

The coevolutionary mosaic of *Betacoronavirus* spillover risk from bat hosts

Norma Forero Rocio Munoz^{1,2,‡} Renata L. Muylaert³ Stephanie N. Seifert⁴ Gregory F. Albery⁵

Daniel J. Becker⁶ Colin J. Carlson^{7,8,9,‡} Timothée Poisot^{1,2,‡}

¹ Université de Montréal ² Québec Centre for Biodiversity Sciences ³ Molecular Epidemiology and Public Health Laboratory, School of Veterinary Science, Massey University, New Zealand ⁴ Paul G. Allen School for Global Health, Washington State University, Pullman, WA, United States ⁵ Department of Biology, Georgetown University, Washington, DC, USA ⁶ Department of Biology, University of Oklahoma, Norman, OK, USA ⁷ Department of Biology, Georgetown University, Washington, DC,

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

‡ These authors contributed equally to the work

Correspondance to:

Timothée Poisot — 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 Spillover risk is multidimensional and complicated. Within a pool of competent hosts, it is driven by a
2 multiplicity of factors (Plowright et al. 2017). Although proxies for the local richness of hosts are
3 commonly analysed (Anthony et al. 2017, Ruiz-Aravena et al. 2022), there is an argument to be made that
4 species who are not (known as) competent hosts of a specific virus genus may not factor into this
5 (Plowright et al. 2015), calling for species-level information. This is especially true as competence data
6 increases predictive power when the taxonomic scope of hosts of a viral family increases (Becker et al.
7 2020, Mull et al. 2022). Similarly, host species who share viruses at different rates should be weighted
8 accordingly (Albery et al. 2020). In mammals, key functional traits (for which phylogeny is a reasonable
9 proxy) are determinants of the spillover potential of their viruses (Olival et al. 2017); these include,
10 notably, body mass, life history, diet, and use for human-occupied spaces (Albery et al. 2022). Finally,
11 especially when the pool of potential hosts spans the entire globe, there may be local host pools that are
12 highly unique; not having been observed in other locations, these can act on the overall risk either by
13 providing novel contact opportunities, reflecting unique host-environment combinations (Engering et al.
14 2013), or facilitating rapid evolutionary changes in specialism of their pathogens (Agosta et al. 2010). In
15 the specific case of generalist pathogens (which is the case for many viruses in the *Betacoronavirus* genus,
16 see e.g. MacLean et al. 2021), there is conceptual and empirical support to the idea that these
17 community-level mechanisms are even more important in driving the overall risk (Power and Mitchell
18 2004). In this paper, we examine the biogeographic structure of bats-betacoronaviruses associations, based
19 on a curated dataset of all confirmed bat hosts of betacoronaviruses, by adopting an analysis framework
20 grounded in concepts from the Geographic Mosaic Theory of Coevolution (GMTC; Thompson 2005). The
21 GMTC posits that coevolutionary processes between pairs (Thompson 1994) or complexes (Janzen 1980)
22 of species are structured in space by the rippling effects of abiotic conditions onto evolutionary
23 mechanism, resulting in spatially fragmented evolutionary dynamics, coupled only by dispersal-related
24 processes (Gomulkiewicz et al. 2000). In turn, these spatially fragmented processes can lead
25 taxonomically homogeneous systems (in our case, the bats-betacoronaviruses complex) to have different
26 structure and dynamics of large spatial extents (see e.g. Price 2002).

27 The GMTC predicts a spatial fragmentation of coevolutionary dynamics under the joint action of three
28 processes (see notably Gomulkiewicz et al. 2007), which all have the potential to act on outbreak
29 potential, pathogen transmission, and disease virulence (Parratt et al. 2016, Turner et al. 2021). First
30 hot/cold spots of coevolution can appear when the intensity of *interaction* (in terms of reciprocal fitness

