

# The coevolutionary mosaic of bat-betacoronaviruses spillover risk

Norma Forero Rocio Munoz<sup>1,2</sup> Renata L. Muylaert<sup>3</sup> Stephanie N. Seifert<sup>4</sup> Gregory F. Albery<sup>5</sup> Colin J. Carlson<sup>5</sup> [Timothée Poisot](#)<sup>1,2,‡</sup>

<sup>1</sup> Université de Montréal <sup>2</sup> Québec Centre for Biodiversity Sciences <sup>3</sup> Molecular Epidemiology and Public Health Laboratory, Hopkirk Research Institute, Massey University, Palmerston North, New Zealand

<sup>4</sup> Paul G. Allen School for Global Health, Washington State University, Pullman, WA, United States <sup>5</sup> ???

‡ These authors contributed equally to the work

## Correspondance to:

Timothée Poisot — [timothee.poisot@umontreal.ca](mailto:timothee.poisot@umontreal.ca)

This work is released by its authors under a CC-BY 4.0 license



Last revision: *April 28, 2022*

Coming soon

1 Spillover risk is not unidimensional. From the standpoint of an animal community, *i.e.* a pool of suitable  
2 hosts, it is driven by a multiplicity of factors (Plowright et al. 2017). The global richness of hosts is one  
3 such component commonly mentioned/analysed (see *e.g.* Anthony et al. 2017 for coronaviruses), but  
4 there is an argument to be made that species who are not competent (or know) hosts of a specific virus  
5 genus may not factor into this (Plowright et al. 2015), or that species who are assumed to share viruses at  
6 different rates should be weighted accordingly (Albery et al. 2020). In mammals, key functional traits (for  
7 which phylogeny is a reasonable proxy) are determinants of the spillover potential (Olival et al. 2017);  
8 these include, notably, body mass and affinity for urban environments (Albery et al. 2022). Finally,  
9 especially when the pool of potential hosts spans the entire globe, there may be local host pools that are  
10 highly unique; not having been observed in other locations, these can act on the overall risk either by  
11 providing novel contact opportunities, reflecting unique host-environment combinations (Engering et al.  
12 2013), or facilitating rapid evolutionary changes in specialism of their pathogens (Agosta et al. 2010). In  
13 the specific case of generalist pathogens, there is conceptual and empirical support to the idea that these  
14 community- level mechanisms are even more important in driving the overall risk (Power and Mitchell  
15 2004).

16 Bats are important reservoir hosts for different classes of microorganisms (Chu 2008, Donaldson 2010, Li  
17 2010), some of which can threaten human health. Chiropterans emerged around 64 million years ago and  
18 are one of the most diverse mammalian orders, with an estimated richness of more than 12000 species,  
19 (Peixoto et al. 2018) and 14325 known species (Simmons and Cirranello 2020). They exhibit a broad  
20 variety of habitat use, behaviour, and feeding strategies, resulting in their playing an essential role in the  
21 delivery of several ecosystem services tied to important ecosystem-derived benefits (Kasso and  
22 Balakrishnan 2013). For example, over two-thirds of bats are known to be either obligate or facultative  
23 insectivorous mammals, therefore playing an important role in the regulation of insect pests that can  
24 affect crops (Williams-Guillén et al. 2008, Voigt and Kingston 2016), and vectors of diseases that put a risk  
25 on human health (Gonsalves et al. 2013a, b). Because bats are globally distributed and have a long  
26 evolutionary history, phylogeographic and biogeographic approaches are required to shed light on the  
27 extant distribution of coevolutionary processes between bats and the pathogens they carry. Not all areas in  
28 which bats, viruses, and human are co-occurring are facing a risk of spillover towards human populations,  
29 and the areas in which this risk exist may not be facing risks of the same nature and magnitude.

