

RESEARCH ARTICLE

Traits, phylogeny and host cell receptors predict *Ebolavirus* host status among African mammals

Mekala Sundaram^{1*}, John Paul Schmidt², Barbara A. Han³, John M. Drake^{2,4}, Patrick R. Stephens¹

1 Department of Integrative Biology, Oklahoma State University, Stillwater, Oklahoma, United States of America, **2** Odum School of Ecology, University of Georgia, Athens, Georgia, United States of America, **3** Cary Institute of Ecosystems Studies, Millbrook, New York, United States of America, **4** Center for the Ecology of Infectious Diseases, University of Georgia, Athens, Georgia, United States of America

* mekala.sundaram@okstate.edu



OPEN ACCESS

Citation: Sundaram M, Schmidt JP, Han BA, Drake JM, Stephens PR (2022) Traits, phylogeny and host cell receptors predict *Ebolavirus* host status among African mammals. *PLoS Negl Trop Dis* 16(12): e0010993. <https://doi.org/10.1371/journal.pntd.0010993>

Editor: Alexander Bukreyev, University of Texas Medical Branch / Galveston National Laboratory, UNITED STATES

Received: July 21, 2022

Accepted: November 28, 2022

Published: December 21, 2022

Copyright: © 2022 Sundaram et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All supporting information methods and results provided in [Supporting Information](#) document. All raw data including mortality and infection databases, NPC1 alignment and accession numbers and predictions of reservoir statuses provided on figshare: <https://doi.org/10.6084/m9.figshare.20250408.v1>.

Funding: This work was funded by NIH R01AI156866 “Spillover of Ebola and other filoviruses at ecological boundaries”, awarded by

Abstract

We explore how animal host traits, phylogenetic identity and cell receptor sequences relate to infection status and mortality from ebolaviruses. We gathered exhaustive databases of mortality from *Ebolavirus* after exposure and infection status based on PCR and antibody tests. We performed ridge regressions predicting mortality and infection as a function of traits, phylogenetic eigenvectors and separately host receptor sequences. We found that mortality from *Ebolavirus* had a strong association to life history characteristics and phylogeny. In contrast, infection status related not just to life history and phylogeny, but also to fruit consumption which suggests that geographic overlap of frugivorous mammals can lead to spread of virus in the wild. Niemann Pick C1 (NPC1) receptor sequences predicted infection statuses of bats included in our study with very high accuracy, suggesting that characterizing NPC1 in additional species is a promising avenue for future work. We combine the predictions from our mortality and infection status models to differentiate between species that are infected and also die from *Ebolavirus* versus species that are infected but tolerate the virus (possible reservoirs of *Ebolavirus*). We therefore present the first comprehensive estimates of *Ebolavirus* reservoir statuses for all known terrestrial mammals in Africa.

Author summary

Identifying the animal hosts of *Ebolavirus* is crucial to preventing future outbreaks. We gathered exhaustive databases of which species die and which species show evidence of past infection from ebolaviruses in published literature. Our approach allowed us to differentiate which species show high mortality from *Ebolavirus* and which species tolerate the infection after exposure. We found that fruit bats are likely reservoirs as they are exposed to the infection but tolerate the virus, whereas primates do not serve as ideal reservoirs because they succumb to the infection once exposed. We also compared different

National Institute of Health (<https://www.nih.gov>) to authors PRS, JPS and JD. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

predictors of infection and conclude that receptors of *Ebolavirus* best predict infection in bats whereas ecological traits predict infection in primates.

Introduction

Ebolaviruses are zoonotic pathogens causing deadly hemorrhagic fever in human and animal populations [1]. Spillover of ebolaviruses into humans have occurred at least since 1976 with known and suspected index cases coming into contact with a wide range of possible animal hosts through hunting, transportation and eating of wild caught mammals [1–3]. Source species implicated in individual outbreaks include gorillas (*Gorilla gorilla*), chimpanzees (*Pan troglodytes*), baboons (*Papio anubis*) and several bats [4–6]. In the last few decades, much research has focused on the wild source of ebolaviruses [7–11].

Though the source of initial infection in several previous outbreaks [1,2,12], gorillas [13] and chimpanzees [14] show extremely high mortality when infected, and are, therefore, not considered to be a major source of transmission to other species [9,15]. While efforts have been made to identify a definitive mammal or arthropod reservoir [7,8,10,16], no reservoir species with high seroprevalence has yet been identified [11]. Rather than a single reservoir species, a network of maintenance hosts may be supporting circulation of ebolaviruses in the wild [11,17,18]. However, no prior study has attempted to estimate the host range of ebolaviruses across all known African mammals.

In previous work, Schmidt et al. [19] used machine learning to explore traits associated with variation in *Ebolavirus* infection status for 119 species sampled in the wild. This study found a high probability of infection for Pteropodid bats, primates and artiodactyls. Similarly, Han et al. [20] estimated the potential global host range of the Filoviridae in bats. While neither study directly incorporated phylogenetic information into models, Schmidt et al. [19] found statistically significant phylogenetic signal in *Ebolavirus* infection status. Moreover, phylogeny has been shown to be a consistent predictor of pathogen sharing among host species in other systems [21–23]. Olivero et al. [24] also concluded that Pteropodid bats, Molossid bats, primates and ungulates were phylogenetically close to known hosts and geographically linked to *Ebolavirus* outbreaks. No study thus far has quantified species-level differences in response to *Ebolavirus* infection, such as differences in susceptibility and mortality. Some species that are likely to test positive for infection tend to show high mortality when infected, while others are able to tolerate infection [19]. Presumably, the latter tolerant group includes the most important reservoirs in the wild such as Pteropodids [11,17,25] as these species will likely sustain the virus in the wild for long periods of time [11].

Intriguingly, a recent laboratory study also suggested that Niemann Pick C1 (hereafter ‘NPC1’) protein sequences may affect variation in infectivity of ebolaviruses at the cellular level [26–28], suggesting a molecular basis for host range. NPC1 is a transmembrane protein, which when functionally impaired in humans leads to lipid accumulation in cellular lysosomes and causes fatal Niemann Pick disease [29]. NPC1 has also been identified as a filovirus receptor fusing to the glycoproteins of filovirus envelopes and facilitating cell infection [26,30]. In a laboratory study of cell lines of two species of bats (specifically ZFBK13-76E and FBKT1 cell lines) and humans (HEK293T cell line), Takadate et al. [26] showed that specific amino acid residues in the loop-1 and loop-2 regions of NPC1 confer resistance to African filoviruses (*Marburgvirus* and *Ebolavirus*) by reducing cellular binding affinity of virus glycoproteins and inhibiting infection. Because of the genetic disorder that it can cause, there is much research interest in NPC1 (e.g., [31–35]), and protein sequence data for more than 100 species of

mammal are available in GenBank for the binding region that Takadate et al. [26] showed to be important [36]. Whether NPC1 can predict the likelihood of infection in the wild for species sampled for ebolaviruses has not been tested.

Here, we compile data on African *Ebolavirus* infection (from antibody and PCR tests) and mortality in mammals. We statistically model infection status and post-infection mortality using mammal species traits and phylogenetic relationships. We also model infection status as a function of NPC1 amino acid sequence data. Finally, we use our best models to predict the likely host range of ebolaviruses across African mammals, differentiating species that fit the profile of secondary amplifying hosts (i.e., that succumb quickly when infected) from better primary reservoir candidates (i.e., species that do not succumb to infection), providing a quantitative assessment of the host species most likely to be involved in maintaining circulation of ebolaviruses in Africa.

Methods

Here we outline research materials and methods. See “Supplemental Materials and Methods” (S1 Text) for additional details and rationale. We gathered exhaustive databases of species mortality after exposure to wild type strains of African ebolaviruses and species infection statuses determined from antibody and PCR tests in field survey studies. We created a binary variable of 1 for high mortality after exposure and 0 for little or no mortality. We created a second binary variable of 1 for positive infection status detected from PCR and antibody tests and 0 for species with only negative test results.

