Title

The Role of Anthropogenic Roosting Ecology in Shaping Viral Outcomes in Bats

Authors

Briana A Betke[1*], Nicole Gottdenker[2], Lauren Ancel Meyers[1], Daniel J Becker[3]

Affiliations

[1]Department of Integrative Biology, University of Texas at Austin, Austin, TX USA

[2]Department of Veterinary Pathology, College of Veterinary Medicine, University of Georgia, Athens, GA, USA

[3]School of Biological Sciences, University of Oklahoma, Norman, OK USA

*Correspondence to: Briana Betke

Email: babetke@utexas.edu

Abstract - 266 words (300 max)

The ability of some wildlife to live in anthropogenic structures in response to rapid land-use change is widely observed across mammals. However, the influence of this adaptation in shaping viral hosting ability and diversity are not well understood, especially for the order Chiroptera. Anthropogenic roosting may have important consequences for predicting virus spillover and spillback risk, particularly as the propensity of bats to roost in anthropogenic structures (e.g., buildings, bridges, homes, and tunnels, etc.) directly relates to human exposure. Here, we integrate novel roosting ecology data with a machine learning approach to assess the importance of anthropogenic roosting in predicting viral outcomes and evaluate if this novel trait improves prediction of undetected but likely host species. Our results show that the importance of anthropogenic roosting varies moderately across viral outcomes. Anthropogenic roosting is most important for predicting virus hosting ability across bats, followed by zoonotic hosting ability, viral richness, and the proportion of viruses that are zoonotic. Anthropogenic roosting status is less important than human population density but more important than most family, diet, and foraging traits of bat species, and models with anthropogenic roosting predict a narrowed list of undetected virus hosts compared to models excluding this trait. We identified 35 bat species likely to host a virus, 18 of which roost in anthropogenic structures. Additionally, we identified 51 undetected zoonotic host species, 30 of which are anthropogenic roosting. Maps of predicted virus host distributions show distinct spatial patterns between anthropogenic and exclusively natural-roosting bats. These findings suggest that anthropogenic roosting has a non-trivial role in shaping viral outcomes in bats, specifically virus hosting ability.

Keywords: synanthropic, urban, machine learning, zoonotic, macroecology

Introduction

As the number of emerging infectious disease (EID) events caused by zoonotic pathogens increase (Jones et al. 2008, Smith et al. 2014), there is a pressing need to identify drivers of pathogen transmission at the human–wildlife interface. Many EIDs are caused by viruses of human, domestic animal, and wildlife health concern (e.g., COVID-19, rinderpest, avian influenza; (Cleaveland et al. 2001, Woolhouse et al. 2008, Smith et al. 2014, Tompkins et al. 2015). Despite increased research efforts, our ability to accurately predict viral emergence is limited by comprehensive data on ecological factors driving viral emergence (Albery et al. 2021). One such factor is wildlife habitat use within human-dominated spaces, specifically wildlife living in anthropogenic structures (e.g., buildings, bridges, homes, and tunnels, etc.). Anthropogenic structures and the wildlife that occupy them represent direct entry points for cross-species transmission at the human–wildlife interface. Given this potential for direct interspecific contact, the propensity of wildlife to live in anthropogenic structures may have consequences for viral emergence by increasing contact between host species and facilitating pathogen introduction to novel hosts.

Prior meta-analyses and comparative studies have used broad urban presence and absence data to demonstrate the propensity for urban-attracted species to have greater diversity of pathogens (including but not limited to viruses) (Murray et al. 2019, Werner and Nunn 2020, Gibb et al. 2020, Albery et al. 2022). Select traits of many urban-adapted species, specifically their fast life histories, may facilitate harboring more pathogens through immune trade-offs and larger population densities (Albery and Becker 2021). In contrast to findings derived from only studying urban presence, how occupancy in human-made structures shapes pathogen transmission in wildlife remains poorly understood, with most work focusing on intra- rather

than interspecific comparisons. One rare interspecific comparison found that the propensity of rodent species to dwell in human-made structures increased their probability of being zoonotic hosts (Ecke et al. 2022). However, further investigation is needed to identify general patterns in the relationship between dwelling in human-made structures and pathogen outcomes across taxa to fully assess risk of both spillover and spillback events.

Relationships between occupying human-made structures and pathogen outcomes are especially poorly understood for bats, the mammalian order second only to rodents in hosting high virus diversity among mammals (Olival et al. 2017). Bats are central-place foragers, returning a roost after feeding to rest, socialize, reproduce, and/or thermoregulate during harsh weather conditions. While roosts provide protective environments for critical life history events, roosting behavior (e.g., social interactions, roost selection) and roost characteristics (e.g., structure type, size, microclimate) can create conditions that promote pathogen transmission. For example, larger roosts can increase contact rates between conspecifics and co-roosting species, driving density-dependent transmission (Laughlin et al. 2019, Ntumvi et al. 2021). Human disturbance further influences bat roosting ecology with implications for pathogen dynamics. Specifically, the redistribution and reduction of available roosting structures due to urbanization have led some species to occupy anthropogenic structures (Betke et al. 2023), many of which are hosts to viruses that are pathogenic to humans and other animals (Passos et al. 1998, Van der Poel et al. 2005, Streicker et al. 2010). Generally, anthropogenic roosting can affect viral dynamics through increased crowding and/or impairment of bat immune responses (Allen et al. 2009, Plowright et al. 2011). Anthropogenic roosting bats are also subject to increased human interactions, resulting in conflict that may further facilitate viral spillover and spillback. For instance, forced evacuation of bats from their anthropogenic roosts caused movement into other

unoccupied buildings, facilitating spatial spread of rabies virus among bat populations (Streicker et al. 2013).

Although several studies have considered the epidemiological importance of intraspecific variation in roosting ecology for a handful species, less is known of interspecific variation when predicting viral outcomes because roosting ecology data has not been compiled and standardized across bat species. This data sparsity challenges efforts to identify drivers of virus outcomes across bats, leading prior studies to omit roosting ecology (natural vs. anthropogenic) when testing hypotheses about bat–virus interactions at scale or building predictive models (Turmelle and Olival 2009, Han et al. 2016, Guy et al. 2019, 2020). In a prior analysis, we addressed this lack of standardized, comparative roosting data by compiling a novel database of natural and anthropogenic roosting status for 1,279 bat species (Betke et al. 2023). We showed anthropogenic roosting is common across bat families and geographies, with over a third of extant bat species roosting in anthropogenic structures. Additionally, anthropogenic roosting is highly predictable on the basis of traits that facilitate living in human-made structures (Betke et al. 2023). Importantly, the trait profile of an anthropogenic roosting bat species is distinct from that of general urban tolerance and is characterized by some slow life history traits, such as relatively small litter size.

In this study, we draw from this novel roosting ecology dataset to compare the importance of anthropogenic roosting to other previously identified factors associated with viral outcomes (i.e., viral richness, zoonotic proportion, and hosting ability) in bats, such as geographic range size, body mass and size, fecundity, and biogeography (Han et al. 2016, Plowright et al. 2019a, Guy et al. 2020). We first compiled a predictor matrix with trait data leveraged from the COalesced Mammal dataBase of INtrinsic and Extrinsic traits (COMBINE;

(Soria et al. 2021), PanTHERIA (Jones et al. 2009), and the International Union for the Conservation of Nature (IUCN; Gearty and Chamberlain 2022, IUCN 2022). Next, we used boosted regression trees to predict virus outcomes extracted from The Global Virome in One Network (VIRION) database (Carlson et al. 2022). We then quantified model performance with and without anthropogenic roosting to test if predictive ability improves with the addition of this epidemiologically relevant trait. Finally, we tested how such models differed in their predictions by also comparing predicted undetected virus and zoonotic virus host species provided by models with and without anthropogenic roosting ability.

