

# The coevolutionary mosaic of betacoronavirus emergence 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>

Daniel J. Becker<sup>6</sup> Colin J. Carlson<sup>7,8,9,‡</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, School of Veterinary Science, Massey University, New Zealand <sup>4</sup> Paul G. Allen School for Global Health, Washington State University, Pullman, WA, United States <sup>5</sup> Department of Biology, Georgetown University, Washington, DC, USA <sup>6</sup> Department of Biology, University of Oklahoma, Norman, OK, USA <sup>7</sup> Department of Biology, Georgetown University, Washington, DC,

<sup>8</sup> Center for Global Health Science and Security, Georgetown University Medical Center, Washington, DC, USA <sup>9</sup> Department of Microbiology and Immunology, Georgetown University Medical Center, Washington, DC, USA

<sup>‡</sup> These authors contributed equally to the work

## Correspondance to:

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

Driven by the need to understand the ecological factors involved in the emergence of *Betacoronavirus* (the genus causing the SARS, MERS, and COVID-19 in human) through bat hosts, we develop an approach to the assessment of spillover risk based on the Geographic Mosaic Theory of Coevolution. In doing so, we provide a global mapping of the spillover risk posed by betacoronaviruses, reflecting the fact that this risk is best understood through the multi-faceted prism of ecological and evolutionary mechanisms. Our framework reveals that reservoir richness alone, although a component of viral hazard, is not a sufficiently integrative predictor of risk. We offer alternative insights based on viral sharing, host compositional uniqueness, and host phylogenetic diversity and phylogeographic regions. By comparing our aggregated measure of risk to a proxy for human density, namely the proportion of each pixel that is covered by urban or built land, we provide a synthetic risk map, allowing the identification of hotspots where the bat-betacoronavirus system may originate novel viruses in humans.

1 Disease emergence is complex, and is driven not only by animal-human contact, but also by the  
2 underlying evolutionary dynamics in viral reservoirs.<sup>1</sup> Although host richness is often used as a superficial  
3 proxy for spillover risk,<sup>2,3</sup> these approaches oversimplify the relevant interspecific heterogeneity in  
4 immunology, behavior, and other traits, and therefore overlook unique host pools that allow for the rapid  
5 evolution of highly divergent viruses.<sup>4</sup> In the case of generalist pathogens like betacoronaviruses, there is  
6 conceptual and empirical support to the idea that these community-level mechanisms are even more  
7 important,<sup>5</sup> particularly given that cross-species transmission may, as a rule, structure viral evolution  
8 more than co-divergence with hosts [Geoghegan]. This creates a disconnect between coevolutionary  
9 theory (including empirical evidence from virology) and most existing ecological frameworks for mapping  
10 spillover risk.

11 The Geographic Mosaic Theory of Coevolution<sup>GMTC;</sup><sup>6</sup> attempts to explicitly connect microevolutionary  
12 dynamics to the macroecology and biogeography of symbiotic interactions. The GMTC posits that  
13 coevolutionary processes among pairs<sup>7</sup> or complexes<sup>8</sup> of species are structured in space by the rippling  
14 effects of abiotic conditions onto evolutionary mechanism, giving rise to fragmented systems with  
15 different structure and ecologically dynamics over large spatial extents.<sup>see e.g.</sup><sup>9</sup> The GMTC predicts a  
16 spatial fragmentation of coevolutionary dynamics under the joint action of three processes:<sup>see notably</sup><sup>10</sup>  
17 coevolutionary hot- and coldspots, which appear when the intensity of *interaction* (in terms of reciprocal  
18 fitness consequences) varies spatially; selection mosaics, wherein the intensity of *selection* varies across  
19 space, driven by both the biotic complexity of the community (locally diverse hosts and viruses are more  
20 biotically complex) and the local favorability of the environment;<sup>11</sup> and trait remixing, which occurs when  
21 coevolutionary dynamics are driven by the arrival (or departure) of *functional traits*, through changes  
22 in community composition due to invasions, meta-community dynamics, and dispersal.

23 Here, we apply the GMTC to explore and explain the global biogeography of betacoronaviruses, the group  
24 that includes SARS-CoV, MERS-CoV, and SARS-CoV-2. In their bat reservoirs, coronaviruses evolve  
25 through a mix of host jumps, recombination among disparate lineages, and, to a lesser degree,  
26 co-divergence with their hosts—<sup>2</sup>a mix of mechanisms that creates a complex and nonlinear relationship  
27 between host diversity and viral emergence. Working from a recently published database of bat hosts of  
28 betacoronaviruses, we develop the first global maps of both host and virus evolutionary distinctiveness  
29 and biogeographic regions for this system. Aiming to explain these patterns, we develop a generalized  
30 framework for applying the GMTC to host-virus interactions, with a specific emphasis on the potential to

31 create independent coevolutionary dynamics (and therefore spatial fragmentation in risk) through  
32 heterogeneity. We develop a trivariate risk assessment system that connects each GMTC mechanism to a  
33 quantifiable aspect of host-virus interactions: (i) viral sharing rates in host communities, representing the  
34 strength of potential interaction between viruses and any one host (i.e., places where viruses undergo  
35 constant host switching may be coevolutionary coldspots); (ii) the phylogenetic diversity of hosts, as a  
36 proxy for variation in the immunological mechanisms that antagonize viruses (i.e., the selection mosaic);  
37 and (iii) the local uniqueness of the bat community, representing the potential for viruses to be exposed to  
38 novel host traits (e.g., variation in receptor sequences). Together, we argue that these can be used to  
39 identify and map the evolutionary drivers that—in conjunction with transmission processes (e.g., viral  
40 prevalence in reservoirs and animal-human contact rates)—determine disease emergence risk.

41 **Host richness does not predict virus distinctiveness**

42 Bats, the second most diverse group of mammals, are found worldwide and serve as the main animal  
43 reservoir for different strains of betacoronaviruses.<sup>12</sup> This has attracted attention to areas where high  
44 diversity of bats, and therefore presumably high diversity of betacoronaviruses, can be an important issue  
45 for human health.<sup>13,14</sup> By overlaying the IUCN range maps for confirmed bat hosts of betacoronaviruses  
46 [fig. 1; top], we see that the main hotspots (both in terms of size and higher values) of host richness are  
47 primarily South-Eastern Asia, parts of Southern Europe, and to a lesser extent parts of Africa in the -25-0  
48 range of latitudes. The description of host richness is an important first step towards understanding risk,  
49 as previous research<sup>2,15</sup> states that locally diverse bat communities could maintain more viruses and  
50 hence, a higher probability of having a pathogen that could represent a risk for human health.

51 [Figure 1 about here.]

52 Nevertheless, locally diverse and virus-rich bat communities could represent an increased risk of spillover  
53 under climate change through the creation of novel interactions,<sup>16</sup> and therefore the diversity of  
54 *Betacoronavirus* strains should similarly be accounted for. In fig. 1 (bottom), we contrast the evolutionary  
55 distinctiveness of bats and viruses – this reveals a slightly different portrait than bat richness alone.  
56 Chiropterans can be classified, from a macro-evolutionary standpoint, as Yangochiroptera and  
57 Yinpterochiroptera elsewhere.<sup>17,18</sup> Specifically, we would expect that the so-called “New World” group of  
58 bats, being more evolutionary distinct, would also have evolutionary distinct viruses. Indeed fig. 1

59 (bottom) reveals it to be the case, and this region harbors a distinct bat-betacoronaviruses complex. This  
60 can be explained by the fact that Yangochiroptera, although not limited to the western hemisphere,  
61 contain the highly diverse adaptive radiation in the Phyllostomidae,<sup>19</sup> which is restricted to the western  
62 hemisphere. By contrast, South-Eastern Asia has a lot of non-evolutionary distinct bats, who nevertheless  
63 hosted evolutionary-distinct viruses.