30 In this paper, we examine the biogeographic structure of bats-betacoronaviruses associations, based on a

31 curated dataset of known and recently discovered hosts. This work is important both as a description of  
32 the bats-betacoronavirus complex, but also because more broadly, bats are known reservoirs for a variety  
33 of emerging viruses and pathogens (Calisher et al. 2006, Melaun et al. 2014), making balancing the needs  
34 for bat conservation and disease prevention a potentially difficult act and a source of human-wildlife  
35 conflicts, especially in more densely populated areas (Stone et al. 2015, Rego et al. 2015). By drawing on  
36 concepts from the Geographic Mosaic Theory of Coevolution (Thompson 2005), we turn these  
37 associations into a spatially explicit additive mapping of zoonotic risk components, which reveals extreme  
38 heterogeneity of risk at the global scale; furthermore, we identify the Amazon and South-Eastern Asia as  
39 hotspots of phylogenetic distinctiveness of betacoronaviruses (Anthony et al. 2017); surprisingly, current  
40 data suggest that viral sharing between hosts is high in the Amazon and low in South-Eastern Asia, which  
41 has the potential to result in different evolutionary dynamics between these two regions.

## 42 **Methods**

### 43 **Known betacoronavirus hosts**

44 We downloaded the data on bats hosts of betacoronaviruses assembled by Becker et al. (2022) from  
45 <https://www.viralemergence.org/betacov> on Apr. 2022, and filtered it to “known” hosts (established  
46 before the emergence of SARS-CoV-2) and “novel” hosts (confirmed through sampling since the  
47 emergence of SARS-CoV-2). The original database was assembled by a combination of data mining and  
48 literature surveys, including automated alerts on the “bats” and “coronavirus” keywords to identify novel  
49 empirical evidence of bats-betacoronaviruses associations.

### 50 **Bats occurrences**

51 We downloaded the rangemap of every extant bat species that was either classified as an empirically  
52 documented host of beta-coronaviruses from the previous step, according to recent IUCN data (IUCN  
53 2021). The range maps were subsequently rasterized using the rasterize function from GDAL (Rouault et  
54 al. 2022) at a resolution of approximately **TK TP**. For every pixel in the resulting raster where at least one  
55 bat host of betacoronavirus was present, we extract the species pool (list of all bat species), which was used  
56 to calculate the following risk assessment components: phylogenetic diversity, bat compositional

57 uniqueness, and predicted viral sharing risk.

## 58 **Bats phylogeography**

59 For every pixel, we measured Faith's Phylogenetic Diversity (Faith 1992) based on a recent synthetic tree  
60 with robust time calibration, covering about 6000 mammalian species (Upham et al. 2019). Faith's PD  
61 measures the sum of unique branches from an arbitrary root to a set of tips, and comparatively larger  
62 values indicate a more phylogenetic diverse species pool. We measured phylogenetic diversity starting  
63 from the root of the entire tree (and not from Chiroptera); this bears no consequences on the resulting  
64 values, since all branches leading up to Chiroptera are only counted one per species pool, and (as we  
65 explain when describing the assembly of the composite risk map), all individual risk components are  
66 ranged in  $[0,1]$ . This measure incorporates a richness component, which we chose not to correct for; the  
67 interpretation of the phylogenetic diversity is therefore a weighted species richness, that accounts for  
68 phylogenetic over/under-dispersal in some places.

## 69 **Bats compositional uniqueness**

70 For every species pool, we measured its Local Contribution to Beta-Diversity (Legendre and De Cáceres  
71 2013); LCBD works from a species-data matrix (traditionally noted as  $\mathbf{Y}$ ), where species are rows and sites  
72 are columns, and a value of 1 indicates occurrence. We extracted the  $\mathbf{Y}$  matrix assuming that every pixel  
73 represents a unique location, and following best practices (Legendre and Condit 2019) transformed it  
74 using Hellinger's distance to account for unequal bat richness at different pixels. The correction of raw  
75 community data is particularly important for two reasons: first, it prevents the artifact of richer sites  
76 having higher importance; second, it removes the effect of overall species richness, which is already  
77 incorporated in the phylogenetic diversity component. High values of LCBD indicate that the pixel has a  
78 community that is on average more dissimilar in species composition than what is expected knowing the  
79 entire matrix, i.e. a more unique community. Recent results by Dansereau et al. (2022) shows that LCBD  
80 measures are robust with regards to spatial scale, and are therefore applicable at the global scale.

## 81 **Viral sharing between hosts**

82 For all bat hosts of betacoronaviruses, we extracted their predicted viral sharing network (Albery et al.  
83 2020). This network stores pairwise values of viral community similarity. To project viral sharing values  
84 into a single value for every pixel, we averaged the pairwise scores. High values of the average sharing  
85 propensity means that this specific extant bat assemblage is likely to be proficient at exchanging viruses.

