator species richness and the historical incidence of *Ebolavirus* and *Marburgvirus* in  
219 Africa from 2000 to 2021, we used the average values of the predictive variables and the total  
220 emergence counts for each grid over that period. At the grid level, the spatial dependencies of filovirus  
221 incidences were assessed using the Moran I statistic; the values for *Ebolavirus* and *Marburgvirus*  
222 were 0.06 and -0.03, respectively, when using a row-standardized neighborhood structure  
223 (Supplementary Fig. 4-5). The Moran I statistic indicated that only the *Ebolavirus* incidence had a  
224 statistically significant spatial or temporal dependency (Supplementary Fig. 4). The Durbin-Watson  
225 test results showed that the *Ebolavirus* incidence data were autocorrelated in terms of the residuals of  
226 the models, with p-values < 0.05 (Supplementary Table 2). The performances of the models in terms  
227 of spatial and spatiotemporal autocorrelations are shown (Supplementary Table 3). Smaller DIC and  
228 WAIC values indicate better performance. Based on these results, we fitted Models 1–5 to display the  
229 association between predator species richness and the historical incidence of *Ebolavirus*. For  
230 *Marburgvirus*, we fitted the model with the average values in line with the Moran I and Durbin-  
231 Watson test results.

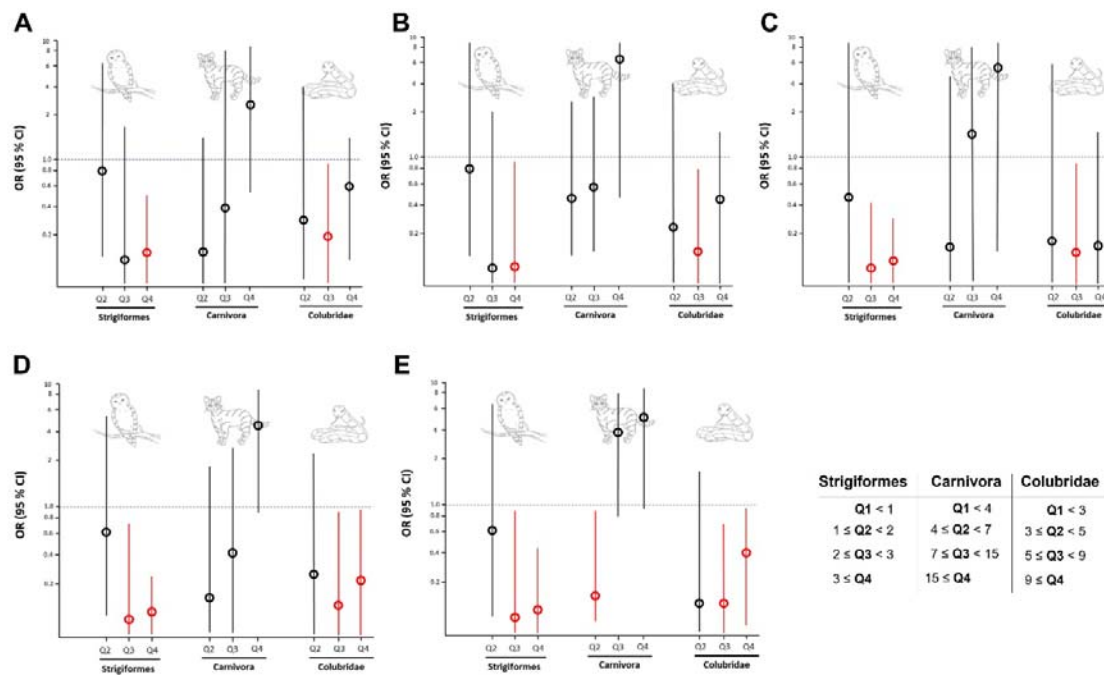
232

233 **Model-estimated association of predator species richness and zoonotic**  
234 **spillover of filoviruses**