We chose life history traits describing ‘slow’ pace-of-life (or slow development and long gestation) vs ‘fast’ pace-of-life (or quick development and reproduction), which have been shown to be important in past studies of infection [19,20], from a near-complete imputed database of mammal traits [37]. We further chose brain mass as a trait representing life history tradeoffs in mammals [38]. Our final list included adult mass (g), brain mass (g), maximum longevity (d), age at first reproduction, gestation length (d), litter size, litters per year and traits reflecting variation in diet including percent diet comprised of scavenged meat, grain, fruit, and plant material. We also included distance of geographic range to a spillover site (m; computed as distance of IUCN range to nearest Schmidt et al. [39] spillover site) and a binary variable of 1 for volant and 0 for non-volant to distinguish bats from other species. For infection status models, we also used summed numbers of individuals sampled across all studies as a measure of sampling effort. To incorporate phylogenetic information, we used the maximum clade credibility tree from a recently published phylogenetic study of all mammals [40], and repeated analyses with a random sample of 100 alternative trees from the Bayesian posterior distribution of possible trees. For each tree we estimated phylogenetic eigenvectors using R package ‘PVR’ [41]. We included the first 48 eigenvectors, which captured 75% of total variation in the phylogeny, in models. We calculated phylogenetic signal in host mortality and infection status using Fritz and Purvis’ D [42] implemented in R package ‘caper’ and tested for significance using null models assuming no phylogenetic structure and a random Brownian process.

We predicted death of mammalian hosts using a logistic ridge regression analysis modified for small sample sizes. The ridge method is a penalized regression approach that typically performs well with correlated predictors [43,44]. We used a modified procedure for selecting the ridge parameter intended for models with small sample sizes [44,45], implemented with logisticRidge function in R package ‘ridge’ [46]. Analyses where the number of predictors greatly exceeds the number of observations are commonplace in genetic studies [44], and this method has been used in previous studies with as few as eight observations [47]. We predicted death of host as a function of pace-of-life traits and the first 48 phylogenetic eigenvectors. We assessed

model accuracy using leave-one-out cross validation method to determine percent observations correctly predicted by our ridge model [48].

We analyzed host infection status using a logistic ridge regression implemented in a machine learning framework. We predicted the binary variable of positive antibody or PCR tests as a function of species pace-of-life traits, sampling effort across studies, and the first 48 phylogenetic eigenvectors. Parameter tuning for our ridge model was done using R packages ‘caret’ and ‘glmnet’ with repeated cross validation method, $k = 5$ folds, $n = 5$ repeats, down sampling to balance the design and with area under curve (AUC) as performance measure. During five-fold cross validation species assigned to one of five random “folds” are excluded during model fitting. The proportion of holdout species accurately predicted by the model, which is fitted to the remainder of the data, is used as a measure of expected model accuracy with new species for which data are not currently available. To ensure that this is not biased by the holdout species chosen, the procedure is repeated using species in each of the five random folds as holdouts. After model fitting, we then supplied the estimated lambda parameter to logisticRidge in R package ‘ridge’ to compute coefficient estimates, t-statistics and accompanying p -values for all predictors [46]. We estimated model accuracy using AUC. We further validated results with sensitivity tests for different subsets of data (e.g., all species for which we had data vs only species sampled using PCR) and by calculating relative contribution of individual predictors (details in supporting information, S1–S3 Tables). We predicted both mortality and infection status for all African mammals with final models.

We mined GenBank for NPC1 protein sequences for 300 species. We aligned sequences with NCBI’s constraint based multiple alignment tool (COBALT) [49]. We identified loops-1 and 2 of NPC1 in AliView v1.28 [50] and converted each residue into dummy variables in R with ‘fastDummies’. We predicted infection status as a function of NPC1 sequences, and then separately NPC1 sequences, distance to spillover and sampling effort using logisticRidge in R package ‘ridge’ [46].

Results

Ebolavirus infection and mortality showed significant phylogenetic structure. Mortality after exposure to *Ebolavirus* ($n = 11$ species of 21 tested, Fig 1A) showed strong phylogenetic structure as measured by Fritz and Purvis’ $D = -0.822$. The value of D was significantly different from a random phylogenetic association ($p = 0.001$) but not significantly different from a Brownian model of evolution ($p = 0.921$) (Fig 1A). Positive infection ($n = 56$ positive species of 363 species tested) showed a weak phylogenetic signal ($D = 0.482$), stronger when compared to a random phylogenetic association ($p < 0.001$) but less structured than a Brownian model of evolution ($p = 0.002$, Fig 1B).

We predicted mortality resulting from exposure to ebolaviruses using pace-of-life traits and phylogeny. Our ridge regression model fit to $n = 21$ species had high accuracy (~90%) using leave-one-out cross validation (Table 1). Non-volant species with long gestation lengths and fewer litters per year were more likely to die from exposure to *Ebolavirus* (Table 2). Four phylogenetic eigenvectors (c3, c6, c12 and c13) predicted the ability to tolerate *Ebolavirus* infection (Table 2; also see S2 Table). This combination of eigenvectors was also associated with a high probability of death for primates and terrestrial Artiodactyla relative to other clades (see S1 Fig).

Infection status was related to pace of life traits, fruit consumption and phylogeny. The final model fit had high accuracy with an estimated AUC of 0.802 (Table 1). Species with large adult body mass, large brain mass, high longevity, older age at first reproduction, long gestation length, small litter sizes and fewer litters per year were more likely to test positive for

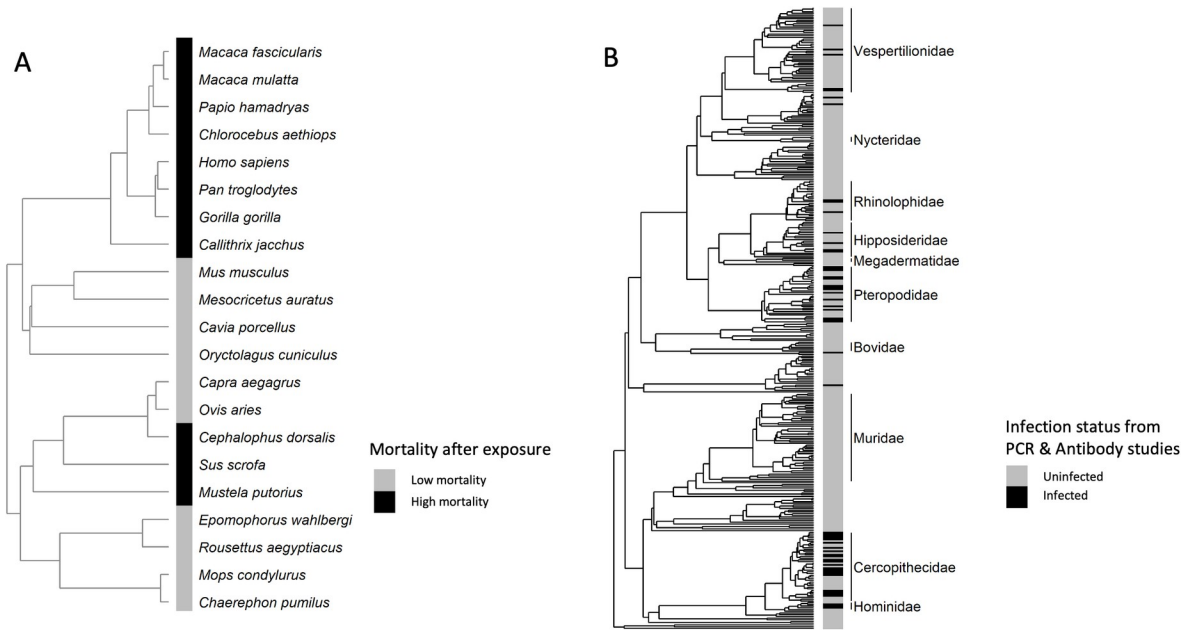


Fig 1. Mortality of species after exposure to *Ebolavirus* plotted on maximum clade credibility tree (A) and infection status of species determined from antibody and PCR tests plotted on maximum clade credibility tree (B).

<https://doi.org/10.1371/journal.pntd.0010993.g001>

Ebolavirus infection than other species across all subsets of data (Table 2). Furthermore, species with high percentage fruit in their diet and species sampled extensively also tested positive for ebolaviruses (Table 2). Multiple phylogenetic eigenvectors were significant predictors of infection status (Table 2) with strong support for three eigenvectors (c3, c11 and c12, S3 Table), corresponding to species in the families Cercopithecidae and Hominidae. Furthermore, fruit bats in the family Pteropodidae showed high infection probabilities (see S1 Fig).