## 86 **Composite risk map**

87 To visualize the aggregated risk at the global scale, we combine the three individual risk components  
88 (phylogenetic diversity, compositional uniqueness, and viral sharing) using an additive color model  
89 (Seekell et al. 2018). In this approach, every risk component gets assigned a component in the RGB color  
90 model (phylogenetic diversity is green, compositional uniqueness is red, and viral sharing is blue). In  
91 order to achieve a valid RGB measure, all components are re-scaled to the [0,1] interval, so that a pixel  
92 with no sharing, no phylogenetic diversity, and no compositional uniqueness is black, and a pixel with  
93 maximal values for each is white. This additive model conveys both the intensity of the overall risk, but  
94 also the nature of the risk as colors diverge towards combinations of values for three risk components. Out  
95 of the possible combinations, the most risky in terms of rapid diversification and spillover potential is high  
96 phylogenetic diversity and low viral sharing (Gomulkiewicz et al. 2000, Cavender-Bares et al. 2009), in  
97 that this allows multiple independent host-virus coevolutionary dynamics to take place in the same  
98 location. In the colorimetric space, this correspond to yellow – because the HSV space is more amenable  
99 to calculations for feature extraction (see *e.g.* Keke et al. 2010), we measured the risk level by calculating  
100 the angular distance of the hue of each pixel to a reference value of 60, and weighted this risk level by the  
101 value component. Specifically, given a pixel with colorimetric coordinates  $(h, s, v)$ , its ranged weighted  
102 risk value is

$$v \times \left[ 1 - \frac{|\text{atan}(\cos(\text{rad}(h)), \sin(\text{rad}(h))) - X|}{2\pi} \right],$$

103 where  $X$  is  $\text{atan}(\cos(\text{rad}(60)), \sin(\text{rad}(60)))$ , a constant approximately equal to 0.5235.

## **Viral phylogeography and evolutionary diversification**

We used the following query to pull all betacoronavirus sequence data from the GenBank Nucleotide database except SARS-CoV-2; (“Betacoronavirus”[Organism] OR betacoronavirus[All Fields]) NOT (“Severe acute respiratory syndrome coronavirus 2”[Organism] OR sars-cov-2[All Fields]). We added a single representative sequence for SARS-CoV-2 and manually curated to remove sequences without the RNA-dependent RNA polymerase (RdRp) sequence or that contained words indicating recombinant or laboratory strains including “patent,” “mutant,” “GFP,” and “recombinant.” We filtered over-represented taxa including betacoronavirus 1, hCoV-OC43, Middle East respiratory syndrome coronavirus, Murine hepatitis virus, and hCoV-HKU1. Curated betacoronavirus RdRp sequences were then aligned using MAFFT v 1.4.0 (Kato and Standley 2013, Supplemental X) and a maximum likelihood tree reconstructed in IQ-TREE v 1.6.12 (Nguyen et al. 2015) with ModelFinder (Kalyaanamoorthy et al. 2017) ultrafast bootstrap approximation (Hoang et al. 2018) and the following parameters (STEPH WILL ADD, Supplemental X).

We first tested the hypothesis that hotspots of viral diversification would track hotspots of bat diversification. To do so, we plotted the number of known bat hosts (specifically only those included in the phylogeny, so there was a 1:1 correspondence between data sources) against the “mean evolutionary distinctiveness” of the associated viruses. To calculate this, we derived the fair proportions evolutionary distinctiveness (Isaac et al. 2007) for each of the viruses in the tree, then averaged these at the bat species level, projected these values onto their geographic distributions, and averaged across every bat found in a given pixel. As such, this can be thought of as a map of the mean evolutionary distinctiveness of the known viral community believed to be associated with a particular subset of bats present.