64 It is noteworthy that outside of South America, viral evolutionary distinctiveness does not accurately track  
65 host diversity, with some areas having over-distinct viruses (eastern China but, oddly, not the rest of  
66 southeast Asia). There are a number of likely explanations. First, given the richness of bats in southeast  
67 Asia, many betacoronaviruses likely remain to be discovered in this region. Indeed, global predictions by<sup>20</sup>  
68 highlight that southeast Asia is a likely hotspot of unconfirmed hosts of betacoronaviruses, which would  
69 likely result in additional viral discoveries. This idea is unsurprising given the growing realization,  
70 especially since the emergence of SARS-CoV-2, that unique lineages of similar viruses are widespread in  
71 bats but still mostly undescribed. The most distinct bats-betacoronaviruses complex is found in South  
72 America, a region with a comparatively lower number of hosts; this matches with the isolation through  
73 variance of the host group, and may highlight a different co-evolutionary dynamic. Alternatively, this  
74 distinctiveness hotspot may be a product of under-sampling: South-America is one of the places where the  
75 fewest *Betacoronavirus* sequences have been discovered,<sup>2,21,22</sup> resulting in sparser phylogenetic tree,  
76 thereby artificially inflating distinctiveness. Adding more viruses would bring the distinctiveness of  
77 known sequences down. Finally, South-America is the range of Phyllostomidae, a group of bats that  
78 underwent explosive diversification events,<sup>19</sup> which may drive the emergence of multiple viral lineages.

## 79 The phylogeographic regions of hosts and their viruses overlap

80 Despite the difference in evolutionary distinctiveness globally, there are reasons to expect that the  
81 phylogeography of bats and betacoronaviruses should show some degree of congruence.<sup>23,24</sup> In particular,  
82 this should be the case if viruses can circulate among hosts and co-evolve with local hosts communities,  
83 making their evolutionary process more than a byproduct of host evolution. High density of hosts sharing  
84 the same virus (albeit possibly different strains) can drive or result from evolution of the bat antiviral  
85 immune system, resulting in spatially distinct immunological responses, as evidenced in several bat  
86 species.<sup>25</sup> Immune characteristics that allow bats to be better adapted to infection by emerging viruses,<sup>26,27</sup>  
87 in addition to being hardcoded in their genome,<sup>28</sup> may be related to a wide variety of diets,<sup>25,29</sup> themselves

88 likely to be driven by spatial effects, especially at the local scale – bats, indeed, occupy a variety of  
89 environments, and therefore display a variety of adaptations to these environments.<sup>30</sup>

90 [Figure 2 about here.]

91 In fig. 2, we show a projection of the phylogeographic signal of bats (top) and viruses (bottom) in space;  
92 the distinct groupings (represented by different colors symbolizing positions in the subspace formed by  
93 the first two axes of the PCoA) are essentially equivalent between the two groups, and can be coarsely  
94 delineated as southeast Asia, Eurasia above a northing of 25, and Africa and south America. These results  
95 suggest that, although the evolutionary distinctiveness of the bat-betacoronaviruses complex varies  
96 spatially, the system shows an important degree of spatial consistency, with a reduced number of  
97 bioregions. Available information describing the spillover of zoonotic betacoronaviruses of bat origin  
98 where data was available before and up through the COVID-19 pandemic puts spillover events of  
99 SARS-CoV-2 in Wuhan, China; SARS-CoV in Guangdong, China based on the presence of closest known  
100 viruses circulating in nature, and a nearby location where serological (antibody) evidence has indicated  
101 human exposure to SARS-like viruses;<sup>31</sup> MERS-CoV in Saudi Arabia based on index cases available from a  
102 recently-published compendium of cases.<sup>32</sup> For the latest event, most if not all index cases are presumed to  
103 be camel-to-human transmission, and the precise origin point (if it exists) of MERS-CoV in bats is  
104 uncertain. Recent recombinant canine coronavirus spillover events in Haiti<sup>33</sup> and Europe<sup>34</sup> are not  
105 relevant here, as bats' involvement in these cycles of transmission have been supposed to be non-existent.  
106 These index cases fall within different phylogeographic bioregions (fig. 2), which further highlight the  
107 issue that different host-virus sub-systems may lead to widespread emergence.

## 108 **Coevolution-informed emergence risk is different in space**

109 As host richness, joint distinctiveness, or phylogeographic structure suggest that the bat-betacoronaviruses  
110 complex is globally fragmented enough to give rise to both different levels of risk (as evidenced by the  
111 spatial location of spillover events) and different types of co-evolutionary dynamics, we turn to the  
112 Geographic Mosaic Theory of Coevolution<sup>6</sup> to provide a measure of risk accounting for multiple processes.  
113 In fig. 3, we overlapped three components of spillover risk: viral sharing, *i.e.* the chance that two bats will  
114 share viruses overall; Local Contribution to Beta Diversity, *i.e.* the fact that a bat community is  
115 compositionally unique compared to the average compositional similarity across the entire system; finally,

116 host phylogenetic diversity, *i.e.* how dispersed the bats in a location are within the tree of life. This  
117 approach leads to the definition of broad biogeographic regions of risk, where the same color represents  
118 the same type of risk. By way of contrast to figures fig. 1 and fig. 2, these regions do not necessarily  
119 overlap with previous spatial partitions of the bat-betacoronaviruses complex.

120 [Figure 3 about here.]

121 From the perspective of spillover risk, the most important combination of factors is a high phylogenetic  
122 diversity of hosts with low viral sharing; this, essentially, means that very different betacoronaviruses  
123 could co-exist within the same place. This is particularly the case given that betacoronaviruses often  
124 evolve and even achieve host shifts through recombination, which requires the co-occurrence of  
125 sufficiently distinct viruses to be a major driver of emergence. In fig. 3, this corresponds to yellow to pale  
126 green areas, which are essentially limited to South-Eastern Asia, and to some part of Sub-Saharan Africa.  
127 Adopting a geographic mosaic theory perspective on risk, other regions of the world are of lesser concern  
128 fig. 4. Our risk decomposition does not account for viral diversity or distinctiveness. The simple rationale  
129 behind it is that the acquisition of viral data is rarely disconnected from the acquisition of host data. There  
130 are more sources of information on hosts than on viruses, allowing to develop a host-centric perspective  
131 on risk (although this estimate would be more accurate with viral traits related to *e.g.* ability to switch hosts  
132 or pathogenic potential). Areas with high bat diversity and high turnover *may* facilitate the evolutionary  
133 radiation of viruses, matching previous findings that the diversification of bat coronaviruses is driven  
134 largely by host shifts (inter-genus or higher levels of cross-species transmission) and, to a lesser degree,  
135 cospeciation and sharing.<sup>intra-genus cross-species transmission;</sup> <sup>2</sup> This diversification is not an actual risk factor for  
136 spillover itself, but acts downstream of a spillover event by increasing the random chance of the  
137 emergence of a virus with the raw genomic components required for the potential to infect humans.

138 [Figure 4 about here.]