We found NPC1 sequences predicted infection status in bats with high accuracy and that key amino acid positions thought to confer resistance to filoviruses were significant predictors of infection status. Mutations in loop-1 at positions 425–427 previously found to confer resistance to *Marburgvirus* [26] were related to infection status in 31 species for which both sequence and infection status data were available (Tables 1, S4, and S5). Specifically, absence of residue T, E and T in positions 425, 426 and 427 was related to higher probability of infection (S4 Table). Furthermore, residue A at position 425 and G at 426 was positively related to infection; these residues are believed to control susceptibility to *Marburgvirus* in laboratory analyses

Table 1. Summary of all models predicting mortality and infection status considered. For each model, table provides prediction accuracy, number of species to which model was fit, method for evaluating model accuracy and lambda parameter used in ridge regression. The two models used to predict the infection characteristics of African mammals in Fig 1 are italicized.

Model	Prediction accuracy	Number of species	Method for evaluating model accuracy	lambda
<i>Mortality</i>	0.905	21	<i>Leave-one-out CV</i>	0.059
<i>Infection status</i>	0.802	363	AUC	1.9
Infection status (≥ 10 individuals sampled)	0.765	152	AUC	2.7
PCR only	0.655	279	AUC	2.9
Infection status (only free-ranging sampled)	0.757	359	AUC	0.3
NPC1	0.645	31	Leave-one-out CV	0.094
NPC1 + distance to spillover	0.58	31	Leave-one-out CV	0.118

<https://doi.org/10.1371/journal.pntd.0010993.t001>

Table 2. Summary of ridge regression models. For each response variable, namely mortality, infection status, infection status for 10 or more individuals, infection status determined from PCR tests only and infection status predicted for free-ranging mammals, table summarizes the t-statistic of coefficient estimates of predictors and accompanying p-value. Coefficients for 31 eigenvectors (c1-c31) provided in table even though model was run with the first 48 phylogenetic eigenvectors; higher eigenvector coefficients for variables c32-c48 were always non-significant and therefore excluded from table. Bold faced numbers represent significance at α of 0.05. Italicized numbers represent significance at α of 0.1. NA represents coefficients that were not fit in model.

Variable	Mortality		Infection status		Infection status (10 individuals)		PCR tested		Free-ranging mammals	
	t-value	Pr (> t)	t-value	Pr (> t)	t-value	Pr (> t)	t-value	Pr (> t)	t-value	Pr (> t)
adult mass g	-0.379	0.705	2.968	0.003	2.131	0.033	5.150	<0.001	1.550	0.121
brain mass g	0.436	0.663	3.321	0.001	0.687	0.492	5.979	<0.001	1.904	0.057
max longevity d	1.170	0.242	5.879	<0.001	3.213	0.001	3.731	<0.001	3.137	0.002
age first reproduction d	1.185	0.236	5.873	<0.001	3.241	0.001	4.813	<0.001	3.663	<0.001
gestation length d	2.176	0.030	6.345	<0.001	4.789	<0.001	2.820	0.005	3.798	<0.001
litter size n	-0.307	0.759	-4.104	<0.001	-3.613	<0.001	-1.312	0.189	-3.710	<0.001
litters per year n	-2.072	0.038	-4.061	<0.001	-3.298	0.001	-0.270	0.787	-1.797	0.072
percent scavenge	NA	NA	-1.463	0.143	-1.096	0.273	-0.276	0.783	-1.208	0.227
percent seed	-0.414	0.679	-0.396	0.692	-1.207	0.227	-1.306	0.191	-0.316	0.752
percent fruit	0.790	0.430	4.415	<0.001	3.275	0.001	<i>1.915</i>	<i>0.056</i>	2.636	0.008
percent plant	-1.372	0.170	-0.018	0.986	0.020	0.984	0.608	0.543	-0.636	0.525
distance to spillover m	-0.026	0.979	-1.937	0.053	-1.028	0.304	-1.611	0.107	-0.922	0.356
terrestrial volant	-2.119	0.034	0.405	0.686	0.821	0.411	-0.001	0.999	2.733	0.006
c1	-0.218	0.827	-2.014	0.044	-2.588	0.010	-0.944	0.345	-2.823	0.005
c2	0.771	0.441	-1.121	0.262	-1.859	0.063	-0.504	0.614	-2.687	0.007
c3	2.724	0.006	2.632	0.008	2.064	0.039	1.243	0.214	-0.032	0.975
c4	-1.756	0.079	-5.622	<0.001	-4.467	<0.001	-1.054	0.292	-2.367	0.018
c5	-0.008	0.994	-0.628	0.530	-0.062	0.951	-0.965	0.334	-1.038	0.299
c6	-2.474	0.013	-5.711	<0.001	-4.139	<0.001	-1.515	0.130	-2.354	0.019
c7	-0.422	0.673	-3.627	<0.001	-3.653	<0.001	-1.316	0.188	-2.640	0.008
c8	0.931	0.352	0.222	0.824	0.153	0.878	0.203	0.839	0.575	0.565
c9	1.303	0.193	-0.890	0.373	-0.175	0.861	-0.324	0.746	-1.069	0.285
c10	1.426	0.154	-3.126	0.002	-1.892	0.059	-0.721	0.471	-3.433	0.001
c11	1.259	0.208	-3.309	0.001	-2.677	0.007	-0.925	0.355	-3.497	<0.001
c12	-2.207	0.027	-4.325	<0.001	-3.610	<0.001	-0.995	0.320	-1.501	0.133
c13	2.162	0.031	4.552	<0.001	3.849	<0.001	0.965	0.335	1.675	0.094
c14	-0.155	0.877	-2.006	0.045	-1.398	0.162	-0.705	0.481	-1.020	0.308
c15	1.443	0.149	-2.455	0.014	-2.554	0.011	-2.278	0.023	-2.042	0.041
c16	-0.181	0.857	-1.149	0.250	-1.623	0.104	-0.512	0.608	-0.919	0.358
c17	-0.848	0.397	1.756	0.079	1.890	0.059	0.270	0.787	1.084	0.278
c18	-0.429	0.668	-1.088	0.277	-1.518	0.129	0.395	0.693	-0.561	0.575
c19	-0.936	0.349	-1.436	0.151	-1.100	0.271	-0.998	0.318	-0.592	0.554
c20	-0.419	0.675	-1.570	0.116	-1.779	<i>0.075</i>	1.615	0.106	-0.874	0.382
c21	-0.862	0.388	0.218	0.828	0.391	0.696	-0.093	0.926	0.199	0.842
c22	-0.828	0.408	0.025	0.980	-0.576	0.564	3.071	0.002	0.293	0.770
c23	-0.617	0.537	1.295	0.195	0.648	0.517	2.817	0.005	1.131	0.258
c24	-0.823	0.410	0.394	0.693	0.258	0.796	0.317	0.751	0.340	0.734
c25	1.196	0.232	0.418	0.676	0.223	0.824	-0.016	0.987	0.298	0.765
c26	0.616	0.538	0.210	0.833	0.364	0.716	-0.893	0.372	-0.018	0.986
c27	0.674	0.500	-1.289	0.197	-1.117	0.264	-0.649	0.516	-0.894	0.371
c28	-0.452	0.651	-6.373	<0.001	-4.417	<0.001	-0.653	0.514	-2.412	0.016
c29	-0.370	0.711	-1.377	0.169	-1.358	0.174	-0.383	0.702	-1.016	0.310

(Continued)

Table 2. (Continued)

Variable	Mortality		Infection status		Infection status (10 individuals)		PCR tested		Free-ranging mammals	
	t-value	Pr (> t)	t-value	Pr (> t)	t-value	Pr (> t)	t-value	Pr (> t)	t-value	Pr (> t)
c30	-1.000	0.317	0.299	0.765	-0.257	0.797	0.206	0.836	0.272	0.786
c31	-0.737	0.461	-1.877	0.061	-2.011	0.044	-0.624	0.533	-1.464	0.143
sample size	NA	NA	6.110	<0.001	3.445	0.001	2.986	0.003	6.045	<0.001

<https://doi.org/10.1371/journal.pntd.0010993.t002>

of bat cell lines [26] (S4 Table). Even models of NPC1 with nuisance variables of sampling effort and distance to spillover site still found significant positive infection in species with residue A at position 425 and marginal significance for other residues in positions 425–427 (S4 Table). Although our NPC1 models showed poor accuracy using leave-one-out cross validation (Table 1), these models predicted infection status of bats with 100% accuracy; whereas trait and phylogenetic models showed only 71% accuracy (Table 3) likely due to the low sensitivity of trait and phylogenetic eigenvector models for this group, meaning the ability to identify a true positive infection status in bats is low (see S6 Table). For primates, the percent accuracy of all models was low at 50–58% (Table 3); however, the sensitivity of trait and phylogenetic eigenvectors models was 1 (S6 Table). Therefore, the low accuracy was due to low specificity (i.e., poor ability to identify a true negative infection status) for primates (S6 Table).