## **Co-distribution of hosts and viral hotspots**

Subsequently, we tested the hypothesis that the biogeography of bat betacoronaviruses should track the biogeography of their hosts. To test this idea, we loosely adapted a method from (Kreft and Jetz 2007, 2010), who proposed a phylogenetic method for the delineation of animal biogeographic regions. In their original method, a distance matrix - where each row or column represents a geographic raster’s grid cell, and the dissimilarity values are the “beta diversity similarity” of their community assemble - undergoes non-metric multidimensional scaling (NMDS); the first two axes of the NMDS are projected

geographically using a four-color bivariate map. Here, we build on this idea with an entirely novel methodology. First, we measure the phylogenetic distance between the different viruses in the betacoronavirus tree by using the cophenetic function in ape (Paradis and Schliep 2019); subsequently, we take a principal components analysis of that distance matrix (readily interchangeable for NMDS in this case) to project the viral tree into an n-dimensional space. We then take the first two principal components and, as with the evolutionary distinctiveness analysis, aggregated these to a mean host value and projected them using a four-color bivariate map.

### **Outbreaks data geo-referencing**

Finally, we provide a summary visualization of what available information describes the spillover of zoonotic betacoronaviruses of bat origin where data was available before and up through the COVID-19 pandemic. The SARS-CoV-2 outbreak was georeferenced to the initial case cluster in Wuhan, China; SARS-CoV was georeferenced based on the cave with the closest known viruses circulating in nature (Hu et al. 2017 PLoS Pathogens), and a nearby location where serological (antibody) evidence has indicated human exposure to SARS-like viruses (Wang et al. 2018 Virologica Sinica). For MERS-CoV, we presented the index cases available from a recently-published compendium of MERS-CoV cases (Ramshaw et al. 2019); these are largely if not all presumed to be camel-to-human transmission, and the precise origin point of MERS-CoV in bats is uncertain. Not shown is a recent case of a recombinant canine coronavirus that showed the ability to infect humans, both because this study was published after the beginning of the COVID-19 pandemic and because bats' involvement in this cycle of transmission has been marginal to non-existent.

## **Results**

### **Host distribution**

Chiroptera are an hyperdiverse group, distributed in a large part of the world, and are an important reservoir for different strains of betacoronaviruses (Drexler et al. 2014); this has attracted attention to areas where high diversity of bats can be an important issue for human health (Calisher et al. 2006, Moratelli and Calisher 2015). Accordingly, we collected the IUCN rangemaps for known hosts of betacoronaviruses,



158 to illustrate where hotspots of host diversity are. These results are presented in Fig xx.a. As per our  
159 current knowledge of which bats are hosts of betacoronaviruses, these hotspots are primarily South-East  
160 Asia, parts of Europe, and to a lesser extent sub-saharan Africa. Even the subset of chiroptera that are  
161 hosts of betacoronaviruses fits the evolutionary timeline of the group. Chiropterans can be classified as  
162 Microchiroptera and macrochiroptera, where macrochiroptera have an older history from an evolutionary  
163 perspective compared to macrochiroptera (Teeling et al. 2005, Springer 2013). South-East Asia has a high  
164 diversity of bats (Kingston, 2010), and our results show that part of that diversity includes betacoronavirus  
165 hosts. High density of hosts sharing the same virus (albeit possibly different strains) calls into question the  
166 evolution of the bat antiviral immune system and its co-evolution with viruses, which may result in  
167 distinct immunological responses in different area, as evidenced in other bat species (Banerjee et al. 2020).

## 168 **Viral evolutionary distinctiveness**

169 Higher host diversity may not result in a higher viral diversity; for this reason, we quantified and mapped  
170 the evolutionary distinctiveness of betacoronaviruses, based on .... Viral evolutionary distinctiveness  
171 largely tracks host diversity, particularly in southern China but oddly not throughout the rest of southeast  
172 Asia, perhaps indicating that many distinctive viruses remain to be discovered in this region (an idea that  
173 is unsurprising given the growing realization, around the emergence of SARS-CoV-2, that a unique  
174 lineage of similar viruses are widespread in bats but still mostly undescribed). The most distinct  
175 betacoronaviruses are found in South America, a region with a comparatively lower number of hosts; this  
176 suggests that the South American bat-betacoronavirus complex has been more isolated, and is probably  
177 undergoing a different co-evolutionary dynamic. Alternatively, this distinctiveness hotspot may be a  
178 product of under-sampling: South-America is one of the places where the fewest betacoronaviruses have  
179 been discovered (Anthony et al. 2017), and adding more viruses would bring the distinctiveness of known  
180 sequences down. Previous work has suggested the Americas may be a hotspot of both undiscovered bat  
181 viruses in general (Olival et al. 2017, Allen et al. 2017) and coronavirus specifically (Anthony et al. 2017),  
182 though not necessarily betacoronaviruses, and particularly not those in clades with notable zoonotic  
183 potential.