Methods

Virus Data

Virus associations for 1,279 extant bat species matched to the most recent bat phylogeny were extracted from The Global Virome in One Network (VIRION) database (Upham et al. 2019, Carlson et al. 2022). VIRION is a comprehensive database that aggregates several host–virus association datasets to create a dynamic and up-to-date collection of host–virus associations with standardized viral taxonomy. These sources include CLOVER, the USAID Emerging Pandemic Threats PREDICT program, NCBI GenBank, and the Global Biotic Interactions (GLOBI) database. We restricted VIRION to NCBI-verified host and virus species names. Two measures of virus richness were derived from VIRION: total viral richness and zoonotic virus richness. Total viral richness is the number of all known viral species for a given bat species (Albery et al. 2022). Zoonotic viruses were identified as any bat viruses also found in humans within VIRION. To identify the trait profile of bat species that disproportionately host zoonotic viruses specifically, we then derived the proportion of zoonotic viruses per host species. Two binary variables for virus and zoonotic host status were derived from the richness variables to predict

potential host and zoonotic host species. Any bat species not known to have any viral associations in VIRION were designated a value of zero (i.e., a pseudoabsense).

Anthropogenic roosting data

We used an established dataset of anthropogenic roosting ecology manually compiled from existing literature and historical records (Betke et al. 2023). This dataset consists of roosting ecology across 80% (n = 1,029) of all 1,460 known bat species to date (Simmons and Cirranello. 2023) and is well represented across bat families, biogeographic regions, and IUCN status. Roosting in human-made structures is described as a binary classification, where one indicates anthropogenic roosting (e.g., mines, buildings, homes, bridges, tunnels, roofs, and attics) and zero indicates exclusively natural roosting (e.g., palm leaves, caves, tree hollows, foliage).

Additional bat trait data

For additional bat traits, including but not limited to those included in prior predictive models of bat virus outcomes (Turmelle and Olival 2009, Han et al. 2016, Plowright et al. 2019a, Guy et al. 2020), we primarily used COMBINE, a recent repository containing 54 mammalian traits aggregated from several commonly used sources (Soria et al. 2021). Traits include important life history variables from PanTHERIA (Jones et al. 2009) and foraging attributes from EltonTraits (Wilman et al. 2014), both of which have commonly been used in prior analyses to identify drivers of pathogen diversity and identify undetected reservoir species (Luis et al. 2013, Ostfeld et al. 2014, Han et al. 2015, Olival et al. 2017, Becker et al. 2018, Becker and Han 2021, Albery et al. 2022, Mull et al. 2022). We restricted COMBINE to non-imputed trait data, as imputation of missing trait data may not account for strong biases in coverage (Johnson et al. 2021), such

biases likely exist for small bats, those with low population size, and those existing far from humans. Biogeographic areas were expanded to binary variables for species occupancy in each biogeographic region (i.e., Afrotropical, Australian, Indomaylan, Nearctic, Neotropical, Oceanian, Palearctic).

Additional geographic variables (i.e., geographic area, human population density, mean precipitation, mean monthly actual evapotranspiration rate [AET], mean monthly potential evapotranspiration rate [PET], minimum and maximum latitude, and mean temperature) were collected from the PanTHERIA database. To account for the shared evolutionary history of bat hosts and their impact on viral outcomes, bat families were included by generating binary variables of all 18 bat families (Han et al. 2016). This trait matrix was merged and taxonomically reconciled with the host–virus association dataset to serve as the trait matrix in this analysis.

Assessing trait data coverage and sampling effort

We evaluated all trait variables for variance and coverage prior to downstream modeling. Any variables found to have no variation or coverage under 30%, indicating mostly missing values, were removed from the dataset (Fig. S1, Supporting Information). We thus excluded 36 out of 99 variables from our analysis, leading to 63 total variables in the dataset in addition to our primary anthropogenic roosting variable (Supporting Information).

Sampling is not equal across hosts and could influence modeling viral outcomes, as some species are studied more due to greater accessibility, proximity to humans, and awareness. Citation count is commonly used as a proxy for sampling bias (i.e., sampling intensity) in trait-based analyses (Olival et al. 2017, Guy et al. 2019, 2020, Albery et al. 2022, Mull et al. 2022). To control for sampling bias in our models, we calculated the number of overall host

citations and virus-related host citations associated with each bat species. A PubMed search was conducted in R with the *easyPubMed* package (Fantini 2019) to calculate both measures of sampling effort for all bat species included in our dataset (Olival et al. 2017, Albery et al. 2022, Mull et al. 2022).

Boosted regression tree analysis

We built boosted regression tree (BRT) models with the *gbm* package for each of our four response variables to identify traits predictive of overall viral richness, the proportion of zoonotic virus, and both measures of host status. Bernoulli-, Poisson-, or Gaussian-distributed error structures were used for binary, count, and proportion data (normalized with the Freeman– Tukey double arcsine transformation via the *metafor* package), respectively. BRTs are machine learning algorithms that determine variables of predictive importance based on an ensemble of regression or classification trees, even in the presence of missing data and non-linear responses, making these models ideal for analyzing ecological trait data (Elith et al. 2008). Trait-based BRTs have also been used to identify unknown hosts of viruses specifically and are found to outperform other modeling approaches for the study of host–parasite associations more generally (Becker et al. 2022).

Before fitting BRT models, we partitioned our data between training (80%) and test (20%) datasets using the *resample* package for internal model validation (measured via the area under the receiver operator characteristic curve [AUC] with the *ROCR* package for binary models), across 100 different random data partitions. Models were run with optimal parameters determined via a grid search with 10-fold cross-validation to prevent overfitting (Fig. S2, Supporting Information). Resulting variable importance values and performance metrics were

196 averaged across these 100 model runs. We extracted importance rankings for anthropogenic
197 roosting and used *t*-tests to compare ranks of this variable across our four virus outcomes, using
198 the Benjamini–Hochberg correction to adjust *p*-values for pairwise comparisons (Benjamini and
199 Hochberg 1995). We used partial dependence plots to visualize the effect of anthropogenic
200 roosting on each viral outcome, after controlling for other features and citation counts.
201 Additionally, host citation count and virus-related host citation count were modeled with the
202 same explanatory variables as a Poisson-distributed response to determine if viral outcomes are
203 heavily influenced by study effort (Han et al. 2015, Plowright et al. 2019a, Mull et al. 2022).
204
205 *Model performance and prediction*
206 Performance metrics were calculated for the appropriate error distributions of each of our four
207 response variables (i.e., AUC, sensitivity, specificity, pseudo R^2). We tested how these
208 performance measures varied between models with and without the anthropogenic roosting
209 variable with unpaired *t*-tests, accompanied by Cohen's *d* effect sizes and associated *p*-values.
210 The *p*-values were again adjusted with the Benjamini–Hochberg correction when evaluating
211 model performance metrics. From our binary models, we also predicted unknown (i.e., not
212 currently known to host any viruses) overall virus and zoonotic virus host species, where species
213 in the 95th percentile were considered likely hosts (Han et al. 2015, Guy et al. 2020, Mull et al.
214 2022). These predicted host species were compared between models with anthropogenic roosting
215 included and excluded. To correct for sampling bias, predictions were made using mean host
216 citation counts and mean virus-related host citation counts across species (Mull et al. 2022,
217 Becker et al. 2022). Lastly, we visualized the spatial distributions of predicted hosts from
218 anthropogenic roosting models if the predictions differed from those of models without

219 anthropogenic roosting. Predicted host species in the 95th percentile were mapped by overlaying
220 shapefiles obtained from the IUCN Red List database of mammal geographic ranges (IUCN
221 2022).