Reservoir status predictions from our ridge regression models showed strong correlations to phylogenetic clades (Fig 2, see S2 Fig for visual comparison of predictions with raw data). The order Perissodactyla, families Ceropithecidae, Hominidae and Suidae showed high likelihood of death after exposure to *Ebolavirus* as well as high probability of past infection as estimated by antibody and PCR tests (Fig 2). We interpreted this category as ‘dead-end hosts’ unlikely to survive after exposure to *Ebolavirus* and therefore unlikely to sustain the pathogen for long periods of time in the wild; a criteria typically considered to be important for a species to serve as a ‘natural reservoir’ for a pathogen [11]. In contrast, fruit-eating bats in family Pteropodidae showed high probability of past infection in antibody and PCR tests and are predicted to have low mortality following exposure (Fig 2). Some members of Bovidae and Afrosoricida also fall into this category. We interpreted this group as potential ‘reservoirs’, rarely succumbing to infection and periodically serving as a source of infection for other hosts with higher mortality. Some species showed both low likelihood of being infected and low mortality even if exposed, which we interpreted as ‘low exposure and susceptibility’ (Fig 2); others showed low probability of infection but high mortality when exposed, which we interpreted as species ‘susceptible but rarely exposed’ to ebolaviruses (Fig 2). These predictions, however, are based on our trait and phylogenetic eigenvector models which has uneven prediction accuracy across clades (S6 Table).

Table 3. Proportion of correct predictions for species in orders Chiroptera and Primates made by three competing models, namely trait and phylogenetic eigenvector model, NPC1 sequence model, and NPC1 model with distance to spillover and sampling effort variables included.

	Proportion correct predictions	
	Chiroptera (n = 14)	Primates (n = 12)
Trait + Phylogenetic eigenvectors	0.714	0.5
NPC1	1	0.583
NPC1 + distance to spillover	1	0.583

<https://doi.org/10.1371/journal.pntd.0010993.t003>

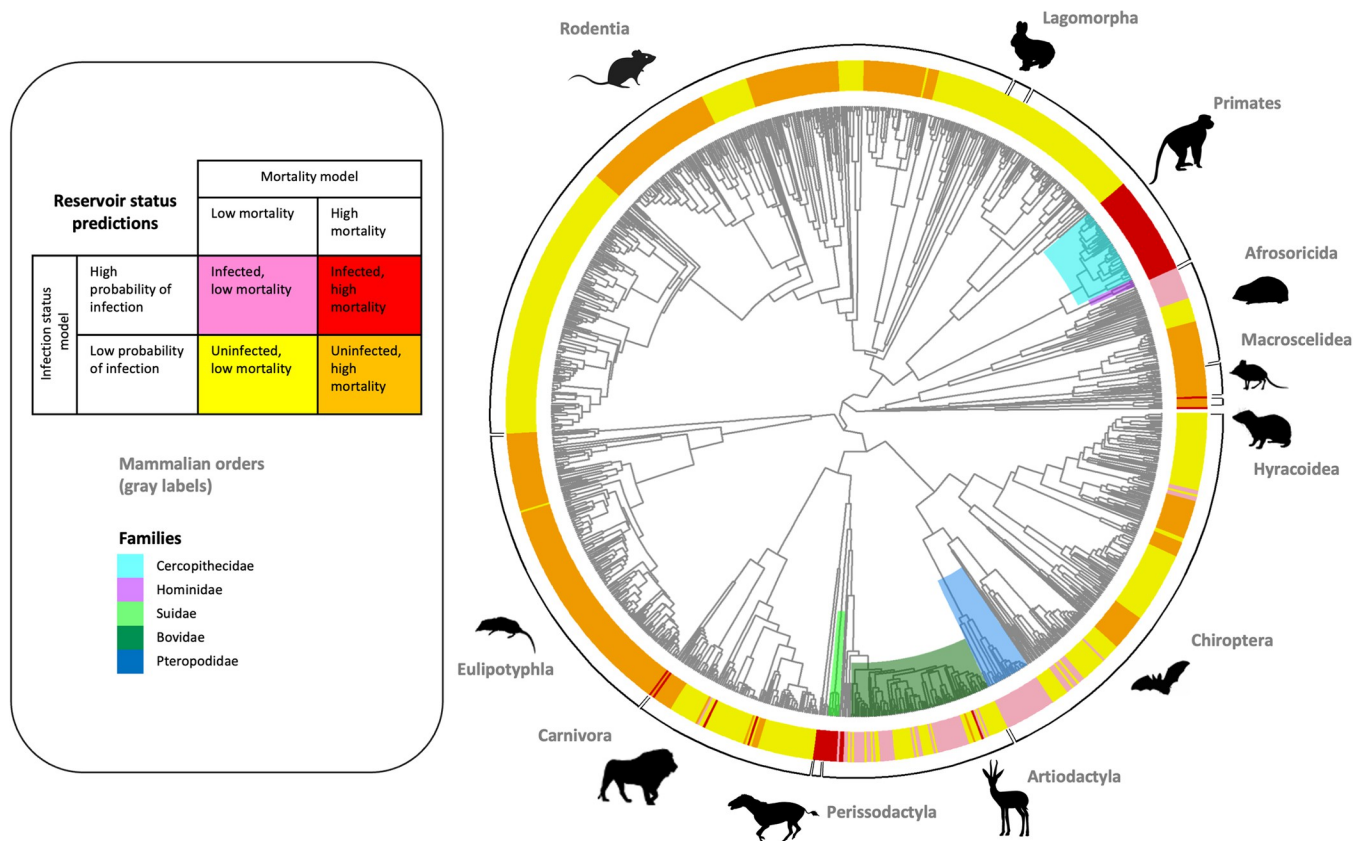


Fig 2. Predictions of reservoir status for all terrestrial African mammals based on ridge model predicting mortality of species after exposure to *Ebolavirus* and ridge model predicting infection status of species. Ridge models used trait and phylogenetic eigenvectors as predictors (see main text for more details of models). Silhouettes used are available under Public Domain (<https://creativecommons.org/publicdomain/zero/1.0/>).

<https://doi.org/10.1371/journal.pntd.0010993.g002>

Discussion

Here, we explore *Ebolavirus* host range across African mammals by combining previously disparate information about host traits and phylogeny, host receptor sequences, and their ability to predict varying degrees of susceptibility to infection. This approach enabled us to determine which trait and phylogenetic (hereafter TP) variables were the best predictors of both host infection probability and mortality when infected, and to estimate the reservoir status of all terrestrial African mammals.

Ebolavirus mortality across species was best predicted by phylogeny and pace-of-life traits. Species vary in pace-of-life along a continuum from slow to fast; some species reproduce rapidly and have shorter life spans, while others reproduce slowly and have longer life spans [51,52]. Fast vs slow species also tend to invest differently in immune functions [53–55]. Our analyses suggest that slow species, particularly primates, with long gestation lengths and few litters per year, were more likely to succumb to infection from *Ebolavirus* (Fig 1A and Table 2). Volant species and fast paced species, such as mice (*Mus musculus*), were more likely to survive (Fig 1A and Table 2), supporting one theory that these species regulate inflammatory immune defenses to fight viral infections [53]. Several studies also suggest that bats have specific immune strategies to fight infections which could allow these species to serve as reservoirs for viruses [56]; including one study providing evidence of bat responses to *Ebolavirus* [57].