## 184 **Geographic Mosaic of bat-betacoronavirus risk**

185 In order to turn the hypotheses based on the Geographic Mosaic Theory of Coevolution into a measure of  
186 risk, we overlapped three components: viral sharing, i.e. the chance that two bats will share viruses  
187 overall; Local Contribution to Beta Diversity, i.e. the fact that a bat community is compositionally unique  
188 compared to the average compositional similarity across the entire system; finally, the phylogenetic  
189 diversity, i.e. how dispersed the bats in a location are within the tree of life. These results are presented  
190 using an additive color mapping in Figure xx, and lead to the definition of broad biogeographic regions of  
191 risk, where the same color represents the same type of risk. Pairwise maps of the three components are  
192 present in supplementary materials.

193 From the perspective of spillover risk, the most important combination of factors is a high phylogenetic  
194 diversity of hosts with low viral sharing; this, essentially, means that very different betacoronavirus could  
195 co-exist within the same place. This is particularly the case given that betacoronaviruses often evolve and  
196 even achieve host shifts through recombination, which requires the co-occurrence of sufficiently distinct  
197 viruses to be a major driver of emergence. In Fig. xx, this corresponds to yellow to pale green areas, which  
198 are essentially limited to South-Eastern Asia, and to some part of Sub-Saharan Africa. Adopting a  
199 geographic mosaic theory perspective on risk, other regions of the world are of lesser concern.

200 Available data on bat betacoronavirus spillover into humans (TP overlay on the figure) is limited and  
201 circumstantial at best for these purposes, but our risk maps suggest that the areas predicted by prior  
202 expectations about host biogeography correspond loosely to those where previous emergence events have  
203 been recorded. Areas with high bat diversity and high turnover may facilitate the evolutionary radiation of  
204 viruses, matching previous findings that the diversification of bat coronaviruses is driven largely by host  
205 shifts (inter-genus or higher levels of cross-species transmission) and, to a lesser degree, cospeciation and  
206 sharing (intra-genus cross-species transmission; Anthony et al. 2017). This diversification - while not an  
207 actual risk factor for spillover itself - likely increases the random chance of a virus with the raw genomic  
208 components required for the potential to infect humans.

## 209 **Global distribution of spillover risk**

210 Based on the previous result, we extracted the yellow component of the risk map (TP add methods), to  
211 provide a single measure of risk varying between 0 and 1. This measure is presented in Fig. xxA. However,

this maps the potential risk, which must be weighed by the potential for contacts with humans. As a proxy for this measure, we used the proportion of build/urban land from the EarthEnv dataset: this is a reasonable proxy for the density of humans per unit area, which increases the probability of pathogen spread more widely (Hazarie et al., 2021). Since human activity is required to amplify the frequency of virus encounters and thus create areas of viral amplification, mapping the potential risk against measures of land use is required to generate a more actionable assessment of risk. This map is presented in Fig. xxB. Most of South America and Europe are at low risk, as although densely populated, settlements tend to be in areas with lower potential risk. However, this mapping reveals that South-East Asia, the Indian subcontinent, and parts of sub-Saharan Africa, are at high risk due to the overlap between built areas and bat communities representing more opportunities for cross-species transmission of betacoronaviruses.

## Discussion

Driven by the need to understand the ecological factors involved in the emergence of viral pathogens, we spatially mapped bat-betacoronavirus interactions worldwide, using (i) a database of known betacov hosts (Becker et al., 2020), and (ii) range maps for the hosts according to IUCN (IUCN 2021). To reflect the fact that the risk posed by viruses has many ecological origins, we quantified the phylogenetic diversity of hosts, their compositional uniqueness, and the expected viral sharing. Because these components of risk matter when contrasted to human density, we compared them to a proxy, namely the proportion of each pixel that is covered by urban or built land. This provides a synthetic risk map, allowing to identifying of hotspots where the bat-betacoronavirus system may originate viruses in humans. SE Asia is one of the regions with the highest risk since, according to our results, several of its conditions could increase the risk of transmission of the virus.