222 **Results**

223 *Data description*

224 Of the 1,279 bat species included in our novel dataset of roosting ecology (Betke et al. 2023),
225 40% of bats are anthropogenic roosting and 42% exclusively use natural roosts, with this trait
226 represented across all bat families (Fig. 1A). Matching VIRION data to our anthropogenic
227 roosting dataset revealed a wide range of virus richness, between 0 to 554 virus species. The bat
228 species with the highest total identified viral richness and zoonotic richness recorded was
229 *Eidolon helvum* at 554 and 22 virus species, respectively. These data contained 381 bat species
230 with known associations with at least one virus, for which 244 of these species were
231 anthropogenic roosting. Similarly, 226 bat species were found to host at least one zoonotic virus,
232 of which 175 species were anthropogenic roosting. We observed distinct patterns in virus
233 richness across bat families and roosting status (Fig. 1B). Descriptive density plots suggested the
234 families Rhinolophidae, Pteropodidae, Mormoopidae, and Miniopteridae had higher identified
235 virus richness among anthropogenic roosting species; however, these trends were less apparent
236 for zoonotic proportions, except for the families Natalidae and Mormoopidae (Fig. 1C).
237

238 *Importance of anthropogenic roosting in predicting viral outcomes*

239 Overall, anthropogenic roosting did not rank highly as a predictor among 62 other features (Fig.
240 2). Our BRTs predicting virus and zoonotic virus host status showed that anthropogenic roosting

status is less important than human population density within a species range but more important than most family, diet, and foraging traits (Fig. S3, Supporting Information). Although anthropogenic roosting ranked only moderately in importance, other traits relating to geography, life history, and taxonomy were among the most important predictors of virus outcomes (Fig. S4, Supporting information).

Next, we used variable ranks derived from our BRTs to compare relative importance of anthropogenic roosting across virus outcomes. Anthropogenic roosting ability was ranked highest for predicting virus host status, with a median ranking of 28 (range, 15-51) out of 63 variables (Fig. 2). The median rankings were similar for virus richness and zoonotic host status models at 37 (range, 26-44) and 35 (range, 28-47), respectively. The lowest median ranking was the double arcsine transformed zoonotic virus richness at 42 (range, 28-50). All pairwise comparisons of average ranking for anthropogenic roosting between viral outcomes were significantly different (Fig. 2). After adjusting for multiple comparisons (Table S2, Supporting Information), the average ranking of anthropogenic roosting for predicting viral richness was significantly higher than zoonotic proportion ($t = -9.42$, $p < 0.001$), but lower than viral host status ($t = 9.93$, $p < 0.001$) and zoonotic host status ($t = 2.47$, $p < 0.05$). For zoonotic proportion, anthropogenic roosting ranked significantly lower than viral host status ($t = 15.34$, $p < 0.001$) and zoonotic host status ($t = 10.97$, $p < 0.001$). Lastly, the average ranking for viral host status was significantly lower than zoonotic host status ($t = -8.13$, $p < 0.001$).

Limited relative importance of anthropogenic roosting was also reflected in partial dependence plots showing the effect of anthropogenic roosting on virus outcomes while fixing all other trait variables and citation counts. The effect of anthropogenic roosting did not substantially differ from natural roosting across viral richness, zoonotic proportion, and zoonotic

host status (Fig. 3). However, a weak effect of anthropogenic roosting was found for viral host status where anthropogenic roosting bats were slightly more likely to host a virus than natural roosting bats (Fig. 3).

Model performance and anthropogenic roosting

BRT models containing anthropogenic roosting predicted overall virus richness with high accuracy (pseudo-R^2 = 0.64 ± 0.02). However, models with our roosting variable poorly predicted zoonotic proportion (pseudo-R^2 = 0.15 ± 0.01). Virus hosts were distinguished with high accuracy ($\overline{AUC}$ = 0.90 ± 0.002), specificity ($\overline{Specificity}$ = 0.93 ± 0.002), and moderate sensitivity ($\overline{Sensitivity}$ = 0.66 ± 0.01). Our models classified zoonotic hosts in a similar fashion to those of virus hosts ($\overline{AUC}$ = 0.91 ± 0.002, $\overline{Specificity}$ = 0.96 ± 0.001 , $\overline{Sensitivity}$ = 0.62 ± 0.01).

When comparing the performance of these models to those without anthropogenic roosting, all differences were negligible. Richness models without anthropogenic roosting had equivalent performance (pseudo-R^2 = 0.64 ± 0.02; t = 0.06, p = 0.998; Cohen's d = 0.01). This lack of difference in performance was also observed for zoonotic proportion models without anthropogenic roosting (pseudo-R^2 = 0.15 ± 0.01; t = 0.01, p = 0.998; Cohen's d = 0.001) and zoonotic host status ($\overline{AUC}$ = 0.91 ± 0.002; t = 0.002, p = 0.998; Cohen's d = 0.0003). Similarly, sensitivity ($\overline{Sensitivity}$ = 0.62 ± 0.01; t = 0.04, p = 0.998; Cohen's d = 0.006) and specificity ($\overline{Specificity}$ = 0.96 ± 0.001; t = -0.05, p = 0.998; Cohen's d = -0.007) did not differ between models for zoonotic host status. Virus host models without anthropogenic roosting also did not vary in performance ($\overline{AUC}$ = 0.90 ± 0.002; t = 0.03, p = 0.998; Cohen's d = 0.004), specificity (

$\overline{Specificity}$ = 0.93 ± 0.002; t = -0.14, p = 0.998; Cohen's d = -0.02), or sensitivity ($\overline{Sensitivity}$ = 0.66 ± 0.01; t = 0.21, p = 0.998; Cohen's d = 0.03).

To assess if trait profiles of viral outcomes were driven by sampling effort, we ran BRT models predicting both host citation count and virus-related host citation count using the same trait predictors as all other models. We found that the included traits poorly predicted both citation count (pseudo-R^2 = 0.12 ± 0.08) and virus citation count (pseudo-R^2 = -0.09 ± 0.14).

Anthropogenic roosting and model predictions

Predicted probabilities for virus hosts varied slightly between models with and without anthropogenic roosting (Fig. 4A). Additionally, predicted probabilities of being a virus host were greater for anthropogenic roosting species (Fig. 4B). There was no variation for the probability of being a zoonotic host between models (Fig. 5A), though we observed a greater density of zoonotic host species for anthropogenic roosting in predicted novel hosts (Fig. 5B).

Our models with anthropogenic roosting predicted 35 virus host species, 18 of which were anthropogenic roosting bats. Models without anthropogenic roosting identified 12 more virus hosts than models with anthropogenic roosting (Table S3, Supporting information). Fifty-one species were predicted to host at least one zoonotic virus, 30 of which were anthropogenic roosting (Table S4, Supporting Information). The zoonotic host models with roost status predicted one additional species than models without roost status. When we stratified these results by bat families, the Vespertilionidae contained the highest proportion of novel virus and zoonotic virus hosts (13/35 and 17/51, respectively). The family with the next-highest number of predicted novel virus host species was the Miniopteridae (8/39). For predicted zoonotic novel host species, the Phyllostomidae followed the Vespertilionidae, with 13 species out of 51. The

309 majority of undetected virus hosts were predicted to be in Afrotropical and Indomalayan
310 biogeographical realms, with 15 and 8 species out of 35, respectively. Neotropical and
311 Afrotropical realms contained 27 and 15 of the 51 undetected zoonotic hosts, respectively.
312 We lastly mapped the distribution of known and novel hosts predicted from the overall
313 virus host model to observe spatial differences in the density of anthropogenic and natural
314 roosting novel host species (Fig. 6). Anthropogenic roosting novel hosts largely had distinct
315 distributions from the novel hosts that exclusively roost in natural structures. This result was
316 especially apparent in the Afrotropical and Indomalayan regions.