139 From another perspective, areas of high host uniqueness and virus sharing (red-to-pink) could provide  
140 hotspots of *Betacoronavirus* risk through mixing of unique viruses (via codivergence) and in turn  
141 recombination. Under our framework, such a hotspot was identified in Madagascar, where most bat  
142 species are endemic following evolutionary divergence from sister species in both African and Asian  
143 continents.<sup>e.g.</sup> <sup>35</sup> Recent surveillance<sup>36</sup> has identified a novel *Betacoronavirus* (in the subgenus

144 *Nobecovirus*) in Madagascar-endemic pteropid bat species (*Pteropus rufus*, *Rousettus madagascariensis*),  
145 emphasizing strong proof of principle in model predictions.

146 **Human occupancy drives different levels of effective risk globally**

147 Based on the previous result, we extracted the risk component from the composite map (see Methods), to  
148 provide a single measure of risk varying between 0 and 1. This measure is presented in fig. 4. As this map  
149 represents the potential risk, it must be weighed by the potential for contacts with humans. As a proxy for  
150 this measure, we used the proportion of build/urban land from the EarthEnv dataset: this is a reasonable  
151 proxy for the density of humans per unit area, which increases the probability of pathogen spread more  
152 widely.<sup>37</sup> Since human activity is required to amplify the frequency of virus encounters and thus create  
153 areas of viral amplification, mapping the potential risk against measures of land use is required to generate  
154 a more actionable assessment of risk. This map is presented in fig. 5. Most of South America and Europe  
155 are at comparatively lower risk, as although densely populated, settlements tend to be in areas with lower  
156 potential risk. Regions like Malaysia and the North coast of Australia have a high risk component, but  
157 should represent a relatively lower effective risk due to low human density. However, this mapping reveals  
158 that South-East Asia, the Indian subcontinent, and parts of sub-Saharan Africa, are at high risk due to the  
159 overlap between built areas and bat communities representing more opportunities for cross-species  
160 transmission of betacoronaviruses. In looking for the origins of SARS in China,<sup>38</sup> present serological  
161 evidence that strongest human-animal contact results in higher risk of virus exposure, regardless of the  
162 animal species, but that different types of contact had different impacts. Ideally, finer-grained information  
163 about human activity (rather than human presence through anthropisation) could allow to partition this  
164 risk further, albeit at the cost of more hypotheses required to estimate the amount of risk represented by  
165 each activity. Our map of risk overlays with recent results from<sup>39</sup> – areas of purported high  
166 risk/diversitification potential (Madagascar, South-America) overlay with sampling gaps for  
167 *Betacoronavirus*, stressing the need for spatially targeted monitoring and discovery.

168

[Figure 5 about here.]

169 **Conclusion**

170 Bats are important reservoir hosts for different classes of microorganisms, many of which a threat to  
171 human health.<sup>23,24</sup> Chiropterans emerged around 64 million years ago and are one of the most diverse  
172 mammalian orders, with an estimated richness of more than 1400 species.<sup>40,41</sup> They exhibit a broad variety  
173 of habitat use, behaviour, and feeding strategies, putting them at key positions in the delivery and  
174 provisioning of several ecosystem services, tied to important ecosystem-derived benefits to human.<sup>42</sup> For  
175 example, bats are an essential component of many seed-dispersal networks.<sup>43</sup> Over two-thirds of bats are  
176 know to be either obligate or facultative insectivores, therefore actively contributing for agricultural pest  
177 control,<sup>44,45</sup> and vectors of pathogens that put a risk on human health.<sup>46,47</sup> Because bats are globally  
178 distributed and have a long evolutionary history, phylogeographic and biogeographic approaches are  
179 required to shed light on the contemporary distribution of coevolutionary processes between bats and the  
180 pathogens they host. Not all areas in which bats, viruses, and human are co-occurring are facing a risk of  
181 spillover towards human populations, and the areas in which this risk exist may not be facing risks of the  
182 same nature and magnitude.

183 Here, we propose a simple framework with broad explanatory power that helps contextualize discoveries  
184 like highly divergent nobecoviruses in Madagascar and the previously-neglected adaptive radiation of  
185 sarbecoviruses outside of southern China and throughout southeast Asia. In doing so, it advances  
186 ecological theory beyond the current state of the art for global maps of emergence risk. For example,  
187 previous studies that have used host richness as proxy have predicted a high diversity of unsampled bat  
188 viruses,<sup>22</sup> bat coronaviruses,<sup>2</sup> and even specifically betacoronaviruses<sup>20</sup> in both the Amazon and southeast  
189 Asia. While we find that both regions are characterized by highly divergent host and viral communities,  
190 our framework identifies key differences between the regions. We find that Latin America is a hotspot of  
191 both host and viral distinctiveness, suggesting that this branch of the bat-betacoronavirus complex may be  
192 undergoing independent evolutionary dynamics from the rest of the global pool, but with limited potential  
193 for viral diversification—a finding that is supported by previous work indicating a higher rate of  
194 codivergence in Latin America.<sup>2</sup> In contrast, in southeast Asia, host richness and viral distinctiveness are  
195 high but sharing is low; this suggests a different type of evolutionary dynamics that could generate high  
196 local diversity of viruses through host switching and viral recombination (see e.g.,<sup>48</sup> as well as the  
197 discovery of recombinant viruses that share genetic material from both the SARS-CoV and SARS-CoV-2

198 branches of the Sarbecovirus lineage).<sup>49</sup> Both of these regions are priority areas for sampling, especially  
199 given predictions that they contain many bat hosts of undiscovered betacoronaviruses.<sup>20,39</sup> However, both  
200 the evolutionary and ecological aspects of emergence risk are likely higher in southeast Asia—a fact that  
201 will only become more relevant, as bats track shifting climates and exchange viruses with other species,  
202 creating a hotspot of cross-species transmission unique to the region.<sup>16</sup>

203 [Tim tinker with para about conservation] There are several factors that drive changes in the diversity of  
204 bats,<sup>50</sup> but human activities' effects on the ecosystem (like modifications of land use) could significantly  
205 decrease it. Therefore, it can be suggested that changes in the diversity of betacoronaviruses in bats are  
206 linked to their biogeographic variation, and human population density and other anthropogenic factors  
207 are decisive moderators for its implications in public health. With the increase of contact between humans  
208 and potential hosts, we also increase the risk of emergence of novel diseases,<sup>51</sup> as previous studies on RNA  
209 viruses suggest the importance of host phylogeography at the time of virus dispersal.<sup>52</sup> One of these  
210 scenarios where interaction between bats and humans can occur can be seed dispersal in tropical  
211 agroecosystems. It opens the discussion of whether the fruits thrown by bats not only disperse seeds but  
212 could also be a source of indirect interaction between viruses of bat origin and humans.<sup>53</sup> This represents a  
213 challenge for conservation strategies and disease ecology since some areas can have both potential for the  
214 acquisition of zoonotic viruses and bat-human interactions; in particular, the challenge lies in the fact that  
215 actual exposure must then be quantified accounting for several transmission scenarios, including both  
216 direct and indirect bat - human interaction.

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

225 **Methods**

226 **Known *Betacoronavirus* hosts**

227 We downloaded the data on bats hosts of *Betacoronavirus* assembled by<sup>20</sup> from  
228 <https://www.viralemergence.org/betacov> on Apr. 2022, and filtered it to “known” hosts (established  
229 before the emergence of SARS-CoV-2) and “novel” hosts (confirmed through sampling and competence  
230 assays since the initial data collection). The original database was assembled by a combination of data  
231 mining and literature surveys, including automated alerts on the “bats” and “coronavirus” keywords to  
232 identify novel empirical evidence of bats-betacoronaviruses associations; this yielded a total of 126 known  
233 hosts, 47 of which were novel hosts.