Bats are found worldwide and are one of the most diverse groups among mammals (Moratelli & Calisher, 2015). Previous research (Anthony et al., 2017; Mollentze & Streicker, 2020) states that locally diverse bat communities could maintain more viruses and hence, a higher probability of having a pathogen that could represent a risk for human health. This probability involves multiple factors, among which the relatedness of hosts (which can make the jumps easier (Longdon et al., 2011; Mollentze et al., 2020; Wolfe et al., 2007), and the overall tendency of hosts within a locality to share viruses, which may limit viral diversity because of within-host competition (Leeks et al., 2018; Sallinen et al., 2020). Species richness, therefore, is not a

240 sufficient measure of viral risk. This is exemplified in our results, where both South America and  
241 South-Eastern Asia have a high species richness of betacov hosts, but only the latter region has a high risk.  
242 Specifically, because previous studies propose that Asia is important when it comes to understanding the  
243 evolutionary origin of various mammalian taxa (Beard C K, 1988). Including bats (Yu et al., 2014), which  
244 could support the relationship between evolutionary time and the development of an immune system  
245 with characteristics that allow them to be better adapted to infection by emerging viruses (Gorbunova et  
246 al., 2020; Irving et al., 2021) may be related to a wide variety of diets (Jones et al., 2022; Moreno Santillán  
247 et al., 2021; Banerjee et al., 2020; Schneeberger et al., 2013).

248 Our study focuses largely on the biogeography of hosts. Yet, we know that viruses with high host  
249 plasticity, that is, the ability of a given virus to adapt to various taxonomic orders and ecological groups  
250 (Kreuder Johnson et al., 2015); are more likely to amplify viral spillover, followed by secondary  
251 human-to-human transmission, and geographical spread (Hazarie et al., 2021). High viral host plasticity is  
252 an especially important trait for RNA viruses such as betacov (Kreuder Johnson et al., 2015; Haddad et al.,  
253 2021). Indeed, our analysis of viral sequences reveals that Latin America is a hotspot of viral  
254 distinctiveness, suggesting that this part of the bats-betacov system may be undergoing independent  
255 evolutionary dynamics (related species sharing viruses that are different from the rest of the global pool).  
256 The other hotspot of viral distinctiveness is S.E. Asia, in which richness is high but sharing is low; this  
257 suggests a different type of evolutionary dynamics (unrelated viruses coevolving with evolutionarily  
258 distinct hosts, generating high diversity locally).

259 This diversity of hosts and how the exchange of viruses occurs between species, is largely affected by the  
260 different environmental changes, as the case of sarbecovirus bats reservoirs (Muylaert et al., 2021) where  
261 they are affected by the area of the cave or the alteration of the forest, which could result in modifications  
262 of host distribution. Additionally, our results highlight the importance of Asia as a betacov hotspot, which  
263 is consistent with recent studies (Muylaert et al., 2021), where projections on this area suggest that new  
264 future events of sarbecovirus viral exchange might be easily spread among species or humans.

265 There are several factors that drive changes in the diversity of bats (Alves et al., 2018), but human  
266 activities' effects on the ecosystem (like modifications of land use) could significantly decrease it.  
267 Therefore, it can be suggested that changes in the diversity of betacovs in bats are linked to their  
268 biogeographic variation, and human population density and other anthropogenic factors are decisive  
269 moderators for its implications in public health. With the increase of contact between humans and

270 potential hosts, we also increase the risk of emergence of novel diseases (Johnson et al., 2020), as previous  
271 studies on RNA viruses suggest the importance of host phylogeography at the time of virus dispersal  
272 (Gryseels et al., 2017).

273 One of these scenarios where interaction between bats and humans can occur can be seed dispersal in  
274 tropical agroecosystems. It opens the discussion of whether the fruits thrown by bats not only disperse  
275 seeds but could also be a source of indirect interaction between viruses of bat origin and humans  
276 (Deshpande et al., 2022) . This represents a challenge for conservation strategies and disease ecology since  
277 we have areas with potential zoonotic viruses and bat-human interaction. However, it must still be taken  
278 into account the quantification of real exposure from several scenarios, where there can be directly or  
279 indirectly bat - human interaction.