317 **Discussion**

318 As the rate of pathogen emergence increases (Jones et al. 2008, Smith et al. 2014), it is important
319 to characterize and synthesize species traits that relate directly to the wildlife–human interface
320 where spillover and spillback can occur. Roosting ecology, particularly the structures in which
321 bats roost, may serve as points of relevant cross-species transmission (McKee et al. 2019, Lunn
322 et al. 2021). Investigating anthropogenic roosting ability provides an opportunity to assess
323 patterns in pathogen outcomes that may not be apparent when considering general urban
324 presence (e.g., only temporary occupancy in an urban environment). We applied a machine
325 learning approach to determine the role of anthropogenic roosting in shaping four pathogen
326 outcomes for viruses specifically (i.e., virus richness, zoonotic proportion host status, and
327 zoonotic host status). We then assessed the importance of anthropogenic roosting in predicting
328 viral outcomes and tested if this novel trait improves prediction of undetected but likely host
329 species.
330 Of 63 total trait variables, roosting ecology was not the most important predictor of any
331 of our four viral outcomes, and assessment of model performance showed that inclusion of this

332 roosting trait only moderately improved predictive capacity for some responses. In particular, we
333 found that anthropogenic roosting bats had a slightly increased probability of hosting at least one
334 virus. These findings do not support our prediction that anthropogenic roosting would promote
335 viral richness and more frequent detection of zoonotic viruses, thus improving predictive ability
336 compared to models without this novel trait. Notably, predictions from virus host models
337 incorporating roosting ecology largely agreed with those from models without this trait but
338 produced a narrowed list of species despite the small improvement in performance. Further, the
339 distribution of these novel hosts differed in subtle but important ways between anthropogenic
340 and natural roosting species. Our results suggest that while anthropogenic roosting ecology is
341 only moderately important in predicting viral outcomes, the inclusion of this trait may provide
342 different predictions compared to models without this trait.

343 Our study is among a growing body of research concerning the role that wildlife living
344 alongside humans have in determining pathogen diversity; however, we uniquely use a more
345 narrowed definition of presence in human-occupied areas to directly assess relationships at the
346 human–wildlife interface. Rather than a broad urban-dwelling criterion (Santini et al. 2019) or
347 variables such as human population density within a species' range to approximate
348 human–wildlife contact (Olival et al. 2017), we focus solely on the ability of a species to roost
349 within human-made structures. Our findings support previous large-scale comparative analyses
350 in which mammals occupying human-dominated areas do not disproportionately host greater
351 zoonotic pathogen diversity (Albery et al. 2022), but our results contradict the finding that urban
352 wildlife (bats specifically) host greater viral richness. Unlike the intraspecific study of the
353 zoonotic hosting capacity of anthropogenic rodents (Ecke et al. 2022), our results do not strongly
354 support that anthropogenic bats are more likely to host zoonotic viruses, suggesting

anthropogenic dwelling may not play as strong of a role in predicting zoonotic bat hosts as is found in rodents.

Our results suggest that anthropogenic roosting bats have slightly greater probabilities of viral hosting than natural roosting bats but do not differ in their viral richness. This weak pattern may be explained by differences in the trait profiles of virus hosts compared to other viral outcomes. With the exception of citation count and other key traits (i.e., geographic range, maximum latitude, and body length), the top determinants of virus richness differed from those of viral hosting ability. Our BRTs predicting virus hosts identified climate variables (i.e., average AET, PET, and temperature) to be among the top ten predictors, while our richness models did not identify these variables. In our previous work, we showed that anthropogenic roosting bat species are characterized by similar climatic factors (Betke et al. 2023). The small difference in marginal effect between anthropogenic and natural roosting bats for viral hosting may therefore reflect this climactic similarity, suggesting that the trait profile of anthropogenic roosting bats somewhat aligns with the profile of viral hosts more than that of bat species with greater viral richness. Additionally, other traits may further explain the weak importance of roosting ecology for predicting viral outcomes. For example, we identified broad geographic range to be a top predictor across viral outcomes, in alignment with prior work on bat–virus associations (Turmelle and Olival 2009, Han et al. 2016, Plowright et al. 2019a, Guy et al. 2020), whereas we previously showed anthropogenic roosting bat species are characterized by narrow geographic ranges (Betke et al. 2023). Large body size and mass were also important predictors of zoonotic proportion and ability to host zoonotic viruses, whereas anthropogenic roosting bats are characterized by small body size (Betke et al. 2023). Lastly, our findings did not challenge the idea that fast-lived species tend to be urban tolerant and more likely to host viruses. Species that

had greater viral richness tended to have larger litter sizes, while anthropogenic bats, conversely, tend to have smaller litter sizes.

Although models with and without roosting ecology performed similarly, we found differences in their predictions of viral hosts. Models without roosting ecology predicted 47 undetected hosts in the 95th percentile, whereas models with roosting ecology detected 35 species. All species predicted by the models with roosting ecology were in agreement with the non-roosting ecology models but excluded 12 species. This reduction in undetected hosts could suggest that exclusion of roosting ecology may lead to false positives. Accurately identifying likely virus hosts is especially important for both study design, given the financial and logistical costs of sampling (Plowright et al. 2019b) as well as for bat conservation, by minimizing unnecessary sampling of vulnerable species. For the latter point, anthropogenic roosting bats may be especially vulnerable to biodiversity loss and are further threatened by forced evacuation due to their reliance on human-made structures (Betke et al. 2023). In contrast to the viral hosting models, zoonotic host models with roosting ecology detected only one additional species than analogous models without roosting ecology. This result was expected, given the lack of difference in performance between models with and without roosting ecology.

The majority of undetected viral hosts in the 95th percentile of predictions were concentrated in the Afrotropical and Indomalayan regions. Despite the lack of empirical support for African bats to serve as reservoir hosts of viruses due to strong sampling biases (Weber et al. 2023), 15 of our 35 undetected species were predicted to host viruses in the Afrotropics. This narrowed species list could serve as a useful starting point to geographically expand sampling effort and accurately confirm viral hosts in this region while reducing the logistical costs associated with broad sampling. The Indomalayan region was second to the Afrotropics, with

401 eight undetected viral host species. As most predicted species in both regions have zero citations,
402 this suggests our models are capable of identifying potential hosts regardless of poor historic
403 sampling effort. Additionally, half of these 35 undetected viral hosts were anthropogenic roosting
404 species. Range maps of these novel viral hosts highlight distinct spatial patterns between natural
405 roosting and anthropogenic roosting bats. While novel hosts occur in Central America and
406 northwestern South America, anthropogenic roosting bats are in the Brazilian Caatinga and
407 Atlantic Forest. Similarly, anthropogenic roosting novel hosts are more concentrated in southern
408 Africa, whereas natural roosting species are distributed in equatorial regions of Africa. Natural
409 roosting species are more clustered in Southeast Asia whereas anthropogenic species are spread
410 across the Malay Archipelago, and in South Asia. The prediction of so many undetected hosts
411 and clear differences in their geographic distribution according to roosting ecology in such
412 disproportionately undersampled regions highlights the need to prioritize sampling effort in the
413 Afrotropical and Indomalayan realms in particular to better understand virus-bat interactions.
414 While all host and virus species names included in our analyses were verified through the
415 NCBI taxonomy, we used a broad definition of hosting to calculate viral outcomes from
416 VIRION; meaning that we did not distinguish between serological and other detection methods
417 (e.g. PCR, virus isolation) when defining associations. Including indicators of exposure such as
418 antibody testing may lead to overcounting viral associations given known issues of
419 cross-reactivity (Gilbert et al. 2013), even if there is no evidence for a species to successfully
420 transmit the associated viruses (i.e., competence; Becker et al. 2020). Integrating data on host
421 competence, and empirically testing predicted hosts for viral competency, could enhance
422 predictions of statistical models, advancing our understanding of spillover and spillback risks.
423 Furthermore, using models informed by host competence may help mitigate stigma against bats

424 labeled as reservoirs solely based on serological or PCR evidence (Shapiro et al. 2021, Weber et
425 al. 2023).