One limitation of our mortality analysis was the relatively low number of species that we could include due in large part to our reliance on laboratory studies. Among 182 observations of mortality in the wild or the laboratory that we located (see raw data provided on figshare <https://doi.org/10.6084/m9.figshare.20250408.v1>), a total of only 21 species were represented. Despite including relatively few observations, based on delete-one cross validation our model was able to predict the mortality of species excluded during model fitting with better than 90% accuracy (Table 1). We speculate that this high accuracy is due to a number of factors. First, we were only attempting to predict mortality at an extremely coarse level, “high” or “low.” Second, mortality once exposed to *Ebolavirus* likely depends largely on inherent characteristics of species and phylogenetic relationships, making it easier to infer mortality compared to infection status which depends on both the susceptibility of species to infection and the frequency with which they happen to encounter the virus in the wild. Finally, the ridge regression method we used [44, 45] was designed for the precise use case of an analysis where the number of predictors exceeds, or even greatly exceeds, the number of observations. Regardless, our study points to a great need for direct observations of variation in mortality after exposure to *Ebolavirus* for additional species.

Infection status was related to a suite of phylogenetic, evolutionary and ecological traits as well as to sampling effort. In contrast to mortality, frugivory and sampling effort, rather than pace-of-life traits and phylogeny alone (Table 2) were important predictors of infection status (S1 Table). This suggests that the geographic overlap of frugivorous species may support *Ebolavirus* spread among hosts in Sub-Saharan Africa. The importance of sampling effort in predicting infection from *Ebolavirus* (also noted in [19]), underscores the need for more systematic sampling across taxa, which may benefit from focused sampling during times of year when synchronous fruiting supports the spatial overlap of multiple frugivorous species that are potential hosts for ebolaviruses.

Though the data on infection status that we present represent the most comprehensive collection of *Ebolavirus* host records we are aware of, these data still have important limitations. The infection status data are likely biased as a result of better surveillance and increased sampling efforts for charismatic species or species of conservation concern. This issue could be mitigated by more systematic and focused sampling efforts in the future. Our data can also be used to identify clades that are undersampled (see S6 Table for sample sizes) and used to guide future efforts. To account for differences in sampling effort, we incorporated total number of individuals sampled as a covariate in our model and we have used repeated cross validation in a machine learning framework to test the robustness of our results. We also note that infection status determined from PCR tests have slightly different interpretations from antibody tests. Specifically, positive PCR samples suggests active circulation or infection from *Ebolavirus*; whereas positive antibody tests cannot distinguish between current or chronic infection and past exposure and infection which has been cleared [58]. Our goal was to model which types of species are likely to be susceptible to infection and therefore included both tests in most of our models. We also modelled PCR tested individuals separately and found qualitatively similar results to models including evidence of infection from any source (Table 2). The primary difference was weaker statistical support as a result of much lower numbers of species with positive PCR tests (Table 1, see S1 Text).

Infection status in bats is likely related to their immune system characteristics, whereas ecological traits of primates play a role in exposure and infection. In particular, our estimates of true positive rates, while high for the frugivorous Pteropodidae, were low for species in the insectivorous bat families Hipposideridae, Vespertilionidae, Molossidae and Minopteridae (see S6 Table). That our NPC1 models predicted the infection status of bats from multiple families with high accuracy (Tables 3 and S5) suggests that available bat trait and phylogenetic data do not adequately capture cross-species differences in immune functioning that can

explain infection status. Better data on traits related to immune functioning in bats is likely to improve predictive models of infection status. Unlike bats, primates showed high true positive and low true negative rates (S6 Table). Our models predicted all Cercopithecids and Homioids are susceptible to *Ebolavirus* infection (Fig 2) even though infections have so far been documented in only a handful of species. Cercopithecoid and Homioid species are also known to succumb to infection once exposed to ebolaviruses [9,15,59]. Temporal and spatial overlap with bats in the use of resources such as fruit trees may explain differential risk of exposure to ebolaviruses in the wild. Detailed behavioral data is hard to find and difficult to gather, but, nonetheless, a critical need given that seroprevalence data captures those species that can be immunologically infected with *Ebolavirus* only as a subset of species that have the ecological opportunity to be infected.

Although the estimated accuracy of our NPC1 model was low across mammals overall (Table 1), the final model successfully predicted infection status of bat species included in the model with high accuracy (Tables 3 and S5). Thus, exploring how NPC1 sequences relates to infection status in bats may improve prediction, especially given the low sensitivity of the alternative TP model. We speculate that the high accuracy of the bat predictions in the NPC1 model may be related to the fact that the loop regions we included in our models were identified by laboratory work performed on bat cell lines [26]. Further, it could also be because immunological characteristics influence infection statuses of bats to a strong degree. Our NPC1 model was less accurate for primates, but similar to the TP results, this was due to low true negative rates for primates (S5 Table) which suggests that behavioral and ecological characteristics of some primates could lower exposure to virus. Infected primates could also frequently die before they are sampled leading to negative infection statuses.

In our NPC1 model, residues that confer high affinity to binding with wild type filovirus glycoproteins strongly determined infection status. We found that residues relating to increased affinity to *Marburgvirus* [26] predicted positive *Ebolavirus* infection in bats (S4 and S5 Tables). However, the specific positions and residues identified by Takadate et al. [26] as binding to *Ebolavirus* do not match our findings. Although straw-colored fruit bats (*Eidolon helvum*) have been found to carry antibodies to *Ebolavirus* in serological studies [60–62], Takadate et al. [26] considered this species to be resistant to *Ebolavirus*. Conversely, an important reservoir for *Marburgvirus*, *Rousettus aegyptiacus* [63,64], does not carry the residues thought to confer resistance, carrying instead residues that increase affinity to the virus [26] (S5 Table). Furthermore, when inoculated with *Marburgvirus*, *R. aegyptiacus* tolerates infection and sheds virus [65,66] suggesting the operation of other mechanisms beyond its NPC1 receptors that allow it to fight infection [57]. This discrepancy of *R. aegyptiacus* not carrying the sequences needed to confer resistance to *Ebolavirus* and *Marburgvirus* was noted by Takadate et al. [26] in their study also leading to the conclusion that unique host factors such as interferons likely influence susceptibility to infection. While the positions and residues identified by Takadate et al. [26] are important predictors of binding affinity to filoviruses and possibly infection status, they do not necessarily confer resistance to filoviruses. Kurosaki et al. [67] also showed that small mutations, affecting only two amino acid residues in *Zaire ebolavirus* compared to wild type strains, can greatly increase infectivity across cell lines expressing NPC1 sequences found in both bats and primates, including humans. More work is clearly needed to identify how binding affinity to NPC1 and host immune responses relates to viral replication within an infected individual (e.g., [26,57]). Sequencing NPC1 gene regions for additional species could also be useful. For example, laboratory *Ebolavirus* inoculation studies for bat species including *Epomops wahlbergi* and two *Tadarida* species have been published [25], however, their NPC1 gene regions have not yet been sequenced. NPC1 sequences are currently available in GenBank for only 31 of the 363 species for which antibody and PCR test results have been published.

We modeled reservoir status across terrestrial African mammals. If we define a reservoir as a species likely to be infected but not to die, our models predict Pteropodid fruit bats, also identified as strong reservoir candidates in other studies [68], as likely reservoirs (Fig 2). In experimental inoculations of fruit and insectivorous bats with ebolaviruses, no evidence exists of either death or illness in bats carrying the virus [25], strongly supporting the idea that bats can naturally tolerate the virus while also serving as a source of infection [11,17]. Our analysis also highlights species from the order Afrosoricida and family Bovidae as potential reservoirs. To our knowledge, none of these species have been sampled for *Ebolavirus* in the wild, though some of them do overlap with known spillover locations. Interestingly, shrew species from the clade Soricidae show evidence of inserted filoviral elements in their genome, which suggests evolutionary history with filoviruses and potential infection of an ancestor [69] and positive infection status has been noted in species such as *Sylvisorex ollula* [70]. Our model is unable to predict positive infection status for this species (S6 Table). Therefore, closer examination of the clade Soricidae could also help clarify host status predictions estimated in this paper. Our analyses identified Cercopithecidae, Hominiidae, Suidae and the order Perissodactyla as “dead-end”, or secondary amplifying hosts that succumb to infection rapidly. While much research has focused on identifying the elusive reservoir of *Ebolavirus* [18], known dead-end hosts including *Pan troglodytes* and *Gorilla gorilla* appear to be the source of several human outbreaks [4,15,71,72], and have also suffered drastic population size reductions after epizootic outbreaks [13,73]. Provisionally, because more testing is still needed to confirm *Ebolavirus* mortality, our study adds species from the order Perissodactyla to the list (Fig 2) of potential dead-end amplifying wild hosts. Better understanding of potential reservoir statuses across mammals (list of reservoir statuses from our models provided on figshare: <https://doi.org/10.6084/m9.figshare.20250408.v1>) will help further refine knowledge of the risk factors for African *Ebolavirus* spillover into human populations.