235 The results of the final models are shown in Figures 3 and 4. The coefficients and ORs of all  
236 covariates are listed (Supplementary Table 4-8). Of all models, Model 2 for *Ebolavirus* exhibited the  
237 smallest DIC and WAIC in terms of spatial autocorrelation. In this model, the fourth quartile (OR =  
238 0.04, 95% CI 0.00–0.98) of Strigiformes species richness and the third quartile (OR = 0.15, 95% CI  
239 0.00–0.81) of Colubridae species richness exhibited significantly lower odds of *Ebolavirus* index  
240 cases (Figure 3B, Supplementary Table 5). This trend was maintained in Model 5, which showed the  
241 lowest DIC and WAIC of all models in terms of spatiotemporal autocorrelation. In this Model, the  
242 third quartile (OR = 0.02, 95% CI 0.00–0.84) and fourth quartile (OR = 0.07, 95% CI 0.00–0.42) of  
243 Strigiformes species richness, the third quartile (OR = 0.15, 95% CI 0.01–0.73) and fourth quartile  
244 (OR = 0.53, 95% CI 0.03–0.85) of Colubridae species richness, and the second quartile (OR = 0.23,  
245 95% CI 0.05–0.94) of Carnivora species richness evidenced significantly lower odds of *Ebolavirus*  
246 index cases (Figure 3E, Supplementary Table 8). However, none of the estimated parameters were  
247 significant for the other quartiles of Carnivora, Colubridae, and Strigiformes. In the models for  
248 *Marburgvirus*, we found no evidence of an association between predator species richness and  
249 *Marburgvirus* spillover (Figure 4, Supplementary Table 9). In addition, negative associations between  
250 predator species richness and *Ebolavirus* emergence were significant for some of the model quartiles  
251 when the “R-INLA” package was used, or when the species richness variables were calculated using  
252 the IUCN polygons (Supplementary Fig. 6, Supplementary Table 10-14; Supplementary Fig. 7,  
253 Supplementary Table 15-19). Models using the species richness variables including predator species  
254 reported to prey on bats did not support any significant association. (Supplementary Fig. 8,  
255 Supplementary Table 20-24). No significant association was revealed in the sensitivity analyses for  
256 *Marburgvirus* (Supplementary Fig. 9-10, Supplementary Table 25-26).

257

258



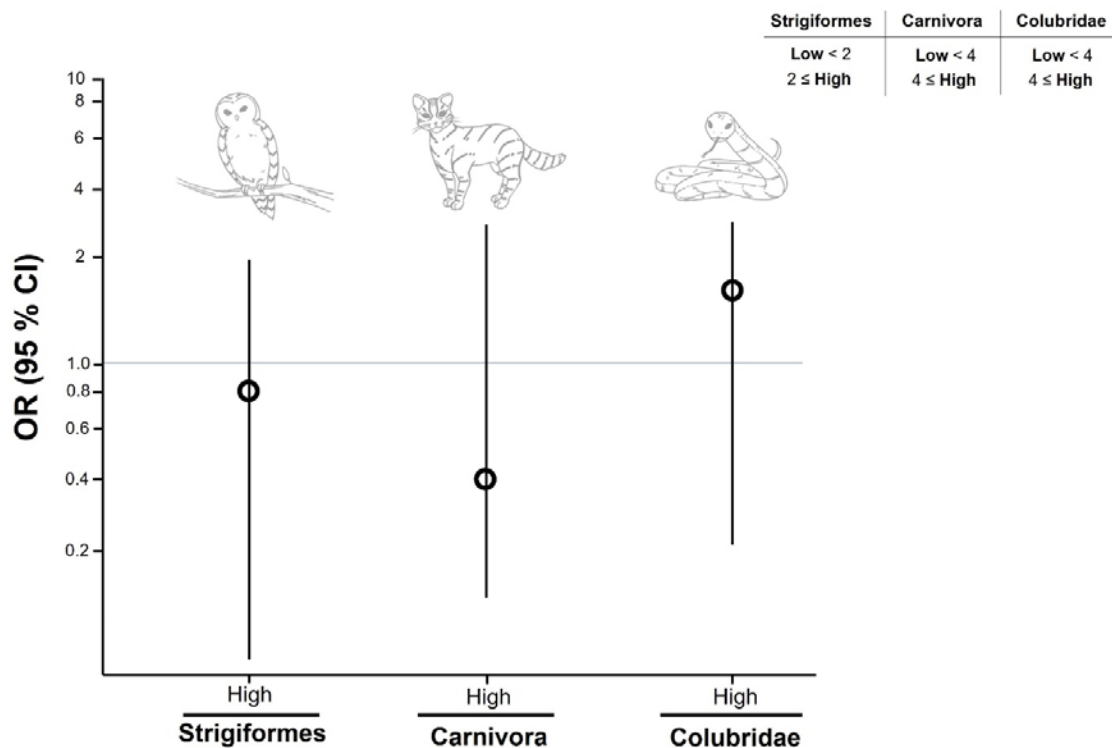
259

260 **Fig 3. Estimated ORs for *Ebolavirus* incidence according to the degree of species richness.** (A)  
 261 The result of Model 1. (B) The result of Model 2. (C) The result of Model 3. (D) The result of Model  
 262 4. (E) The result of Model 5. Model 2 and Model 5 were the best-fitting models with the greatest DIC  
 263 and WAIC, considering spatial and spatio-temporal autocorrelation, respectively. The dots indicate the  
 264 estimated ORs, with error bars representing the corresponding 95 % Wald's credible intervals. Red  
 265 means that the error bar does not intersect 1. The y-axis is shown on a logarithmic scale. The authors  
 266 generated draws of each predator.