426 To evaluate anthropogenic roosting at the host species scale, we collapsed the bat–virus
427 network, leading to a loss of information relating to viral traits. The inclusion of viral
428 characteristics such as nucleic acid (DNA vs RNA), enveloped vs non-enveloped, or
429 transmission mode (vector, environmental, direct contact) could demonstrate variation in the
430 importance of anthropogenic roosting for predicting associations with particular viruses. Future
431 models could focus explicitly on include link prediction to better integrate both host and virus
432 characteristics into analyses of the host–virus network. Additionally, our modeling approach with
433 roosting ecology and link prediction suggested here could be readily adapted to trait-based
434 analyses of other pathogens. Bat–bacterial associations are especially understudied, as research
435 effort is increasingly focused on determinants of viral outcomes (Shaw et al. 2020, Szentivanyi et
436 al. 2023).

437 In conclusion, we leveraged recently compiled data on roosting ecology to explore how
438 anthropogenic roosting shapes viral outcomes across bats. While roosting ecology was not a
439 highly predictive trait, our results show that predicted viral hosts are distinctly distributed based
440 on their roosting ecology and that anthropogenic roosts are a meaningful interface where
441 bat–human contact occurs. A better understanding of how anthropogenic roosting influences
442 bat–virus interactions through the application of macroecological tools and further integration
443 with studies of bat–human interactions within ecological, cultural, behavioral and socioeconomic
444 contexts can provide holistic frameworks to understand as well as prevent transmission of
445 bat-borne zoonoses or viral pathogens of concern to threatened species. The incorporation of
446 methodologies such as link prediction models with roosting ecology as well as greater coverage

of viral trait and host competency data could allow for improved predictions of viral hosts while minimizing invasive sampling of vulnerable wildlife species.

Author Contributions

Conceptualization, D.B., B.B.; Methodology, D.B. and B.B.; Formal Analysis, B.B.; Investigation, B.B. and D.B.; Data Curation, B.B. and D.B.; Writing–Original Draft, B.B. and D.B., Writing –Review & Editing; B.B., D.B., L.M. and N.G.; Supervision, D.B.

Acknowledgements

B.B. and D.J.B. were supported by funding to the Viral Emergence Research Initiative (VERENA) Institute, including NSF BII 2021909 and NSF BII 2213854.

Declaration of Interests

The authors declare no competing interests.

Data Availability

The viral association dataset used in this analysis can be accessed at the VIRION GitHub repository (https://github.com/viralemergence/virion). All code and data associated with this analysis can be found at the bathaus GitHub repository (https://github.com/viralemergence/bathaus).

Figures

Figure 1. Distribution of anthropogenic roosting and viral outcomes across bat families. (A) Number of bats per roosting type across all families. (B) Density ridgeline plots of log virus richness of bats faceted by families with observations greater than 5. (C) Density ridgeline plots of zoonotic virus richness faceted by families with observations greater than 5. Distributions are colored by anthropogenic (green), natural (purple), and unknown roosting bats (gray).

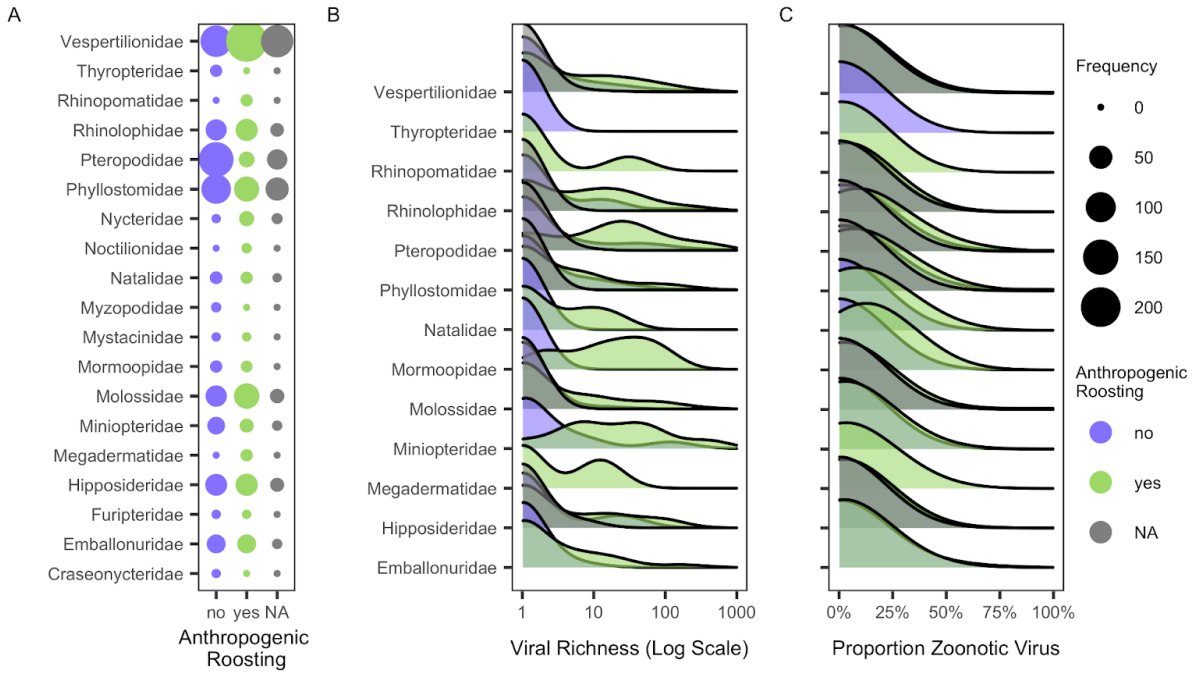


Figure 2. Differences in ranking of our anthropogenic roosting variable across viral outcomes. Colored dots represent rankings for each individual model run (100 random splits of training and test data). Violin plots show the distribution of rankings for each respective viral outcome. Pairwise comparisons between viral outcomes were evaluated using unpaired *t*-tests adjusted for multiple comparisons (**** $P < 0.0001$).