Supporting information

S1 Text. Supplementary Methods and Results.

(PDF)

S1 Table. Relative importances of predictor variables in ridge models.

(PDF)

S2 Table. Sensitivity of mortality of animal host model to uncertainty in phylogenetic relationships.

(PDF)

S3 Table. Sensitivity of infection status model to uncertainty in phylogenetic relationships.

(PDF)

S4 Table. Ridge regression predicting infection status of animal host on the basis of Niemann-Pick C1 amino acid residues, distance to spillover site and sampling effort.

(PDF)

S5 Table. Niemann-Pick C1 residues at key positions identified by Takadate et al. [26] for species for whom infection status is also known.

(PDF)

S6 Table. True positive rates (or sensitivity) and true negative rates (or specificity) provided for trait and phylogenetic eigenvector ridge model predicting infection status by mammalian clade.

(PDF)

S1 Fig. Plots of significant phylogenetic eigenscores for all mammalian clades. Phylogenetic eigenvector 3 or c3 plotted on maximum clade credibility tree (A), phylogenetic eigenvector 11 or c11 plotted on maximum clade credibility tree (B), and phylogenetic eigenvector 12 or c12 plotted on maximum clade credibility tree (C). These eigen scores consistently predict host mortality and host infection status; with c3 positively being related to host status, c11 and c12 being negatively related to host status.

(PDF)

S2 Fig. Predictions of reservoir status for all known terrestrial African mammals and accompanying accuracy metrics. (A) Predictions of reservoir status which are based on ridge model predicting mortality of species after exposure to *Ebolavirus* and ridge model predicting infection status of species. Ridge models used trait and phylogenetic eigenvectors as predictors (see main text for more details of models). (B) Accuracy of infection status predictions by mammal clade. Silhouettes used are available under Public Domain (<https://creativecommons.org/publicdomain/zero/1.0/>).

(PDF)

Author Contributions

Conceptualization: Mekala Sundaram, John Paul Schmidt, Barbara A. Han, John M. Drake, Patrick R. Stephens.

Data curation: Mekala Sundaram, Patrick R. Stephens.

Formal analysis: Mekala Sundaram.

Funding acquisition: John Paul Schmidt, John M. Drake, Patrick R. Stephens.

Investigation: Mekala Sundaram, John Paul Schmidt, Barbara A. Han, John M. Drake, Patrick R. Stephens.

Methodology: Mekala Sundaram, Barbara A. Han, John M. Drake, Patrick R. Stephens.

Project administration: Patrick R. Stephens.

Resources: Mekala Sundaram, John M. Drake, Patrick R. Stephens.

Software: Mekala Sundaram.

Supervision: Patrick R. Stephens.

Validation: Mekala Sundaram.

Visualization: Mekala Sundaram.

Writing – original draft: Mekala Sundaram, Patrick R. Stephens.

Writing – review & editing: Mekala Sundaram, John Paul Schmidt, Barbara A. Han, John M. Drake, Patrick R. Stephens.

References

1. Kuhn JH. Filoviruses: a compendium of 40 years of epidemiological, clinical, and laboratory studies. Calisher CH, editor. New York: SpringerWien; 2008.
2. Leroy EM, Epelboin A, Mondonge V, Pourrut X, Gonzalez JP, Muyembe-Tamfum JJ, et al. Human ebola outbreak resulting from direct exposure to fruit bats in Luebo, Democratic Republic of Congo, 2007. Vector-Borne Zoonotic Dis. 2009; 9: 723–728. <https://doi.org/10.1089/vbz.2008.0167> PMID: 19323614

3. Pourrut X, Kumulungui B, Wittmann T, Moussavou G, Délicat A, Yaba P, et al. The natural history of Ebola virus in Africa. *Microbes Infect.* 2005; 7: 1005–1014. <https://doi.org/10.1016/j.micinf.2005.04.006> PMID: 16002313
4. Lahm SA, Kombila M, Swanepoel R, Barnes RFW. Morbidity and mortality of wild animals in relation to outbreaks of Ebola haemorrhagic fever in Gabon, 1994–2003. *Trans R Soc Trop Med Hyg.* 2007; 101: 64–78. <https://doi.org/10.1016/j.trstmh.2006.07.002> PMID: 17010400
5. World Health Organization. Outbreak of Ebola haemorrhagic fever in Yambio, south Sudan, April–June 2004. *Wkly Epidemiol Rec.* 2005; 80: 370–375. PMID: 16285261
6. Saéz AM, Weiss S, Nowak K, Lapeyre V, Zimmermann F, Dux A, et al. Investigating the zoonotic origin of the West African Ebola epidemic. *EMBO Mol Med.* 2015; 7: 17–23. Available: https://www.academia.edu/24080902/Investigating_the_zoonotic_origin_of_the_West_African_Ebola_epidemic. <https://doi.org/10.15252/emmm.201404792> PMID: 25550396
7. Reiter P, Turell M, Coleman R, Miller B, Maupin G, Liz J, et al. Field investigations of an outbreak of Ebola hemorrhagic fever, Kikwit, Democratic Republic of the Congo, 1995: Arthropod studies. *J Infect Dis.* 1999; 179: 148–154. <https://doi.org/10.1086/514304> PMID: 9988178
8. Breman JG, Johnson KM, Van Der Groen G, Robbins CB, Szczeniowski M V., Ruti K, et al. A search for Ebola virus in animals in the Democratic Republic of the Congo and Cameroon: Ecologic, virologic, and serologic surveys, 1979–1980. *J Infect Dis.* 1999; 179: 1979–1980. <https://doi.org/10.1086/514278> PMID: 9988177
9. Groseth A, Feldmann H, Strong JE. The ecology of Ebola virus. *Trends Microbiol.* 2007; 15: 408–416. <https://doi.org/10.1016/j.tim.2007.08.001> PMID: 17698361
10. Leirs H, Mills JN, Krebs JW, Childs JE, Akaibe D, Woollen N, et al. Search for the Ebola virus reservoir in Kikwit, Democratic Republic of the Congo: Reflections on a vertebrate collection. *J Infect Dis.* 1999; 179: 155–163. <https://doi.org/10.1086/514299> PMID: 9988179
11. Amman BR, Swanepoel R, Nichol ST, Towner JS. Ecology of Filoviruses. In: Mühlberger E, Towner JS, Henley LL, editors. *Marburg- and Ebolaviruses From Ecosystems to Molecules*. Cham, Switzerland: Springer International; 2017. pp. 23–61.
12. Georges AJ, Leroy EM, Renaut AA, Benissan CT, Nabias RJ, Ngoc MT, et al. Ebola hemorrhagic fever outbreaks in Gabon, 1994–1997: Epidemiologic and health control issues. *J Infect Dis.* 1999; 179: 1994–1997. <https://doi.org/10.1086/514290> PMID: 9988167
13. Bermejo M, Rodríguez-Teijeiro JD, Illera G, Barroso A, Vilà C, Walsh PD. Ebola outbreak killed 5000 gorillas. *Science* (80-). 2006; 314: 1564. <https://doi.org/10.1126/science.1133105> PMID: 17158318
14. Formenty P, Boesch C, Wyers M, Steiner C, Donati F, Dind F, et al. Ebola virus outbreak among wild chimpanzees living in a rain forest of Cote d'Ivoire. *J Infect Dis.* 1999; 179: 120–126. <https://doi.org/10.1086/514296> PMID: 9988175
15. Hayman DTS. African Primates: Likely Victims, Not Reservoirs, of Ebolaviruses. *J Infect Dis.* 2019; 220: 1547–1550. <https://doi.org/10.1093/infdis/jiz007> PMID: 30657949
16. Commission M of the I. Ebola haemorrhagic fever in Zaire, 1976. Report of an international commission. *Bull World Health Organ.* 1978; 56: 271–293.
17. Caron A, Bourgarel M, Cappelle J, Liégeois F, De Nys HM, Roger F. Ebola virus maintenance: if not (Only) bats, what else? *Viruses.* 2018; 10. <https://doi.org/10.3390/v10100549> PMID: 30304789
18. Peterson TT, Carroll DS, Mills JN, Johnson KM. Potential mammalian filovirus reservoirs. *Emerg Infect Dis.* 2004; 10: 2073–2081. <https://doi.org/10.3201/eid1012.040346> PMID: 15663841
19. Schmidt JP, Maher S, Drake JM, Huang T, Farrell MJ, Han BA. Ecological indicators of mammal exposure to *Ebolavirus*. *Philos Trans R Soc B Biol Sci.* 2019; 374. <https://doi.org/10.1098/rstb.2018.0337> PMID: 31401967
20. Han BA, Schmidt JP, Alexander LW, Bowden SE, Hayman DTS, Drake JM. Undiscovered Bat Hosts of Filoviruses. *PLoS Negl Trop Dis.* 2016; 10: 1–10. <https://doi.org/10.1371/journal.pntd.0004815> PMID: 27414412
21. Cooper N, Griffin R, Franz M, Omotayo M, Nunn CL. Phylogenetic host specificity and understanding parasite sharing in primates. *Ecol Lett.* 2012; 15: 1370–1377. <https://doi.org/10.1111/j.1461-0248.2012.01858.x> PMID: 22913776
22. Huang S, Bininda-Emonds ORP, Stephens PR, Gittleman JL, Altizer S. Phylogenetically related and ecologically similar carnivores harbour similar parasite assemblages. *J Anim Ecol.* 2014; 83: 671–680. <https://doi.org/10.1111/1365-2656.12160> PMID: 24289314
23. Stephens PR, Altizer S, Ezenwa VO, Gittleman JL, Moan E, Han B, et al. Parasite sharing in wild ungulates and their predators: Effects of phylogeny, range overlap, and trophic links. *J Anim Ecol.* 2019; 88: 1017–1028. <https://doi.org/10.1111/1365-2656.12987> PMID: 30921468