234 **Bat occurrences**

235 We downloaded the rangemap of every current bat species that was classified as an empirically  
236 documented host of *Betacoronavirus* from the previous step, according to recent IUCN data.<sup>54</sup> The range  
237 maps were subsequently rasterized using the rasterize function from GDAL<sup>55</sup> at a resolution of  
238 approximately 100kmx100km. For every pixel in the resulting raster where at least one bat host of  
239 *Betacoronavirus* was present, we extract the species pool (list of all known bat hosts), which was used to  
240 calculate the following risk assessment components: bat phylogenetic diversity, bat compositional  
241 uniqueness, and predicted viral sharing risk.

242 **Bat phylogenetic diversity**

243 For every pixel, we measured Faith’s Phylogenetic Diversity<sup>56</sup> based on a recent synthetic tree with robust  
244 time calibration, covering about 6000 mammalian species.<sup>57</sup> Faith’s PD measures the sum of unique  
245 branches from an arbitrary root to a set of tips, and comparatively larger values indicate a more  
246 phylogenetic diverse species pool. We measured phylogenetic diversity starting from the root of the entire  
247 tree (and not from Chiroptera); this bears no consequences on the resulting values, since all branches  
248 leading up to Chiroptera are only counted one per species pool, and (as we explain when describing the  
249 assembly of the composite risk map), all individual risk components are ranged in [0,1]. This measure  
250 incorporates a richness component, which we chose not to correct for; the interpretation of the

251 phylogenetic diversity is therefore a weighted species richness, that accounts for phylogenetic  
252 over/under-dispersal in some places.

## 253 **Bat compositional uniqueness**

254 For every species pool, we measured its Local Contribution to Beta-Diversity;<sup>58</sup> LCBD works from a  
255 species-data matrix (traditionally noted as  $\mathbf{Y}$ ), where species are rows and sites are columns, and a value of  
256 1 indicates occurrence. We extracted the  $\mathbf{Y}$  matrix assuming that every pixel represents a unique location,  
257 and following best practices<sup>59</sup> transformed it using Hellinger's distance to account for unequal bat  
258 richness at different pixels. The correction of raw community data is particularly important for two  
259 reasons: first, it prevents the artifact of richer sites having higher importance; second, it removes the effect  
260 of overall species richness, which is already incorporated in the phylogenetic diversity component. High  
261 values of LCBD indicate that the pixel has a community that is on average more dissimilar in species  
262 composition than what is expected knowing the entire matrix, i.e. a more unique community. Recent  
263 results by<sup>60</sup> shows that LCBD measures are robust with regards to spatial scale, and are therefore  
264 applicable at the global scale.

## 265 **Viral sharing between hosts**

266 For all bat hosts of *Betacoronavirus*, we extracted their predicted viral sharing network, generated from a  
267 previously published generalized additive mixed model of virus sharing by a tensor function of  
268 phylogenetic distance and geographic range overlap.<sup>61</sup> This network stores pairwise values of viral  
269 community similarity. To project viral sharing values into a single value for every pixel, we averaged the  
270 pairwise scores. High values of the average sharing propensity means that this specific extant bat  
271 assemblage is likely to be proficient at exchanging viruses.

## 272 **Composite risk map**

273 To visualize the aggregated risk at the global scale, we combine the three individual risk components  
274 (phylogenetic diversity, compositional uniqueness, and viral sharing) using an additive color model.<sup>62</sup> In  
275 this approach, every risk component gets assigned a component in the RGB color model (phylogenetic  
276 diversity is green, compositional uniqueness is red, and viral sharing is blue). In order to achieve a valid

277 RGB measure, all components are re-scaled to the [0,1] interval, so that a pixel with no sharing, no  
278 phylogenetic diversity, and no compositional uniqueness is black, and a pixel with maximal values for  
279 each is white. This additive model conveys both the intensity of the overall risk, but also the nature of the  
280 risk as colors diverge towards combinations of values for three risk components. Out of the possible  
281 combinations, the most risky in terms of rapid diversification and spillover potential is high phylogenetic  
282 diversity and low viral sharing,<sup>63,64</sup> in that this allows multiple independent host-virus coevolutionary  
283 dynamics to take place in the same location. In the colorimetric space, this corresponds to yellow – because  
284 the HSV space is more amenable to calculations for feature extraction,<sup>see e.g.</sup><sup>65</sup> we measured the risk level  
285 by calculating the angular distance of the hue of each pixel to a reference value of 60, and weighted this  
286 risk level by the value component. Specifically, given a pixel with colorimetric coordinates  $(h, s, v)$ , its  
287 ranged weighted risk value is

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

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

## 289 **Viral phyogeography and evolutionary diversification**

290 To next represent phyogeography of betacoronaviruses in bats, we aggregated and analyzed  
291 betacoronavirus sequence data. We used the following query to pull all *Betacoronavirus* sequence data  
292 from the GenBank Nucleotide database except SARS-CoV-2; (“Betacoronavirus”[Organism] OR  
293 betacoronavirus[All Fields]) NOT (“Severe acute respiratory syndrome coronavirus 2”[Organism] OR  
294 sars-cov-2[All Fields]). We added a single representative sequence for SARS-CoV-2 and manually curated  
295 to remove sequences without the RNA-dependent RNA polymerase (RdRp) sequence or that contained  
296 words indicating recombinant or laboratory strains including “patent”, “mutant”, “GFP”, and  
297 “recombinant”. We filtered over-represented taxa including betacoronavirus 1, hCoV-OC43, Middle East  
298 respiratory syndrome coronavirus, Murine hepatitis virus, and hCoV-HKU1. Curated betacoronavirus  
299 RdRp sequences were then aligned using MAFFT v 1.4.0 [<sup>66</sup>; Algorithm FFT-NS-2, Scoring matrix  
300 200PAM / k=2, gap open penalty 1.53m offset value 0.123] and a maximum likelihood tree reconstructed  
301 in IQ-TREE v 1.6.12<sup>67</sup> with ModelFinder<sup>68</sup> ultrafast bootstrap approximation<sup>69</sup> with a general time  
302 reversible model with empirical base frequencies and the 5-discrete-rate-category FreeRaye model of

303 nucleotide substitution (GTR+F+R5).

304 We first tested the hypothesis that hotspots of viral diversification would track hotspots of bat  
305 diversification. To do so, we plotted the number of known bat hosts (specifically only those included in the  
306 phylogeny, so there was a 1:1 correspondence between data sources) against the “mean evolutionary  
307 distinctiveness” of the associated viruses. To calculate this, we derived the fair proportions evolutionary  
308 distinctiveness<sup>70</sup> for each of the viruses in the tree, then averaged these at the bat species level, projected  
309 these values onto their geographic distributions, and averaged across every bat found in a given pixel. As  
310 such, this can be thought of as a map of the mean evolutionary distinctiveness of the known viral  
311 community believed to be associated with a particular subset of bats present.

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

313 Subsequently, we tested the hypothesis that the biogeography of bat betacoronaviruses should track the  
314 biogeography of their hosts. To test this idea, we loosely adapted a method from,<sup>71,72</sup> who proposed a  
315 phylogenetic method for the delineation of animal biogeographic regions. In their original method, a  
316 distance matrix - where each row or column represents a geographic raster’s grid cell, and the dissimilarity  
317 values are the “beta diversity similarity” of their community assemble - undergoes non-metric  
318 multidimensional scaling (NMDS); the first two axes of the NMDS are projected geographically using a  
319 four-color bivariate map. Here, we build on this idea with an entirely novel methodology. First, we  
320 measure the phylogenetic distance between the different viruses in the betacoronaviruses tree by using the  
321 cophenetic function in ape;<sup>73</sup> subsequently, we take a principal components analysis of that distance  
322 matrix (readily interchangeable for NMDS in this case) to project the viral tree into an n-dimensional  
323 space. We then take the first two principal components and, as with the evolutionary distinctiveness  
324 analysis, aggregated these to a mean host value and projected them using a four-color bivariate map.