280 Comparing scenarios of high viral exchange vs low viral exchange, open the discussion to consider if the  
281 best scenario is where viruses easily adapted to multiple hosts but with low virulence or easily ignored by  
282 the immune system of the host, or where we have viruses specialized to a specific host, but highly virulent  
283 when invade a new host. Accordingly, the understanding of viral-host interactions from a taxonomic and  
284 phylogenetic contributes to improving zoonoses surveillance programs.

285 **Acknowledgements:** We acknowledge that this study was conducted on land within the traditional  
286 unceded territory of the Saint Lawrence Iroquoian, Anishinabewaki, Mohawk, Huron-Wendat, and  
287 Omàmiwininiwak nations. This work was supported by funding to the Viral Emergence Research  
288 Initiative (VERENA) consortium including NSF BII 2021909 and a grant from Institut de Valorisation des  
289 Données (IVADO). This research was enabled in part by support provided by Calcul Québec  
290 (www.calculquebec.ca) and Compute Canada (www.computecanada.ca). NF is funded by the NSERC  
291 BIOS<sup>2</sup> CREATE program. TP and NF are funded by the Courtois Foundation. RLM was supported by  
292 Bryce Carmine and Anne Carmine (née Percival), through the Massey University Foundation.

## 293 **References**

- 294 Agosta, S. J. et al. 2010. How specialists can be generalists: Resolving the “parasite paradox” and  
295 implications for emerging infectious disease. - *Zoologia (Curitiba)* 27: 151–162.
- 296 Albery, G. F. et al. 2020. Predicting the global mammalian viral sharing network using phylogeography. -

297 Nature Communications 11: 2260.

298 Albery, G. F. et al. 2022. Urban-adapted mammal species have more known pathogens. in press.

299 Allen, T. et al. 2017. Global hotspots and correlates of emerging zoonotic diseases. - Nature  
300 Communications in press.

301 Anthony, S. J. et al. 2017. Global patterns in coronavirus diversity. - Virus Evolution in press.

302 Banerjee, A. et al. 2020. Novel Insights Into Immune Systems of Bats. - Frontiers in Immunology 11: 26.

303 Becker, D. J. et al. 2022. Optimising predictive models to prioritise viral discovery in zoonotic reservoirs. -  
304 The Lancet Microbe in press.

305 Calisher, C. H. et al. 2006. Bats: Important Reservoir Hosts of Emerging Viruses. - Clinical Microbiology  
306 Reviews 19: 531–545.

307 Cavender-Bares, J. et al. 2009. The merging of community ecology and phylogenetic biology. - Ecol. Lett.  
308 12: 693–715.

309 Dansereau, G. et al. 2022. Evaluating ecological uniqueness over broad spatial extents using species  
310 distribution modelling. - Oikos n/a: e09063.

311 Drexler, J. F. et al. 2014. Ecology, evolution and classification of bat coronaviruses in the aftermath of  
312 SARS. - Antiviral Research 101: 45–56.

313 Engering, A. et al. 2013. Pathogen–host–environment interplay and disease emergence. - Emerging  
314 Microbes & Infections 2: e5.

315 Faith, D. P. 1992. Conservation evaluation and phylogenetic diversity. - Biological Conservation 61: 1–10.

316 Gomulkiewicz, R. et al. 2000. Hot Spots, Cold Spots, and the Geographic Mosaic Theory of Coevolution. -  
317 The American Naturalist 156: 156–174.

318 Gonsalves, L. et al. 2013a. Do mosquitoes influence bat activity in coastal habitats? - Wildlife Research 40:  
319 10–24.

320 Gonsalves, L. et al. 2013b. Mosquito Consumption by Insectivorous Bats: Does Size Matter? - PLOS ONE  
321 8: e77183.

322 Isaac, N. J. B. et al. 2007. Mammals on the EDGE: Conservation Priorities Based on Threat and Phylogeny.  
323 - PLOS ONE 2: e296.

324 IUCN 2021. The IUCN Red List of Threatened Species.

325 Kasso, M. and Balakrishnan, M. 2013. Ecological and Economic Importance of Bats (Order Chiroptera). -  
 326 ISRN Biodiversity 2013: e187415.

327 Keke, S. et al. 2010. Study on skin color image segmentation used by Fuzzy-c-means arithmetic. - 2010  
 328 Seventh International Conference on Fuzzy Systems and Knowledge Discovery 2: 612–615.