24. Olivero J, Fa JE, Real R, Farfán MÁ, Márquez AL, Vargas JM, et al. Mammalian biogeography and the Ebola virus in Africa. *Mamm Rev*. 2017; 47: 24–37. <https://doi.org/10.1111/mam.12074>
25. Swanepoel R, Leman PA, Burt FJ, Zachariades NA, Braack LEO, Ksiazek TG, et al. Experimental Inoculation of Plants and Animals with Ebola Virus. *Emerg Infect Dis*. 1996; 2: 321–325. <https://doi.org/10.3201/eid0204.960407> PMID: 8969248
26. Takadate Y, Kondoh T, Igarashi M, Maruyama J, Manzoor R, Ogawa H, et al. Niemann-Pick C1 Heterogeneity of Bat Cells Controls Filovirus Tropism. *Cell Rep*. 2020; 30: 308–319.e5. <https://doi.org/10.1016/j.celrep.2019.12.042> PMID: 31940478
27. Ng M, Ndungo E, Kaczmarek ME, Herbert AS, Binger T, Kuehne AI, et al. Filovirus receptor NPC1 contributes to species-specific patterns of *ebolavirus* susceptibility in bats. *Elife*. 2015; 4: 1–22. <https://doi.org/10.7554/eLife.11785> PMID: 26698106
28. Ndungo E, Herbert AS, Raaben M, Obernosterer G, Biswas R, Miller EH, et al. A single residue in Ebola virus receptor NPC1 influences cellular host range in reptiles. *mSphere*. 2016; 1: 1–15. <https://doi.org/10.1128/mSphere.00007-16> PMID: 27303731
29. Carstea ED, Polymeropoulos MH, Parker CC, Deterawadleigh SD, Oneill RR, Patterson MC, et al. Linkage of Niemann-Pick disease Type-C to human chromosome-18. *Proc Natl Acad Sci U S A*. 1993; 90: 2002–2004. <https://doi.org/10.1073/pnas.90.5.2002> PMID: 8446622
30. Ng M, Ndungo E, Jangra RK, Cai Y, Postnikova E, Sheli R, et al. Cell entry by a novel European filovirus requires host endosomal cysteine proteases and Niemann-Pick C1. *Virology*. 2014; 0: 637–646. <https://doi.org/10.1016/j.virol.2014.08.019> PMID: 25310500
31. Yamamoto T, Nanba E, Ninomiya H, Higaki K, Taniguchi M, Zhang H, et al. NPC1 gene mutations in Japanese patients with Niemann-Pick disease type C. *Hum Genet*. 1999; 105: 10–16. <https://doi.org/10.1007/s004399900059> PMID: 10480349
32. Scott C, Ioannou YA. The NPC1 protein: Structure implies function. *Biochim Biophys Acta—Mol Cell Biol Lipids*. 2004; 1685: 8–13. <https://doi.org/10.1016/j.bbalip.2004.08.006> PMID: 15465421
33. Fluegel ML, Parker TJ, Pallanck LJ. Mutations of a drosophila NPC1 gene confer sterol and ecdysone metabolic defects. *Genetics*. 2006; 172: 185–196. <https://doi.org/10.1534/genetics.105.046565> PMID: 16079224
34. Liu B, Li H, Repa JJ, Turley SD, Dietschy JM. Genetic variations and treatments that affect the lifespan of the NPC1 mouse. *J Lipid Res*. 2008; 49: 663–669. <https://doi.org/10.1194/jlr.M700525-JLR200> PMID: 18077828
35. Polese-Bonatto M, Bock H, Farias ACS, Mergener R, Matte MC, Gil MS, et al. Niemann-Pick Disease Type C: Mutation Spectrum and Novel Sequence Variations in the Human NPC1 Gene. *Mol Neurobiol*. 2019; 56: 6426–6435. <https://doi.org/10.1007/s12035-019-1528-z> PMID: 30820861
36. Benson DA, Cavanaugh M, Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, et al. GenBank. *Nucleic Acids Res*. 2013; 41: 36–42. <https://doi.org/10.1093/nar/gks1195> PMID: 23193287
37. Soria CD, Pacifici M, Di Marco M, Stephen SM, Rondinini C. COMBINE: a coalesced mammal database of intrinsic and extrinsic traits. *Ecology*. 2021; 102: 13028255. <https://doi.org/10.1002/ecy.3344> PMID: 33742448
38. Barton RA, Capellini I. Maternal investment, life histories, and the costs of brain growth in mammals. *Proc Natl Acad Sci U S A*. 2011; 108: 6169–6174. <https://doi.org/10.1073/pnas.1019140108> PMID: 21444808
39. Schmidt JP, Park AW, Kramer AM, Han BA, Alexander LW, Drake JM. Spatiotemporal fluctuations and triggers of ebola virus spillover. *Emerg Infect Dis*. 2017; 23: 415–422. <https://doi.org/10.3201/eid2303.160101> PMID: 28221131
40. Upham NS, Esselstyn JA, Jetz W. Inferring the mammal tree: Species-level sets of phylogenies for questions in ecology, evolution, and conservation. *PLoS Biology*. 2019. <https://doi.org/10.1371/journal.pbio.3000494> PMID: 31800571
41. Santos T, Diniz-Filho JA, Bini TR e LM. Package “PVR.” 2012. pp. 1–13.
42. Fritz SA, Purvis A. Selectivity in mammalian extinction risk and threat types: A new measure of phylogenetic signal strength in binary traits. *Conserv Biol*. 2010; 24: 1042–1051. <https://doi.org/10.1111/j.1523-1739.2010.01455.x> PMID: 20184650
43. Frank LE, Friedman JH. A statistical view of some chemometrics regression tools. *Technometrics*. 1993; 35: 109–135. <https://doi.org/10.1080/00401706.1993.10485033>
44. Cule E, Vineis P, De Iorio M. Significance testing in ridge regression for genetic data. *BMC Bioinformatics*. 2011; 12. <https://doi.org/10.1186/1471-2105-12-372> PMID: 21929786
45. Cule E, De Iorio M. Ridge regression in prediction problems: Automatic choice of the ridge parameter. *Genet Epidemiol*. 2013; 37: 704–714. <https://doi.org/10.1002/gepi.21750> PMID: 23893343