267

268





269

270 **Fig 4. Estimated ORs for *Marburgvirus* incidence according to the degree of species richness.**

271 The dots indicate the estimated ORs, with error bars representing the corresponding 95 % Wald's  
272 credible intervals. The y-axis is shown on a logarithmic scale. The authors generated draws of each  
273 predator.

274

## 275 Discussion

276 We evaluated the association between predator species richness and filovirus spillover in Africa. The  
277 results showed that higher species richness in the in the order Strigiformes and family Colubridae was  
278 associated with lower odds of *Ebolavirus* spillover compared with that in regions with lower predator  
279 species richness. Regardless of the approach taken to calculate species diversity, this association was  
280 robust.

281 The negative association between predator species richness and the risk of *Ebolavirus* spillover  
282 suggests top-down regulation of *Ebolavirus* reservoir hosts (i.e., bats) by predators. The crucial roles  
283 played by predators in terms of the functional diversity of ecological communities and the control of

284 populations of disease reservoir hosts have been reported previously (9, 10, 12). The greater the  
285 predator species richness (i.e., the numbers of predator species within an area), the greater the cascade  
286 effect on prey species. The growing body of research on bat predation is slowly improving our  
287 understanding of bat predators and the effects of predation on bat populations (12-14). Natural bat  
288 predators may include birds, snakes, and mammals. Although few vertebrate predators are known to  
289 specialize on bats, and bat predation appears to be mostly opportunistic in nature, generalist and  
290 opportunistic predators may substantially impact bat ecology (9, 12, 14) via both direct predation and  
291 non-lethal cascade effects, also termed trait-mediated indirect interactions. Thus, predators control the  
292 abundance, density, and behavior patterns of prey species, eventually reducing the rate of contact  
293 between reservoir hosts and humans and thus mitigating the risk of zoonotic spillover (9). Such  
294 suppression is relatively strong in regions wherein ecological diversity is well-maintained.

295 The predator species richness of the order Strigiformes was significantly and negatively associated  
296 with the risk of *Ebolavirus* spillover. This is consistent with previous studies suggesting that owls are  
297 primary predators of bats (13, 14). Snakes are also supposed to prey on bats, via two strategies:  
298 positioning themselves near bat passage routes (i.e., near the entrances to bat roosts) and entering the  
299 refuges (12). Most such behaviors have been reported in tropical regions (29), perhaps because  
300 tropical bats roost by hiding among leaves or in open canopies that are accessible to most vertebrate  
301 predators. Bat predation is poorly understood; bats fly at night and hide by day. However, it appears  
302 that predation of bats by snakes in our study area is more significant than previously thought. More  
303 ecological research is required.

304 Predator species richness was not significantly associated with *Ebolavirus* cases in models that  
305 considered only the species reported to prey on bats. This may be attributable to a lack of information  
306 on all bat predators. Although the number of known predators is increasing, such research is limited  
307 by the ecological characteristics of bats, which render observations of predation difficult (12). Also,  
308 *Marburgvirus* occurrences were not consistently associated with predator species richness. The  
309 composition of bat species in the *Marburgvirus* regions may explain these results. Given the high bat  
310 diversity in the study region, *R. aegypticus*, the primary reservoir host of *Marburgvirus*, would not be  
311 the dominant bat species there. Therefore, the extent of predator richness may not have had any  
312 discernible effect on bat activities (30). Further studies of bat ecology, diversity, and abundance,  
313 especially of *R. aegypticus*, are needed.