### 325 **References**

326 1.

327 Plowright, R. K. *et al.* Pathways to zoonotic spillover. *Nature Reviews Microbiology* **15**, 502–510 (2017).

328 2.

- 329 Anthony, S. J. *et al.* Global patterns in coronavirus diversity. *Virus Evolution* **3**, (2017).
- 330 3.
- 331 Ruiz-Aravena, M. *et al.* Ecology, evolution and spillover of coronaviruses from bats. *Nature Reviews Microbiology* **20**, 299–314 (2022).
- 332 4.
- 333 Agosta, S. J., Janz, N. & Brooks, D. R. How specialists can be generalists: Resolving the ‘parasite paradox’ and implications for emerging infectious disease. *Zoologia (Curitiba)* **27**, 151–162 (2010).
- 334 5.
- 335 Power, A. G. & Mitchell, C. E. Pathogen Spillover in Disease Epidemics. *The American Naturalist* **164**, S79–S89 (2004).
- 336 6.
- 337 Thompson, J. N. *The Geographic Mosaic of Coevolution*. (University Of Chicago Press, 2005).
- 338 7.
- 339 Thompson, J. N. *The Coevolutionary Process*. (University of Chicago Press, 1994).
- 340 8.
- 341 Janzen, D. H. When is it Coevolution? *Evolution* **34**, 611–612 (1980).
- 342 9.
- 343 Price, P. W. *Macroevolutionary Theory on Macroecological Patterns*. (Cambridge University Press, 2002).  
doi:[10.1017/CBO9780511615030](https://doi.org/10.1017/CBO9780511615030).
- 344 10.
- 345 Gomulkiewicz, R. *et al.* Dos and don’ts of testing the geographic mosaic theory of coevolution. *Heredity* **98**, 249–258 (2007).
- 346 11.
- 347 Thrall, P. H., Hochberg, M. E., Burdon, J. J. & Bever, J. D. Coevolution of symbiotic mutualists and parasites in a community context. *Trends in Ecology & Evolution* **22**, 120–126 (2007).
- 348 12.
- 349 Drexler, J. F., Corman, V. M. & Drosten, C. Ecology, evolution and classification of bat coronaviruses in the aftermath of SARS. *Antiviral Research* **101**, 45–56 (2014).
- 350 13.
- 351 Calisher, C. H., Childs, J. E., Field, H. E., Holmes, K. V. & Schountz, T. Bats: Important Reservoir Hosts of Emerging Viruses. *Clinical Microbiology Reviews* **19**, 531–545 (2006).

- 352 14.
- 353 Moratelli, R. & Calisher, C. H. Bats and zoonotic viruses: Can we confidently link bats with emerging deadly  
viruses? *Memórias do Instituto Oswaldo Cruz* **110**, 1–22 (2015).
- 354 15.
- 355 Mollentze, N. & Streicker, D. G. Viral zoonotic risk is homogenous among taxonomic orders of mammalian  
and avian reservoir hosts. *Proceedings of the National Academy of Sciences* **117**, 9423–9430 (2020).
- 356 16.
- 357 Carlson, C. J. *et al.* Climate change increases cross-species viral transmission risk. *Nature* 1–1 (2022)  
doi:[10.1038/s41586-022-04788-w](https://doi.org/10.1038/s41586-022-04788-w).
- 358 17.
- 359 Springer, M. S. Phylogenetics: Bats United, Microbats Divided. *Current Biology* **23**, R999–R1001 (2013).
- 360 18.
- 361 Teeling, E. C. *et al.* A Molecular Phylogeny for Bats Illuminates Biogeography and the Fossil Record. *Science*  
(New York, N.Y.) **307**, 580–584 (2005).
- 362 19.
- 363 Villalobos, F. & Arita, H. T. The diversity field of New World leaf-nosed bats (Phyllostomidae). *Global  
Ecology and Biogeography* **19**, 200–211 (2010).
- 364 20.
- 365 Becker, D. J. *et al.* Optimising predictive models to prioritise viral discovery in zoonotic reservoirs. *The  
Lancet Microbe* (2022) doi:[10.1016/S2666-5247\(21\)00245-7](https://doi.org/10.1016/S2666-5247(21)00245-7).
- 366 21.
- 367 Allen, T. *et al.* Global hotspots and correlates of emerging zoonotic diseases. *Nature Communications* **8**,  
(2017).
- 368 22.
- 369 Olival, K. J. *et al.* Host and viral traits predict zoonotic spillover from mammals. *Nature* **546**, 646–650 (2017).
- 370 23.
- 371 Van Brussel, K. & Holmes, E. C. Zoonotic disease and virome diversity in bats. *Current Opinion in Virology*  
**52**, 192–202 (2022).
- 372 24.
- 373 Letko, M., Seifert, S. N., Olival, K. J., Plowright, R. K. & Munster, V. J. Bat-borne virus diversity, spillover  
and emergence. *Nature Reviews Microbiology* **18**, 461–471 (2020).
- 374 25.
- 375 Banerjee, A. *et al.* Novel Insights Into Immune Systems of Bats. *Frontiers in Immunology* **11**, 26 (2020).

- 376 26.
- 377 Gorbunova, V., Seluanov, A. & Kennedy, B. K. The World Goes Bats: Living Longer and Tolerating Viruses. *Cell Metabolism* **32**, 31–43 (2020).
- 378 27.
- 379 Irving, A. T., Ahn, M., Goh, G., Anderson, D. E. & Wang, L.-F. Lessons from the host defences of bats, a unique viral reservoir. *Nature* **589**, 363–370 (2021).
- 380 28.
- 381 Jebb, D. *et al.* Six reference-quality genomes reveal evolution of bat adaptations. *Nature* **583**, 578–584 (2020).
- 382 29.
- 383 Moreno Santillán, D. D. *et al.* Large-scale genome sampling reveals unique immunity and metabolic adaptations in bats. *Molecular Ecology* **mec.16027** (2021) doi:[10.1111/mec.16027](https://doi.org/10.1111/mec.16027).
- 384 30.
- 385 Muylaert, R. L. *et al.* Present and future distribution of bat hosts of sarbecoviruses: Implications for conservation and public health. *Proceedings of the Royal Society B: Biological Sciences* **289**, 20220397 (2022).
- 386 31.
- 387 Wang, N. *et al.* Serological Evidence of Bat SARS-Related Coronavirus Infection in Humans, China. *Virologica Sinica* **33**, 104–107 (2018).
- 388 32.
- 389 Ramshaw, R. E. *et al.* A database of geopositioned Middle East Respiratory Syndrome Coronavirus occurrences. *Scientific Data* **6**, 318 (2019).
- 390 33.
- 391 Lednicky, J. A. *et al.* Isolation of a Novel Recombinant Canine Coronavirus from a Visitor to Haiti: Further Evidence of Transmission of Coronaviruses of Zoonotic Origin to Humans. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America* **ciab924** (2021) doi:[10.1093/cid/ciab924](https://doi.org/10.1093/cid/ciab924).
- 392 34.
- 393 Vlasova, A. N. *et al.* Animal alphacoronaviruses found in human patients with acute respiratory illness in different countries. *Emerging Microbes & Infections* **11**, 699–702 (2022).
- 394 35.
- 395 Shi, J. J. *et al.* A Deep Divergence Time between Sister Species of Eidolon (Pteropodidae) with Evidence for Widespread Panmixia. *Acta Chiropterologica* **16**, 279–292 (2014).
- 396 36.
- 397 Kettenburg, G. *et al.* Full Genome Nobecovirus Sequences From Malagasy Fruit Bats Define a Unique Evolutionary History for This Coronavirus Clade. *Frontiers in Public Health* **10**, (2022).