46. Moritz S, Cule E, Frankowski D. ridge: Ridge regression with automatic selection of the penalty parameter. R package; 2022.
47. Tu-Chan AP, Natraj N, Godlove J, Abrams G, Ganguly K. Effects of somatosensory electrical stimulation on motor function and cortical oscillations. *J. Neuroeng. Rehabilitation*, 2017; 14: 1–9. <https://doi.org/10.1186/s12984-017-0323-1> PMID: 29132379
48. Ugarte MD, Militino AF, Arnholt AT. Probability and statistics with R. 2nd ed. Boca Raton: CRC Press; 2016.
49. Papadopoulos JS, Agarwala R. COBALT: Constraint-based alignment tool for multiple protein sequences. *Bioinformatics*. 2007; 23: 1073–1079. <https://doi.org/10.1093/bioinformatics/btm076> PMID: 17332019
50. Larsson A. AliView: A fast and lightweight alignment viewer and editor for large datasets. *Bioinformatics*. 2014; 30: 3276–3278. <https://doi.org/10.1093/bioinformatics/btu531> PMID: 25095880
51. Promislow DE., Harvey PH. Living fast and dying young: a comparative analysis of life-history variation among mammals. *J Zool*. 1990; 220: 417–437. <https://doi.org/10.1111/j.1469-7998.1990.tb04316.x>
52. Ricklefs R, Wikelski M. Biodiversity reflects in part the diversification of life histories. *Trends Ecol Evol*. 2002; 17: 462–468. Available: https://ac.els-cdn.com/S0169534702025788/1-s2.0-S0169534702025788-main.pdf?_tid=9d9d8202-a223-11e7-9008-00000aacb35e&acdnat=1506366131_56ecc99336113d5fe5d529b4572abb84.
53. Lee KA. Linking immune defenses and life history at the levels of the individual and the species. *Integr Comp Biol*. 2006; 46: 1000–1015. <https://doi.org/10.1093/icb/ici049> PMID: 21672803
54. Tieleman BI, Williams JB, Ricklefs RE, Klasing KC. Constitutive innate immunity is a component of the pace-of-life syndrome in tropical birds. *Proc R Soc B-Biological Sci*. 2005; 272: 1715–1720. <https://doi.org/10.1098/rspb.2005.3155> PMID: 16087427
55. Previtali MA, Ostfeld RS, Keesing F, Jolles AE, Hanselmann R, Martin LB. Relationship between pace of life and immune responses in wild rodents. *Oikos*. 2012; 121: 1483–1492. <https://doi.org/10.1111/j.1600-0706.2012.020215.x>
56. Schountz T, Baker ML, Butler J, Munster V. Immunological control of viral infections in bats and the emergence of viruses highly pathogenic to humans. *Front Immunol*. 2017; 8: 1098. <https://doi.org/10.3389/fimmu.2017.01098> PMID: 28959255
57. Jayaprakash AD, Ronk AJ, Prasad AN, Covington MF, Stein KR, Schwarz TM, et al. Marburg and Ebola virus infections elicit a muted inflammatory state in bats. *bioRxiv*. 2021. Available: <https://www.biorxiv.org/content/early/2021/04/29/2020.04.13.039503>.
58. Abbas AK, Lichtman AH, Pillai S. Cellular and molecular immunology E-book. Elsevier Health Sciences; 2022.
59. Bennett RS, Huzella LM, Jahrling PB, Bollinger L, Olinger GG Jr, Hensley LE. Nonhuman primate models of Ebola virus disease. In: Mühlberger E, Towner JS, Henley LL, editors. *Marburg- and Ebolaviruses From Ecosystems to Molecules*. Cham, Switzerland: Springer International; 2017. pp. 171–194.
60. Hayman DTS, Emmerich P, Yu M, Wang LF, Suu-Ire R, Fooks AR, et al. Long-term survival of an urban fruit bat seropositive for ebola and lagos bat viruses. *PLoS One*. 2010; 5: 2008–2010. <https://doi.org/10.1371/journal.pone.0011978> PMID: 20694141
61. Ogawa H, Miyamoto H, Nakayama E, Yoshida R, Nakamura I, Sawa H, et al. Seroepidemiological Prevalence of Multiple Species of Filoviruses in Fruit Bats (*Eidolon helvum*) Migrating in Africa. *J Infect Dis*. 2015; 212: S101–S108. <https://doi.org/10.1093/infdis/jiv063> PMID: 25786916
62. De Nys HM, Mbala Kingebeni P, Keita AK, Butel C, Thaurignac G, Villabona-Arenas CJ, et al. Survey of ebola viruses in frugivorous and insectivorous bats in Guinea, Cameroon, and the democratic republic of the Congo, 2015–2017. *Emerg Infect Dis*. 2018; 24: 2228–2240. <https://doi.org/10.3201/eid2412.180740> PMID: 30307845
63. Towner JS, Amman BR, Sealy TK, Reeder Carroll SA, Comer JA, Kemp A, et al. Isolation of genetically diverse Marburg viruses from Egyptian fruit bats. *PLoS Pathog*. 2009; 5. <https://doi.org/10.1371/journal.ppat.1000536> PMID: 19649327
64. Amman BR, Carroll SA, Reed ZD, Sealy TK, Balinandi S, Swanepoel R, et al. Seasonal Pulses of Marburg Virus Circulation in Juvenile *Rousettus aegyptiacus* Bats Coincide with Periods of Increased Risk of Human Infection. *PLoS Pathog*. 2012; 8. <https://doi.org/10.1371/journal.ppat.1002877> PMID: 23055920
65. Amman BR, Jones MEB, Sealy TK, Uebelhoefer LS, Schuh AJ, Bird BH, et al. Oral shedding of Marburg virus in experimentally infected Egyptian fruit bats (*Rousettus aegyptiacus*). *J Wildl Dis*. 2015; 51: 113–124. <https://doi.org/10.7589/2014-08-198> PMID: 25375951

66. Jones MEB, Schuh AJ, Amman BR, Sealy TK, Zaki SR, Nichol ST, et al. Experimental inoculation of Egyptian rousette bats (*Rousettus aegyptiacus*) with viruses of the *ebolavirus* and *marburgvirus* genera. *Viruses*. 2015; 7: 3420–3442. <https://doi.org/10.3390/v7072779> PMID: 26120867
67. Kurosaki Y, Ueda MT, Nakano Y, Yasuda J, Koyanagi Y, Sato K, et al. Different effects of two mutations on the infectivity of Ebola virus glycoprotein in nine mammalian species. *J Gen Virol*. 2018; 99: 181–186. <https://doi.org/10.1099/jgv.0.000999> PMID: 29300152
68. Leroy EM, Kumulungui B, Pourrut X, Rouquet P, Hassanin A, Yaba P, et al. Fruit bats as reservoirs of Ebola virus. *Nature*. 2005; 438: 575–576. <https://doi.org/10.1038/438575a> PMID: 16319873
69. Taylor DJ, Leach RW, Bruenn J. Filoviruses are ancient and integrated into mammalian genomes. *BMC Evol Biol*. 2010; 10. <https://doi.org/10.1186/1471-2148-10-193> PMID: 20569424
70. Morvan JM, Deubel V, Gounon P, Nakouné E, Barrière P, Murri S, et al. Identification of Ebola virus sequences present as RNA or DNA in organs of terrestrial small mammals of the Central African Republic. *Microbes Infect*. 1999; 1: 1193–1201. [https://doi.org/10.1016/s1286-4579\(99\)00242-7](https://doi.org/10.1016/s1286-4579(99)00242-7) PMID: 10580275
71. Wittmann TJ, Biek R, Hassanin A, Rouquet P, Reed P, Yaba P, et al. Isolates of *Zaire ebolavirus* from wild apes reveal genetic lineage and recombinants. *Proc Natl Acad Sci U S A*. 2007; 104: 19656. <https://doi.org/10.1073/pnas.0704076104> PMID: 17942693
72. Rouquet P, Froment JM, Bermejo M, Kilbourn A, Karesh W, Reed P, et al. Wild animal mortality monitoring and human ebola outbreaks, Gabon and Republic of Congo, 2001–2003. *Emerg Infect Dis*. 2005; 11: 283–290. <https://doi.org/10.3201/eid1102.040533> PMID: 15752448
73. Leroy EM, Rouquet P, Formenty P, Souquière S, Kilbourne A, Froment JM, et al. Multiple Ebola Virus Transmission Events and Rapid Decline of Central African Wildlife. *Science* (80). 2004; 303: 387–390. <https://doi.org/10.1126/science.1092528> PMID: 14726594