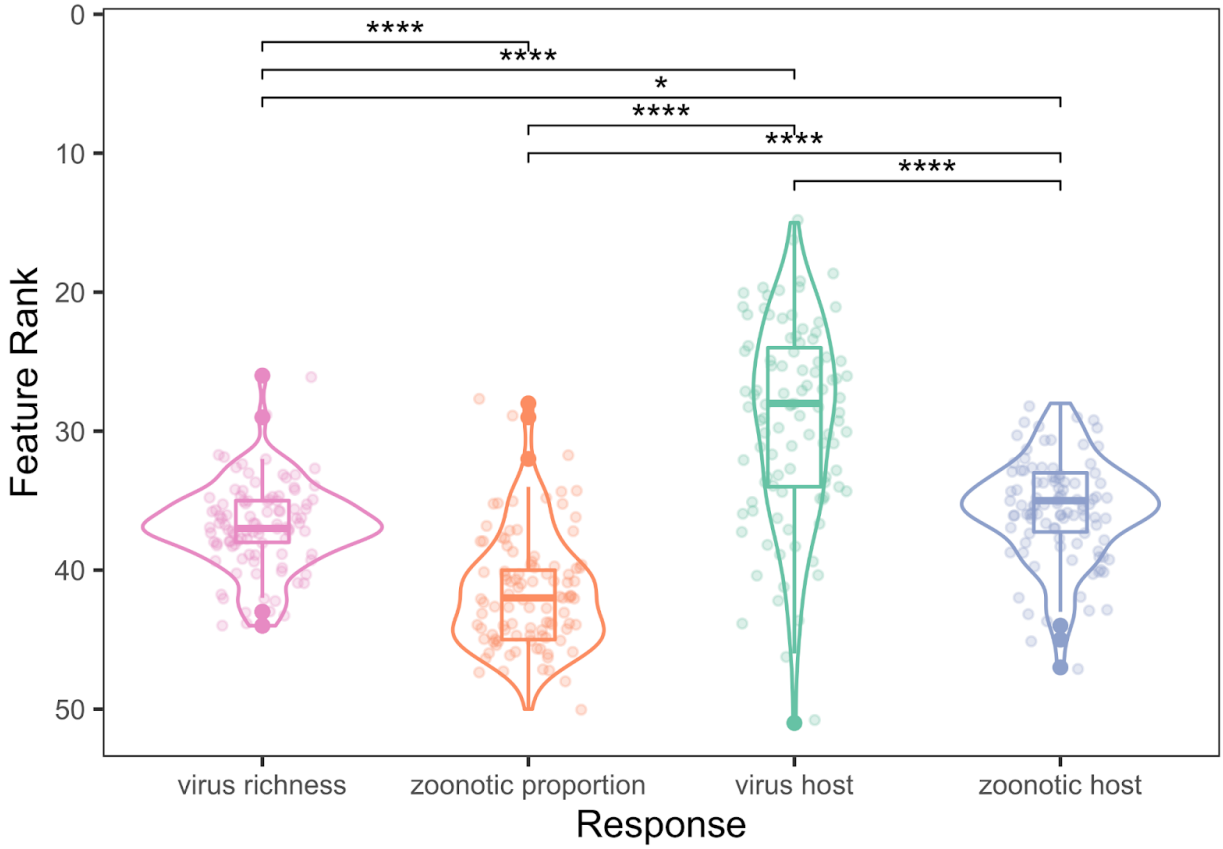


Figure 3. Partial dependence plots for anthropogenic roosting for all viral outcomes. Colored dots represent marginal effects for each individual model run (100 random splits of training and test data), and the means are indicated with black squares. Dot colors correspond to viral outcome where virus richness is pink (A), zoonotic proportion is orange (B), virus host status is green (C), and zoonotic host is blue (D).

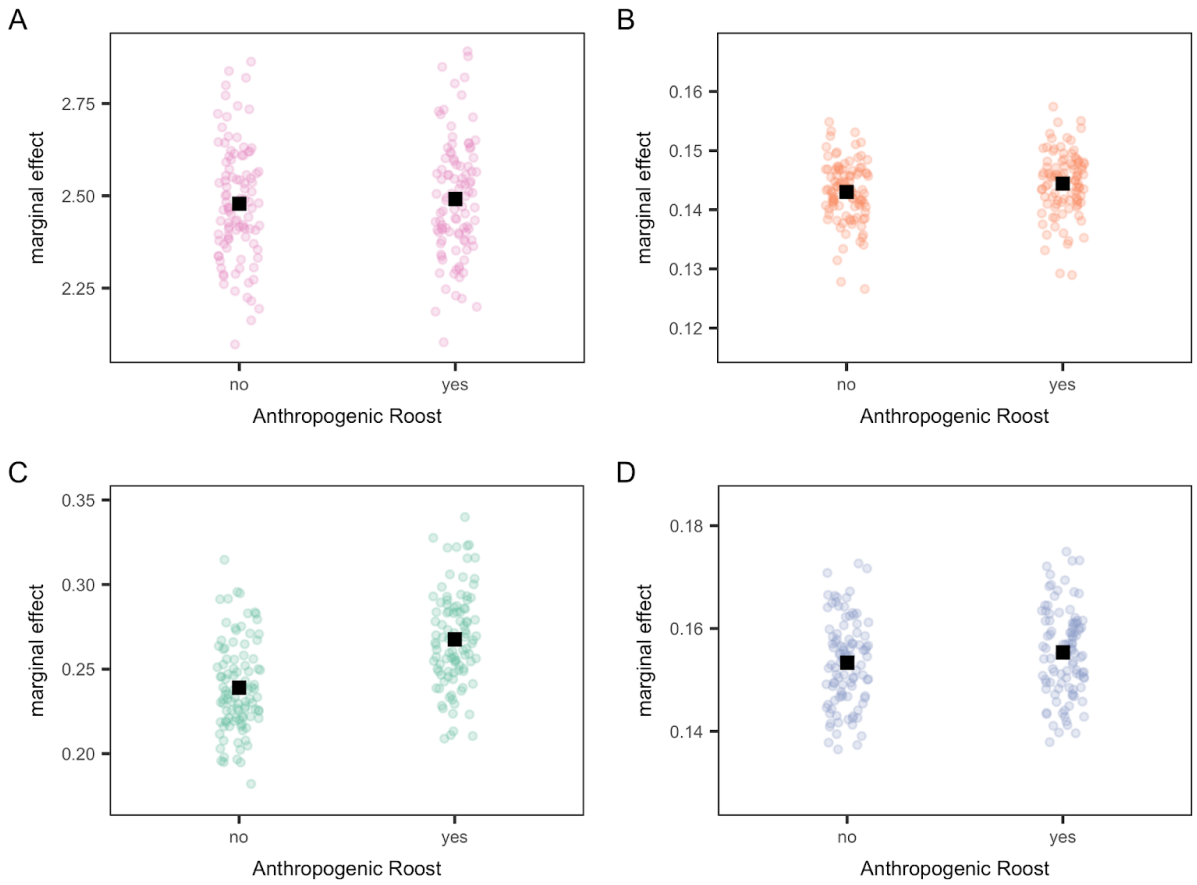


Figure 4. Predicted probability of hosting a virus for models with and without anthropogenic roosting as a variable. (A) The predicted probabilities of hosting of known and unknown hosts from anthropogenic roosting models stratified by anthropogenic roosting status. (B) The differences in average predictions between the models with anthropogenic roosting and the models without.