- 398 37.
- 399 Hazarie, S., Soriano-Paños, D., Arenas, A., Gómez-Gardeñes, J. & Ghoshal, G. Interplay between population  
density and mobility in determining the spread of epidemics in cities. *Communications Physics* **4**, 191 (2021).
- 400 38.
- 401 Xu, R.-H. *et al.* Epidemiologic Clues to SARS Origin in China. *Emerging Infectious Diseases* **10**, 1030–1037  
(2004).
- 402 39.
- 403 Cohen, L. E., Fagre, A. C., Chen, B., Carlson, C. J. & Becker, D. J. Sampling strategies and pre-pandemic  
surveillance gaps for bat coronaviruses. 2022.06.15.496296 (2022) doi:[10.1101/2022.06.15.496296](https://doi.org/10.1101/2022.06.15.496296).
- 404 40.
- 405 Peixoto, F. P., Braga, P. H. P. & Mendes, P. A synthesis of ecological and evolutionary determinants of bat  
diversity across spatial scales. *BMC ecology* **18**, 18 (2018).
- 406 41.
- 407 Simmons, N. B. & Cirranello, A. L. Bat Species of the World: A taxonomic and geographic database. [https://  
//batnames.org/](https://batnames.org/) (2020).
- 408 42.
- 409 Kasso, M. & Balakrishnan, M. Ecological and Economic Importance of Bats (Order Chiroptera). *ISRN  
Biodiversity* **2013**, e187415 (2013).
- 410 43.
- 411 Mello, M. A. R. *et al.* The Missing Part of Seed Dispersal Networks: Structure and Robustness of Bat-Fruit  
Interactions. *PLOS ONE* **6**, e17395 (2011).
- 412 44.
- 413 *Bats in the Anthropocene: Conservation of Bats in a Changing World.* (Springer International Publishing,  
2016). doi:[10.1007/978-3-319-25220-9](https://doi.org/10.1007/978-3-319-25220-9).
- 414 45.
- 415 Williams-Guillén, K., Perfecto, I. & Vandermeer, J. Bats Limit Insects in a Neotropical Agroforestry System.  
*Science* **320**, 70–70 (2008).
- 416 46.
- 417 Gonsalves, L., Bicknell, B., Law, B., Webb, C. & Monamy, V. Mosquito Consumption by Insectivorous Bats:  
Does Size Matter? *PLOS ONE* **8**, e77183 (2013).
- 418 47.
- 419 Gonsalves, L., Lamb, S., Webb, C., Law, B. & Monamy, V. Do mosquitoes influence bat activity in coastal  
habitats? *Wildlife Research* **40**, 10–24 (2013).

- 420 48.
- 421 Latinne, A. *et al.* Origin and cross-species transmission of bat coronaviruses in China. *bioRxiv: The Preprint*  
*Server for Biology* 2020.05.31.116061 (2020) doi:[10.1101/2020.05.31.116061](https://doi.org/10.1101/2020.05.31.116061).
- 422 49.
- 423 Wu, Z. *et al.* A comprehensive survey of bat sarbecoviruses across China for the origin tracing of SARS-CoV  
and SARS-CoV-2. (2021) doi:[10.21203/rs.3.rs-885194/v1](https://doi.org/10.21203/rs.3.rs-885194/v1).
- 424 50.
- 425 Alves, D. M. C. C., Diniz-Filho, J. A. F., da Silva e Souza, K., Gouveia, S. F. & Villalobos, F. Geographic  
variation in the relationship between large-scale environmental determinants and bat species richness.  
*Basic and Applied Ecology* **27**, 1–8 (2018).
- 426 51.
- 427 Johnson, C. K. *et al.* Global shifts in mammalian population trends reveal key predictors of virus spillover  
risk. *Proceedings of the Royal Society B: Biological Sciences* **287**, 20192736 (2020).
- 428 52.
- 429 Gryseels, S. *et al.* When Viruses Don't Go Viral: The Importance of Host Phylogeographic Structure in the  
Spatial Spread of Arenaviruses. *PLOS Pathogens* **13**, e1006073 (2017).
- 430 53.
- 431 Deshpande, K., Vanak, A. T., Devy, M. S. & Krishnaswamy, J. Forbidden fruits? Ecosystem services from  
seed dispersal by fruit bats in the context of latent zoonotic risk. *Oikos (Copenhagen, Denmark)* oik.08359  
(2022) doi:[10.1111/oik.08359](https://doi.org/10.1111/oik.08359).
- 432 54.
- 433 IUCN. The IUCN Red List of Threatened Species. <https://www.iucnredlist.org> (2021).
- 434 55.
- 435 Rouault, E. *et al.* *GDAL/OGR Geospatial Data Abstraction software Library.* (Zenodo, 2022).  
doi:[10.5281/ZENODO.5884351](https://doi.org/10.5281/ZENODO.5884351).
- 436 56.
- 437 Faith, D. P. Conservation evaluation and phylogenetic diversity. *Biological Conservation* **61**, 1–10 (1992).
- 438 57.
- 439 Upham, N. S., Esselstyn, J. A. & Jetz, W. Inferring the mammal tree: Species-level sets of phylogenies for  
questions in ecology, evolution, and conservation. *PLOS Biology* **17**, e3000494 (2019).
- 440 58.
- 441 Legendre, P. & De Cáceres, M. Beta diversity as the variance of community data: Dissimilarity coefficients  
and partitioning. *Ecology Letters* **16**, 951–963 (2013).

- 442 59.
- 443 Legendre, P. & Condit, R. Spatial and temporal analysis of beta diversity in the Barro Colorado Island forest  
dynamics plot, Panama. *Forest Ecosystems* **6**, 7 (2019).
- 444 60.
- 445 Dansereau, G., Legendre, P. & Poisot, T. Evaluating ecological uniqueness over broad spatial extents using  
species distribution modelling. *Oikos* **n/a**, e09063 (2022).
- 446 61.
- 447 Albery, G. F., Eskew, E. A., Ross, N. & Olival, K. J. Predicting the global mammalian viral sharing network  
using phylogeography. *Nature Communications* **11**, 2260 (2020).
- 448 62.
- 449 Seekell, D. A., Lapierre, J.-F. & Cheruvilil, K. S. A geography of lake carbon cycling. *Limnology and  
Oceanography Letters* **3**, 49–56 (2018).
- 450 63.
- 451 Cavender-Bares, J., Kozak, K. H., Fine, P. V. A. & Kembel, S. W. The merging of community ecology and  
phylogenetic biology. *Ecol. Lett.* **12**, 693–715 (2009).
- 452 64.
- 453 Gomulkiewicz, R., Thompson, J. N., Holt, R. D., Nuismer, S. L. & Hochberg, M. E. Hot Spots, Cold Spots,  
and the Geographic Mosaic Theory of Coevolution. *The American Naturalist* **156**, 156–174 (2000).
- 454 65.
- 455 Keke, S., Peng, Z. & Guohui, L. Study on skin color image segmentation used by Fuzzy-c-means arithmetic.  
in *2010 Seventh International Conference on Fuzzy Systems and Knowledge Discovery* vol. 2 612–615 (2010).
- 456 66.
- 457 Katoh, K. & Standley, D. M. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in  
Performance and Usability. *Molecular Biology and Evolution* **30**, 772–780 (2013).
- 458 67.
- 459 Nguyen, L.-T., Schmidt, H. A., von Haeseler, A. & Minh, B. Q. IQ-TREE: A Fast and Effective Stochastic  
Algorithm for Estimating Maximum-Likelihood Phylogenies. *Molecular Biology and Evolution* **32**, 268–274  
(2015).
- 460 68.
- 461 Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A. & Jermiin, L. S. ModelFinder: Fast model  
selection for accurate phylogenetic estimates. *Nature Methods* **14**, 587–589 (2017).
- 462 69.