314 Despite the strengths of this ecological study, several limitations should be noted. First, we estimated  
315 the diversity of predator species using stacked (aggregated) species distribution models. These models

316 may systematically overestimate site-level species richness (31). Therefore, we adjusted for bias using  
317 categorical values of predator species richness. Second, we did not include the temporal variations in  
318 species numbers from 2000 to 2021. However, such temporal changes can be ignored because most  
319 species considered are classified as IUCN “Least concern” (i.e., low risk of extinction). Third, when  
320 measuring species diversity, we simply calculated the numbers of species; we excluded the relative  
321 abundances of the predator species. Future research should employ other indicators of diversity such  
322 as the Simpson diversity index. Fourth, we considered only three bat species (*E. franqueti*, *H.*  
323 *monstrosus*, and *M. torquate*) that tested positive by PCR as primary reservoir hosts of *Ebolavirus*.  
324 Other probable reservoirs (bat species positive using serological methods) should be included in  
325 future studies. Finally, our study units were  $1^\circ \times 1^\circ$  grids; the use of a different scale (such as  $0.5^\circ \times$   
326  $0.5^\circ$ ) could have affected the results. This is the well-known modifiable area unit problem.

327 The world is still struggling to exit the unprecedented COVID-19 pandemic. It is predicted that the  
328 probability of pandemics caused by spillovers may increase in the coming decades, given the tectonic  
329 shifts in climate change and anthropogenic environmental degradation. However, although  
330 environmental and biodiversity changes may affect the spread of zoonotic diseases via various  
331 mechanisms, prevention of outbreaks still depends on containment, i.e., human disease surveillance,  
332 vaccines, and therapeutics. Here, we suggest that predator species richness may play a crucial role in  
333 mitigating the risk of filovirus spillover. Therefore, attempts to reduce the impacts of zoonotic  
334 diseases on public health should incorporate the concept of conservation epidemiology when deriving  
335 sustainable solutions that both maintain biodiversity and prevent zoonotic spillover, benefiting both  
336 humans and the environment.

337

338

339

340 **Data availability**

341 The datasets used and/or analysed during the current study are available from the  
342 corresponding author on reasonable request.

343

344 **Code availability**

345 We share the R codes on <https://github.com/TaeHChang/R-codes-for-paper-1>

346

347 **Acknowledgements**

348 This research was supported by a National Research Foundation of Korea (NRF) grant funded by the  
349 Korea government (MSIT) (no. NRF-2021R1C1C2012611). The funders had no role in the design and  
350 conduct of the study; collection, management, analysis, and interpretation of the data; preparation,  
351 review, or approval of the paper; and decision to submit the paper for publication.

352

353 **Author contributions**

354 K.D.M. and S.C. conceived, designed, and supervised the study. T.C. collected and analyzed the data.  
355 T.C. wrote the drafts of the paper. K.D.M. and S.C. commented on and revised drafts of the paper. All  
356 authors read and approved the final report.

357

358 **Competing interests**

359 The authors declare no competing interests.

360

361

362

363

## 364 **References**

- 365 1. Feldmann H, Geisbert TW. Ebola haemorrhagic fever. *Lancet*. 2011;377(9768):849-62.  
366 2. Languon S, Quaye O. Impacts of the Filoviridae family. *Curr Opin Pharmacol*. 2021;60:268-74.  
367 3. Ebola Virus Disease Distribution Map: cases of Ebola Virus Disease in Africa Since 1976  
368 [Internet]. 2022. Available from: <https://www.cdc.gov/vhf/ebola/history/distribution-map.html>  
369 (accessed Aug 1, 2022).
- 370 4. History of Marburg Virus Disease (MVD) Outbreaks: Known Cases and Outbreaks of Marburg  
371 virus disease, in Chronological Order [Internet]. 2022. Available from:  
372 <https://www.cdc.gov/vhf/marburg/outbreaks/chronology.html> (accessed Aug 1, 2022).
- 373 5. Koch LK, Cunze S, Kochmann J, Klimpel S. Bats as putative Zaire ebolavirus reservoir hosts  
374 and their habitat suitability in Africa. *Sci Rep*. 2020;10(1):14268.
- 375 6. Languon S, Quaye O. Filovirus Disease Outbreaks: A Chronological Overview. *Virology (Auckl)*.  
376 2019;10:1178122X19849927.
- 377 7. Pigott DM, Golding N, Mylne A, Huang Z, Henry AJ, Weiss DJ, et al. Mapping the zoonotic  
378 niche of Ebola virus disease in Africa. *Elife*. 2014;3:e04395.
- 379 8. Pigott DM, Golding N, Mylne A, Huang Z, Weiss DJ, Brady OJ, et al. Mapping the zoonotic  
380 niche of Marburg virus disease in Africa. *Trans R Soc Trop Med Hyg*. 2015;109(6):366-78.
- 381 9. Ostfeld RS, Holt RD. Are predators good for your health? Evaluating evidence for top-down  
382 regulation of zoonotic disease reservoirs. *Frontiers in Ecology and the Environment*.  
383 2004;2(1):13-20.
- 384 10. O'Bryan CJ, Braczkowski AR, Magalhaes RJS, McDonald-Madden E. Conservation  
385 epidemiology of predators and scavengers to reduce zoonotic risk. *Lancet Planet Health*.  
386 2020;4(8):e304-e5.
- 387 11. Klemola T, Tanhuanpää M, Korpimäki E, Ruohomäki K. Specialist and generalist natural  
388 enemies as an explanation for geographical gradients in population cycles of northern  
389 herbivores. *Oikos*. 2002;99(1):83-94.
- 390 12. Lima SL, O'Keefe JM. Do predators influence the behaviour of bats? *Biological Reviews*.  
391 2013;88(3):626-44.
- 392 13. Breviglieri CP, Piccoli GC, Uieda W, Romero GQ. Predation-risk effects of predator identity on  
393 the foraging behaviors of frugivorous bats. *Oecologia*. 2013;173(3):905-12.
- 394 14. Mikula P, Morelli F, Lučan RK, Jones DN, Tryjanowski P. Bats as prey of diurnal birds: a global  
395 perspective. *Mammal Review*. 2016;46(3):160-74.
- 396 15. The IUCN Red List of Threatened Species 2022 [Internet]. 2022. Available from:  
397 <https://www.iucnredlist.org/resources/spatial-data-download> (accessed Aug 5, 2022).
- 398 16. R Core Team. R: A language and environment for statistical computing. 2020.
- 399 17. Robert J. Hijmans SP, John Leathwick, Jane Elith. Dismo: Species Distribution Modeling. R  
400 Package Version 1.0-12. 2015.
- 401 18. Lee D. CARBayes: An R Package for Bayesian Spatial Modeling with Conditional  
402 Autoregressive Priors. *Journal of Statistical Software*. 2013;55(13):1-24.
- 403 19. Peterson AT, Samy AM. Geographic potential of disease caused by Ebola and Marburg viruses  
404 in Africa. *Acta Tropica*. 2016;162:114-24.