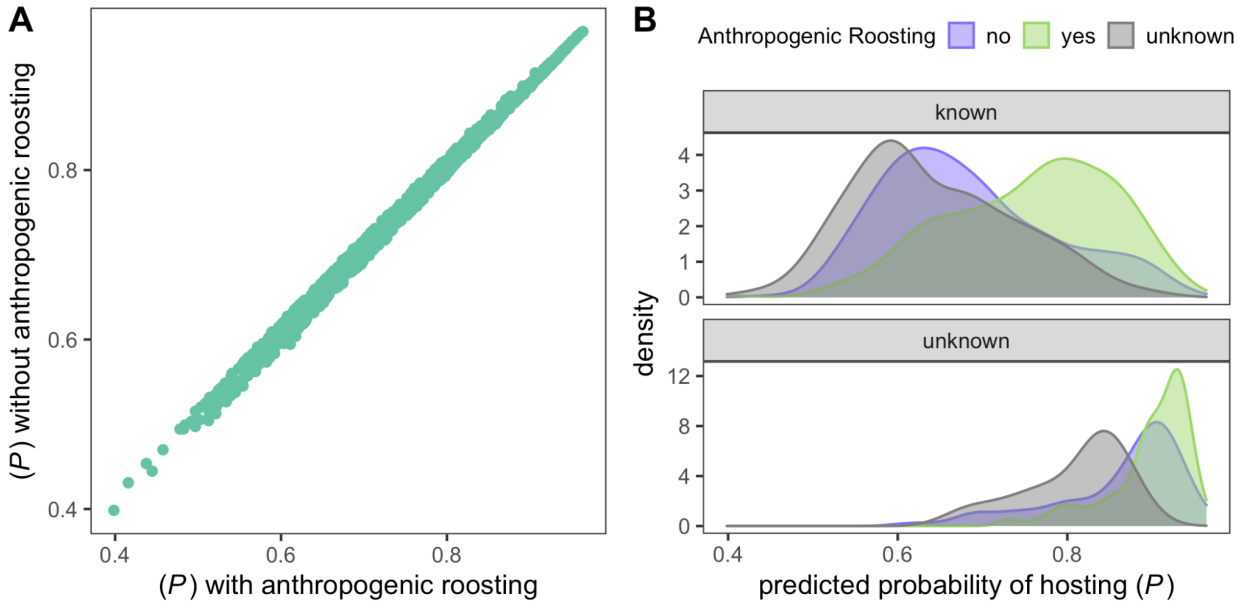


Figure 5. Predicted probability of hosting a zoonotic virus for models with and without anthropogenic roosting as a variable. (A) The predicted probabilities of hosting of known and unknown hosts from anthropogenic roosting models stratified by anthropogenic roosting status. (B) The differences in average predictions between the models with anthropogenic roosting and the models without.

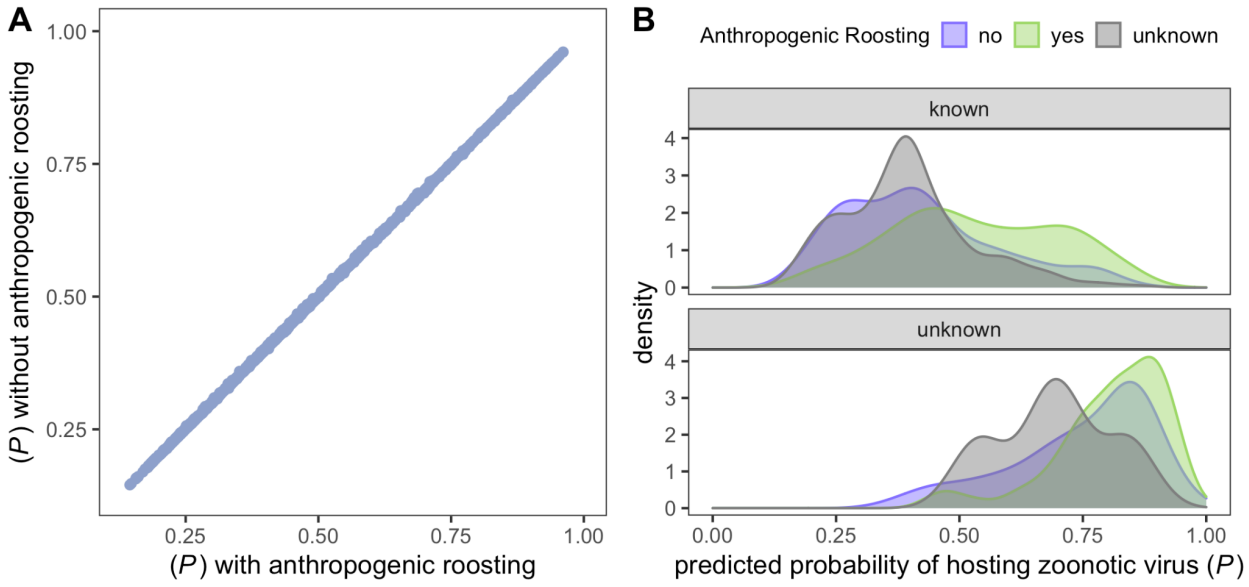
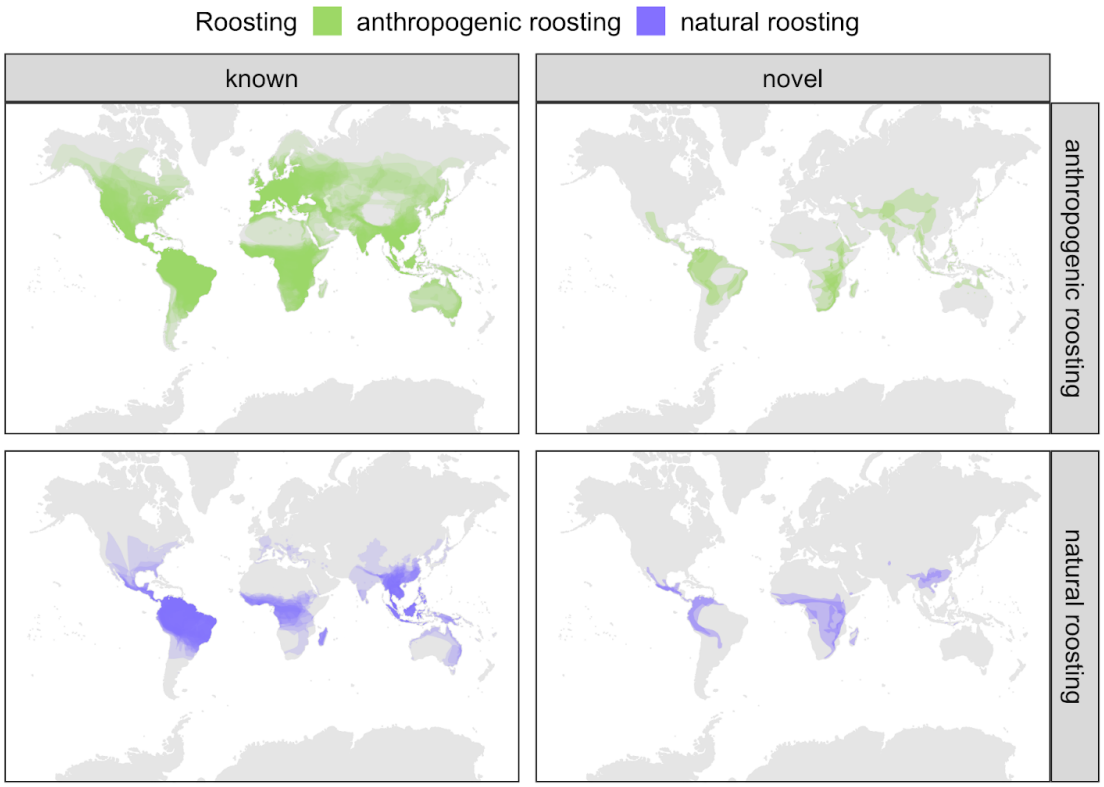


Figure 6. Distribution of virus hosts. The known column represents the distribution of known viral hosts for natural (top left) and anthropogenic roosting species (bottom left). The novel column represents the distribution of novel hosts predicted in the 95th percentile by models with anthropogenic roosting for natural (top right) and anthropogenic roosting species (bottom right).



References

Albery, G. F. and Becker, D. J. 2021. Fast-lived Hosts and Zoonotic Risk. - Trends Parasitol. 37: 117–129.

Albery, G. F. et al. 2021. The science of the host–virus network. - Nature Microbiology 2021 6:12 6: 1483–1492.

Albery, G. F. et al. 2022. Urban-adapted mammal species have more known pathogens. - Nature Ecology & Evolution: 794–801.

Allen, L. C. et al. 2009. Roosting ecology and variation in adaptive and innate immune system function in the Brazilian free-tailed bat (Tadarida brasiliensis). - J. Comp. Physiol. B 179: 315–323.

Becker, D. J. and Han, B. A. 2021. The macroecology and evolution of avian competence for Borrelia burgdorferi. - Glob. Ecol. Biogeogr. 30: 710–724.

Becker, D. J. et al. 2018. Using host species traits to understand the consequences of resource provisioning for host–parasite interactions. - J. Anim. Ecol. in press.