- 463 Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q. & Vinh, L. S. UFBoot2: Improving the Ultrafast  
Bootstrap Approximation. *Molecular Biology and Evolution* **35**, 518–522 (2018).
- 464 70.
- 465 Isaac, N. J. B., Turvey, S. T., Collen, B., Waterman, C. & Baillie, J. E. M. Mammals on the EDGE: Conservation  
Priorities Based on Threat and Phylogeny. *PLOS ONE* **2**, e296 (2007).
- 466 71.
- 467 Kreft, H. & Jetz, W. Global patterns and determinants of vascular plant diversity. *Proceedings of the National  
Academy of Sciences* **104**, 5925–5930 (2007).
- 468 72.
- 469 Kreft, H. & Jetz, W. A framework for delineating biogeographical regions based on species distributions.  
*Journal of Biogeography* **37**, 2029–2053 (2010).
- 470 73.
- 471 Paradis, E. & Schliep, K. Ape 5.0: An environment for modern phylogenetics and evolutionary analyses in  
R. *Bioinformatics* **35**, 526–528 (2019).

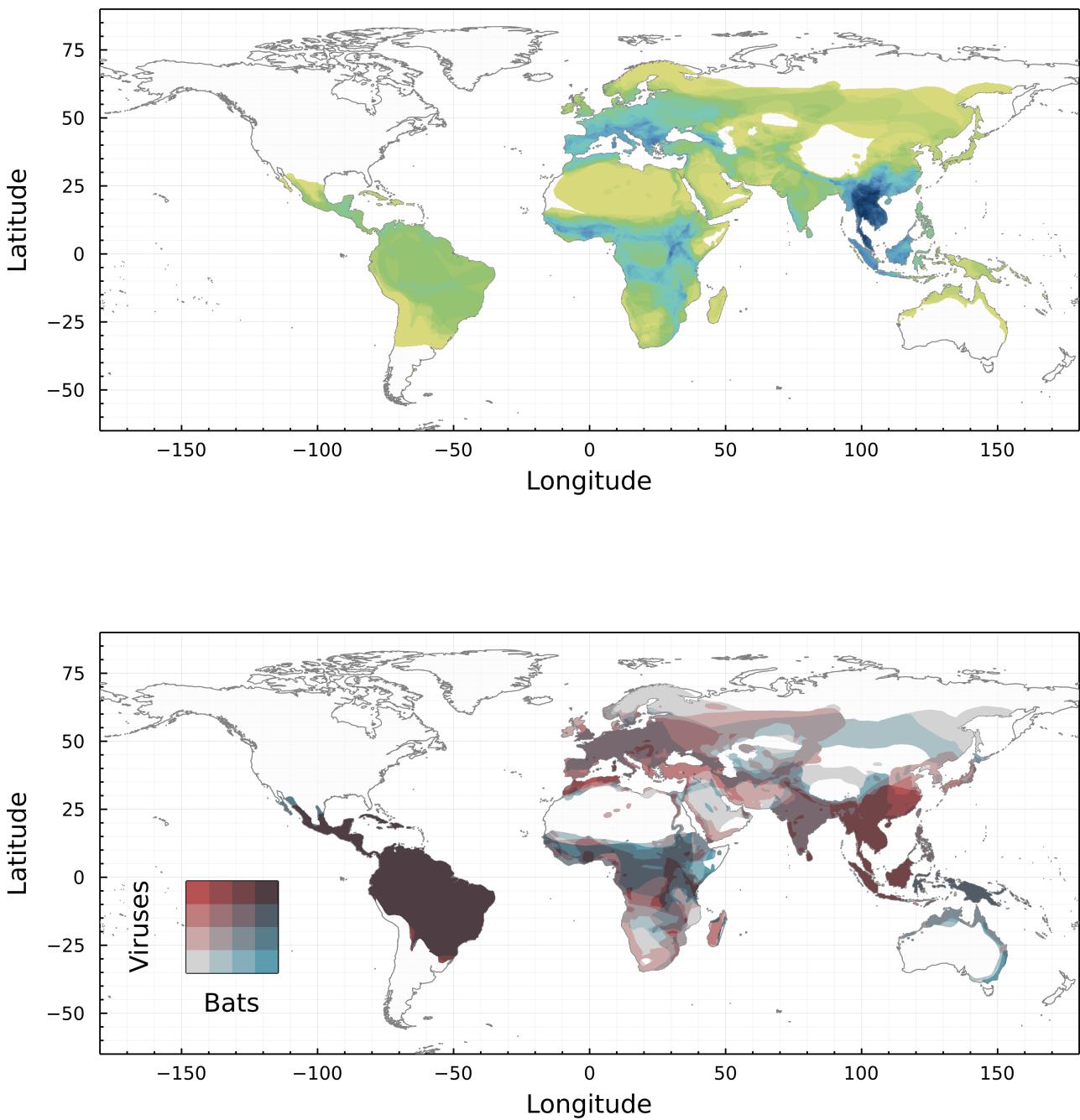


Figure 1: Top panel: relative diversity of known bat hosts of betacoronaviruses. This map shows that the region with the largest number of possible hosts is South-Eastern Asia. Bottom panel: congruence between the evolutionary distinctiveness of the hosts (grey to blue) and the viruses (grey to red). By contrast to the richness map, this reveals that South America has the most evolutionary distinct hosts *and* viruses, whereas South-Eastern Asia and the Rift Valley region have mostly distinct viruses. This is congruent with known results about New World bats being evolutionary distinct, and suggests that they similarly have distinct viruses.



Figure 2: Phylogeographic regions of bats (top) and viruses (bottom) based on the joint analysis of their occurrence and phylogenetic relatedness. The different colors show tendencies to separate alongside the first two components of a PCoA. Note that the PCoA for the bats and viruses are independent, and so cannot be compared directly – that being said, the regions can be compared across maps.