- 405 20. Phillips SJ, Anderson RP, Schapire RE. Maximum entropy modeling of species geographic  
406 distributions. *Ecological Modelling*. 2006;190(3):231-59.
- 407 21. Fick SE, Hijmans RJ. WorldClim 2: new 1-km spatial resolution climate surfaces for global land  
408 areas. *International Journal of Climatology*. 2017;37(12):4302-15.
- 409 22. Hole-filled seamless SRTM data V4, International Centre for Tropical Agriculture (CIAT)  
410 [Internet]. 2008. Available from: <http://srtm.csi.cgiar.org> (accessed Aug 6, 2022).
- 411 23. MCD12C1 MODIS/Terra+Aqua Land Cover Type Yearly L3 Global 0.05Deg CMG [Internet].  
412 2018. Available from: <http://doi.org/10.5067/MODIS/MCD12C1.006> (accessed Aug 6, 2022).
- 413 24. Hansen MC, Potapov PV, Moore R, Hancher M, Turubanova SA, Tyukavina A, et al. High-  
414 Resolution Global Maps of 21st-Century Forest Cover Change. *Science*. 2013;342(6160):850-3.
- 415 25. Kummu M, Taka M, Guillaume JHA. Gridded global datasets for Gross Domestic Product and  
416 Human Development Index over 1990–2015. *Scientific Data*. 2018;5(1):180004.
- 417 26. Mu H, Li X, Wen Y, Huang J, Du P, Su W, et al. A global record of annual terrestrial Human  
418 Footprint dataset from 2000 to 2018. *Scientific Data*. 2022;9(1):176.
- 419 27. WorldPop, open data for spatial demography [Internet]. 2018. Available from:  
420 [www.worldpop.org](http://www.worldpop.org) (accessed Aug 6, 2022).
- 421 28. Blangiardo M, Cameletti M. *Spatial and Spatio-temporal Bayesian Models with R* INLA.  
422 John Wiley & Sons. 2015. p. 235-58.
- 423 29. Barti L, Péter Á, Csősz I, Sándor AD. Snake predation on bats in Europe: new cases and a  
424 regional assessment. *Mammalia*. 2019;83(6):581-5.
- 425 30. Saleem M, Fetzter I, Dormann CF, Harms H, Chatzinotas A. Predator richness increases the  
426 effect of prey diversity on prey yield. *Nat Commun*. 2012;3:1305.
- 427 31. Hortal J, De Marco Jr P, Santos AMC, Diniz-Filho JAF. Integrating biogeographical processes  
428 and local community assembly. *Journal of Biogeography*. 2012;39(4):627-8.