Becker, D. J. et al. 2020. Beyond Infection: Integrating Competence into Reservoir Host Prediction. - Trends Ecol. Evol. in press.

Becker, D. J. et al. 2022. Optimising predictive models to prioritise viral discovery in zoonotic reservoirs. - The Lancet Microbe 3: e625–e637.

Benjamini, Y. and Hochberg, Y. 1995. Controlling the false discovery rate: A practical and powerful approach to multiple testing. - J. R. Stat. Soc. 57: 289–300.

Betke, B. A. et al. 2023. Ecological and Evolutionary Characteristics of Anthropogenic Roosting Ability in Bats of the World. - bioRxiv: 2023.10.15.562433.

Carlson, C. J. et al. 2022. The Global Virome in One Network (VIRION): an Atlas of Vertebrate-Virus Associations. - MBio 13: e0298521.

Cleaveland, S. et al. 2001. Diseases of humans and their domestic mammals: pathogen characteristics, host range and the risk of emergence. - Philos. Trans. R. Soc. Lond. B Biol. Sci. 356: 991.

Ecke, F. et al. 2022. Population fluctuations and synanthropy explain transmission risk in rodent-borne zoonoses. - Nat. Commun. 13: 7532.

Elith, J. et al. 2008. A working guide to boosted regression trees. - J. Anim. Ecol. 77: 802–813.

Fantini, D. 2019. easyPubMed: Search and Retrieve Scientific Publication Records from PubMed. in press.

Gearty, W. and Chamberlain, S. 2022. rredlist: “IUCN” Red List Client. - R package version 0.7.1. https://CRAN.R-project.org/package=rredlist in press.

Gibb, R. et al. 2020. Zoonotic host diversity increases in human-dominated ecosystems. - Nature 584: 398–402.

Gilbert, A. T. et al. 2013. Deciphering serology to understand the ecology of infectious diseases in wildlife. - Ecohealth 10: 298–313.

Guy, C. et al. 2019. Phylogeny matters: Revisiting “a comparison of bats and rodents as reservoirs of zoonotic viruses.” - Royal Society Open Science in press.

Guy, C. et al. 2020. The influence of bat ecology on viral diversity and reservoir status. - Ecol. Evol. 10: 5748–5758.

Han, B. A. et al. 2015. Rodent reservoirs of future zoonotic diseases. - Proc. Natl. Acad. Sci. U. S. A. 112: 7039–7044.

Han, B. A. et al. 2016. Undiscovered Bat Hosts of Filoviruses. - PLoS Negl. Trop. Dis. in press.

IUCN 2022. The IUCN Red List of Threatened Species. Version 2022-2. in press.

Johnson, T. F. et al. 2021. Handling missing values in trait data. - Glob. Ecol. Biogeogr. 30: 51–62.
Jones, K. E. et al. 2008. Global trends in emerging infectious diseases. - Nature 451: 990–993.
Jones, K. E. et al. 2009. PanTHERIA: a species−level database of life history, ecology, and geography of extant and recently extinct mammals. - Ecology 90: 2648–2648.
Laughlin, A. J. et al. 2019. Ecological determinants of pathogen transmission in communally roosting species. - Theor. Ecol. 12: 225–235.
Luis, A. D. et al. 2013. A comparison of bats and rodents as reservoirs of zoonotic viruses: are bats special? - Proc. Biol. Sci. 280: 20122753.
Lunn, T. J. et al. 2021. Spatial dynamics of pathogen transmission in communally roosting species: Impacts of changing habitats on bat-virus dynamics. - J. Anim. Ecol. 90: 2609–2622.
McKee, C. D. et al. 2019. Host Phylogeny, Geographic Overlap, and Roost Sharing Shape Parasite Communities in European Bats. - Frontiers in Ecology and Evolution in press.
Mull, N. et al. 2022. Virus isolation data improve host predictions for New World rodent orthohantaviruses. - J. Anim. Ecol. 91: 1290–1302.
Murray, M. H. et al. 2019. City sicker? A meta-analysis of wildlife health and urbanization. - Front. Ecol. Environ. 17: 575–583.
Ntumvi, N. F. et al. 2021. Evaluation of bat adenoviruses suggests co-evolution and host roosting behaviour as drivers for diversity. - Microb Genom in press.
Olival, K. J. et al. 2017. Host and viral traits predict zoonotic spillover from mammals. - Nature in press.
Ostfeld, R. S. et al. 2014. Life history and demographic drivers of reservoir competence for three tick-borne zoonotic pathogens. - PLoS One 9: e107387.
Passos, E. C. et al. 1998. Isolation of rabies virus in an insectivorous bat Nyctinomops macrotis, in Southeastern Brazil. - Rev. Saúde Pública 32: 74–76.
Plowright, R. K. et al. 2011. Urban habituation, ecological connectivity and epidemic dampening: the emergence of Hendra virus from flying foxes (Pteropus spp.). - Proc. Biol. Sci. 278: 3703–3712.
Plowright, R. K. et al. 2019a. Prioritizing surveillance of Nipah virus in India. - PLoS Negl. Trop. Dis. 13: e0007393.
Plowright, R. K. et al. 2019b. Sampling to elucidate the dynamics of infections in reservoir hosts. - Philos. Trans. R. Soc. Lond. B Biol. Sci. 374: 20180336.
Santini, L. et al. 2019. One strategy does not fit all: determinants of urban adaptation in mammals. - Ecol. Lett. 22: 365–376.
Shapiro, J. T. et al. 2021. Setting the Terms for Zoonotic Diseases: Effective Communication for Research, Conservation, and Public Policy. - Viruses in press.
Shaw, L. P. et al. 2020. The phylogenetic range of bacterial and viral pathogens of vertebrates. - Mol. Ecol. 29: 3361–3379.
Simmons, N. B. and Cirranello., A. L. 2023. Bat Species of the World: A taxonomic and geographic database. Version 1.3.
Smith, K. F. et al. 2014. Global rise in human infectious disease outbreaks. - J. R. Soc. Interface in press.
Soria, C. D. et al. 2021. COMBINE: a coalesced mammal database of intrinsic and extrinsic traits. - Ecology 102: e03344.
Streicker, D. G. et al. 2010. Host phylogeny constrains cross-species emergence and

establishment of rabies virus in bats. - Science 329: 676–679.
Szentivanyi, T. et al. 2023. Trends in Bacterial Pathogens of Bats: Global Distribution and Knowledge Gaps. - Transbound. Emerg. Dis. in press.
Tompkins, D. M. et al. 2015. Emerging infectious diseases of wildlife: a critical perspective. - Trends Parasitol. 31: 149–159.
Turmelle, A. S. and Olival, K. J. 2009. Correlates of Viral Richness in Bats (Order Chiroptera). - Ecohealth 6: 522.
Upham, N. S. et al. 2019. Inferring the mammal tree: Species-level sets of phylogenies for questions in ecology, evolution, and conservation. - PLoS Biol. 17: e3000494.
Van der Poel, W. H. M. et al. 2005. European bat lyssaviruses, The Netherlands. - Emerg. Infect. Dis. 11: 1854–1859.
Weber, N. et al. 2023. Robust evidence for bats as reservoir hosts is lacking in most African virus studies: a review and call to optimize sampling and conserve bats. - Biol. Lett. 19: 20230358.
Werner, C. S. and Nunn, C. L. 2020. Effect of urban habitat use on parasitism in mammals: a meta-analysis. - Proceedings of the Royal Society B: Biological Sciences in press.
Wilman, H. et al. 2014. EltonTraits 1.0: Species-level foraging attributes of the world's birds and mammals. - Ecology 95: 2027–2027.
Woolhouse, M. E. J. et al. 2008. Temporal trends in the discovery of human viruses. - Proceedings of the Royal Society B: Biological Sciences 275: 2111.