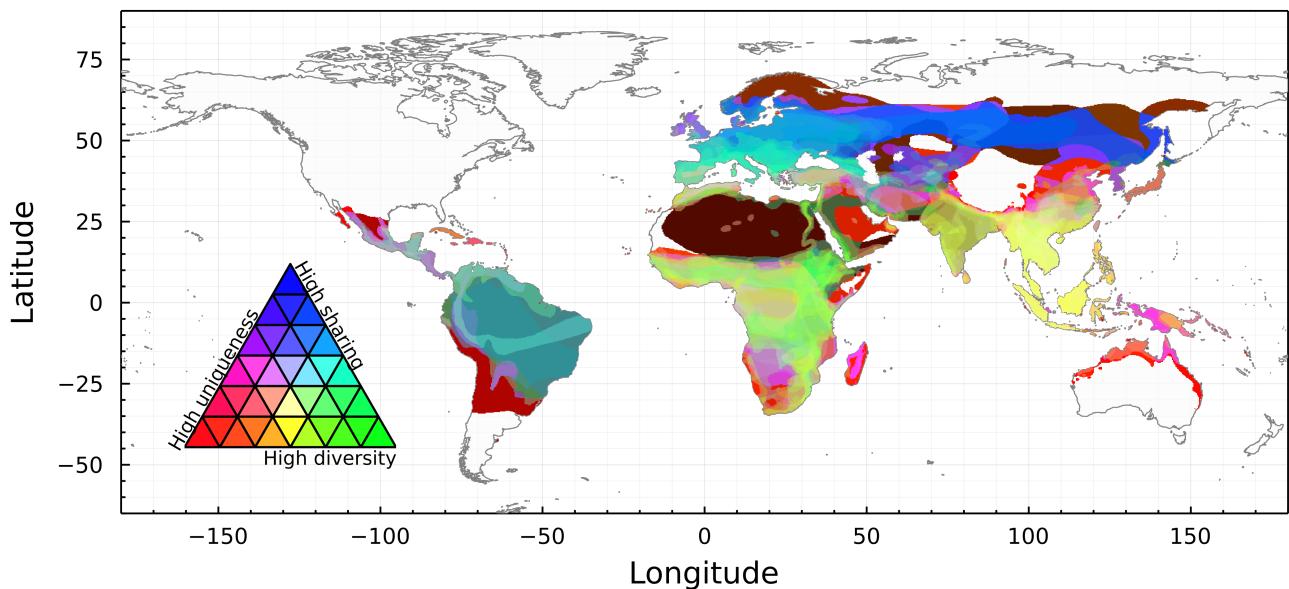


Figure 3: Trivariate additive mapping of the components of risk in the red/green/blue, where high virus sharing is encoded in the blue channel, host phylogenetic diversity in the green channel, and compositional uniqueness in the red channel. A pixel that would maximize all measures (highest possible risk) would be a pure white (specifically  $\text{RGB}(1.0, 1.0, 1.0)$ ), and a pixel with the lowest possible values would be pure black (specifically  $\text{RGB}(0.0, 0.0, 0.0)$ ). Therefore, lighter values (the sum of the three channels gets closer to 3) indicate higher risk, and the color indicates the proportional distribution of the factors making up the total risk.

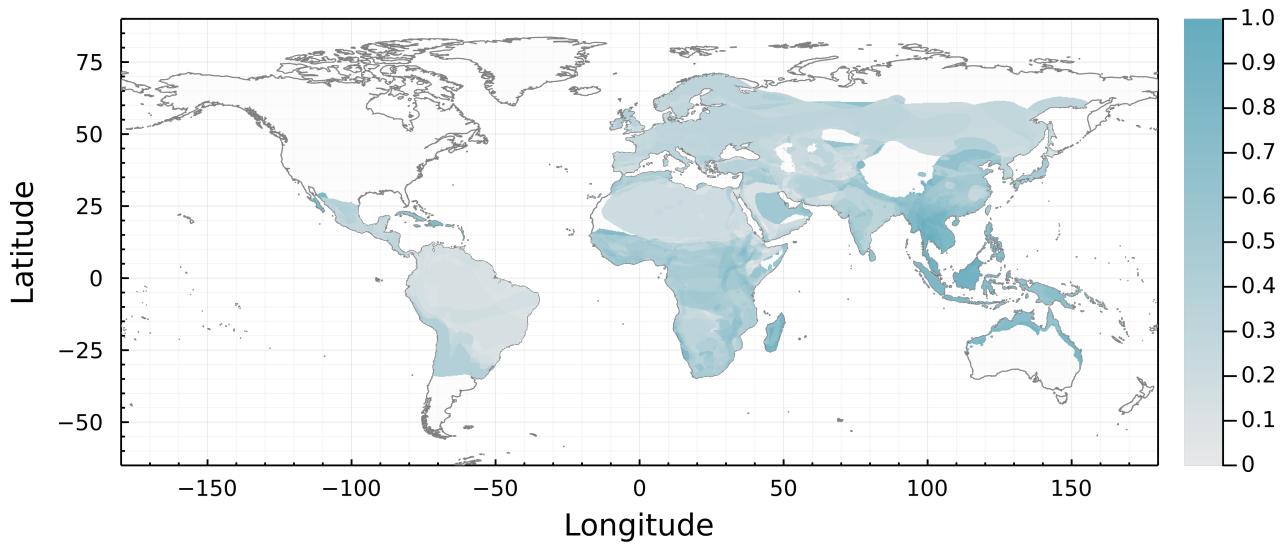


Figure 4: Extraction of a measure of *Betacoronavirus* spillover risk from bat hosts based on the colorimetric space from fig. 3. The risk is a composite measure of the color value and angular distance to the yellow hue, as defined in the methods, ranged in the unit space. Based on this analyses, regions at high risk of spillover are southeast Asia and Madagascar.

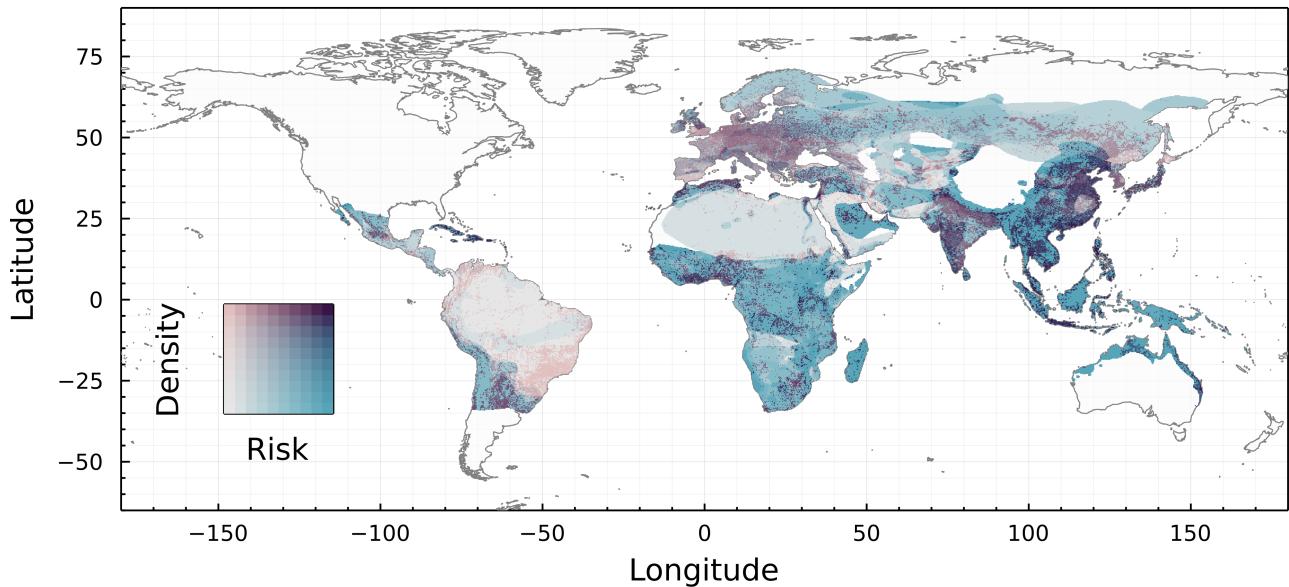


Figure 5: Overlap of the percent of each pixel occupied by urbanized structures, representing the degree of settlement, on the spillover risk map (where the risk comes only from wildlife, and ignores multi-hosts chains of transmissions including non-bats hosts). Darker pixels correspond to more risk, in that the GMTC-derived risk of fig. 4 is high *and* the pixel is densely occupied by human populations. This approach increases the relative risk of several regions in Africa, and highlights the risk in India, southeast China, and the Arabian peninsula where areas of high to moderate risk overlap with areas of denser population.