329 Kreft, H. and Jetz, W. 2007. Global patterns and determinants of vascular plant diversity. - Proceedings of  
 330 the National Academy of Sciences 104: 5925–5930.

331 Kreft, H. and Jetz, W. 2010. A framework for delineating biogeographical regions based on species  
 332 distributions. - Journal of Biogeography 37: 2029–2053.

333 Legendre, P. and De Cáceres, M. 2013. Beta diversity as the variance of community data: Dissimilarity  
 334 coefficients and partitioning (H Morlon, Ed.). - Ecology Letters 16: 951–963.

335 Legendre, P. and Condit, R. 2019. Spatial and temporal analysis of beta diversity in the Barro Colorado  
 336 Island forest dynamics plot, Panama. - Forest Ecosystems 6: 7.

337 Melaun, C. et al. 2014. Bats as Potential Reservoir Hosts for Vector-Borne Diseases. - In: Klimpel, S. and  
 338 Mehlhorn, H. (eds), Bats (Chiroptera) as Vectors of Diseases and Parasites: Facts and Myths.  
 339 Parasitology Research Monographs. Springer, pp. 25–61.

340 Moratelli, R. and Calisher, C. H. 2015. Bats and zoonotic viruses: Can we confidently link bats with  
 341 emerging deadly viruses? - Memórias do Instituto Oswaldo Cruz 110: 1–22.

342 Olival, K. J. et al. 2017. Host and viral traits predict zoonotic spillover from mammals. - Nature 546:  
 343 646–650.

344 Paradis, E. and Schliep, K. 2019. Ape 5.0: An environment for modern phylogenetics and evolutionary  
 345 analyses in R. - Bioinformatics 35: 526–528.

346 Peixoto, F. P. et al. 2018. A synthesis of ecological and evolutionary determinants of bat diversity across  
 347 spatial scales. - BMC ecology 18: 18.

348 Plowright, R. K. et al. 2015. Ecological dynamics of emerging bat virus spillover. - Proceedings of the  
 349 Royal Society B: Biological Sciences 282: 20142124.

350 Plowright, R. K. et al. 2017. Pathways to zoonotic spillover. - Nature Reviews Microbiology 15: 502–510.

351 Power, A. G. and Mitchell, C. E. 2004. Pathogen Spillover in Disease Epidemics. - The American  
 352 Naturalist 164: S79–S89.

353 Rego, K. M. da C. et al. 2015. Assessing human-bat interactions around a protected area in northeastern  
 354 Brazil. - Journal of Ethnobiology and Ethnomedicine 11: 80.

355 Rouault, E. et al. 2022. GDAL/OGR Geospatial Data Abstraction software Library. - Zenodo.

356 Seekell, D. A. et al. 2018. A geography of lake carbon cycling. - Limnology and Oceanography Letters 3:  
 357 49–56.

358 Simmons, N. B. and Cirranello, A. L. 2020. Bat Species of the World: A taxonomic and geographic  
 359 database.

360 Springer, M. S. 2013. Phylogenetics: Bats United, Microbats Divided. - Current Biology 23: R999–R1001.

361 Stone, E. et al. 2015. Managing Conflict between Bats and Humans: The Response of Soprano Pipistrelles  
 362 (*Pipistrellus pygmaeus*) to Exclusion from Roosts in Houses. - PLoS ONE 10: e0131825.

363 Teeling, E. C. et al. 2005. A Molecular Phylogeny for Bats Illuminates Biogeography and the Fossil Record.  
 364 - Science (New York, N.Y.) 307: 580–584.

365 Thompson, J. N. 2005. The Geographic Mosaic of Coevolution. - University Of Chicago Press.

366 Upham, N. S. et al. 2019. Inferring the mammal tree: Species-level sets of phylogenies for questions in  
 367 ecology, evolution, and conservation. - PLOS Biology 17: e3000494.

368 2016. Bats in the Anthropocene: Conservation of Bats in a Changing World (CC Voigt and T Kingston,  
 369 Eds.). - Springer International Publishing.

370 Williams-Guillén, K. et al. 2008. Bats Limit Insects in a Neotropical Agroforestry System. - Science 320:  
 371 70–70.