consequences) varies spatially, because of e.g. partial range overlap between organisms (Nuismer et al. 2003); hot/cold spots are notoriously difficult to identify in nature (Laine 2005), especially when the structure of range overlap is complex, and the reciprocal fitness effects are relatively low. Both of these conditions are met in the bats-coronaviruses complex. Second, GMTC supposes the existence of selection mosaics, wherein the intensity of *selection* (in terms of selective consequences of fitness effects) varies across space; the strength of reciprocal selection responds to the biotic complexity of the community (locally diverse hosts and viruses are more biotically complex; Thrall et al. 2007) *and* to the local favorability of the environment (Hochberg and Baalen 1998). In a system with near global distribution and variations in relative richness, like the bats-betacoronaviruses complex, the GMTC predicts the establishment of many different coevolutionary trajectories. The final GMCT process is trait remixing, under which changes in coevolutionary dynamics happen due to the arrival (or departure) of functional traits, through changes in community composition due to invasions, meta-community dynamics, and dispersal; adopting an host-centric perspective, high viral sharing would expose the host immune system to multiple new viral traits, thereby changing the selection landscape for all viruses.

We turn the processes on the GMTC into definitions of spillover risk from viruses to hosts (focusing on the bats-betacoronavirus complex), with a specific emphasis on the potential to create independant coevolutionary dynamics (and therefore spatial fragmentation in the risk) through heterogeneity. Our components of risk are (i) the phylogenetic diversity of hosts, representing the diversification potential of the system (hotspots, selection mosaics), under the assumption that bat diversification is not primarily driven by viruses (Gorbunova et al. 2020); (ii) the local uniqueness of the bat community, representing the potential for viruses to be exposed to novel host traits in unusual combinations (cold spots, trait remixing); finally, (iii) viral sharing among hosts, representing the potential for hosts to be exposed to complex and functionaly diverse communities (trait remixing, biotic complexity). We turn these predicates into a spatially explicit additive mapping of zoonotic risk components, revealing the extreme heterogeneity of risk at the global scale. Explicitely framing the notion of spillover risk based on propositions from the GMTC (which is to say, based on a framework linking interactions between species to change within species) is a novel idea, that should be relatively general. Indeed, it only assumes the action of well described evolutionary mechanisms. The benefit of this approach is to provide the potential for a more dynamic and nuanced understanding of risk: not only on ecological timescales, but also by providing clues about which areas can change over micro-evolutionary timescales. This provides a way to look at

61 spatial structure by accounting for more notions than species richness/similarity, but also a way to identify
62 spatial areas of higher risk.

63 **Methods**

64 **Known *Betacoronavirus* hosts**

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

72 **Bats occurrences**

73 We downloaded the rangemap of every current bat species that was classified as an empirically
74 documented host of *Betacoronavirus* from the previous step, according to recent IUCN data (IUCN 2021).
75 The range maps were subsequently rasterized using the rasterize function from GDAL (Rouault et al.
76 2022) at a resolution of approximately 100kmx100km. For every pixel in the resulting raster where at least
77 one bat host of *Betacoronavirus* was present, we extract the species pool (list of all competent bat hosts),
78 which was used to calculate the following risk assessment components: bat phylogenetic diversity, bat
79 compositional uniqueness, and predicted viral sharing risk.

80 **Bats phylogenetic diversity**

81 For every pixel, we measured Faith’s Phylogenetic Diversity (Faith 1992) based on a recent synthetic tree
82 with robust time calibration, covering about 6000 mammalian species (Upham et al. 2019). Faith’s PD
83 measures the sum of unique branches from an arbitrary root to a set of tips, and comparatively larger
84 values indicate a more phylogenetic diverse species pool. We measured phylogenetic diversity starting

85 from the root of the entire tree (and not from Chiroptera); this bears no consequences on the resulting
86 values, since all branches leading up to Chiroptera are only counted one per species pool, and (as we
87 explain when describing the assembly of the composite risk map), all individual risk components are
88 ranged in [0,1]. This measure incorporates a richness component, which we chose not to correct for; the
89 interpretation of the phylogenetic diversity is therefore a weighted species richness, that accounts for
90 phylogenetic over/under-dispersal in some places.

91 **Bats compositional uniqueness**

92 For every species pool, we measured its Local Contribution to Beta-Diversity (Legendre and De Cáceres
93 2013); LCBD works from a species-data matrix (traditionally noted as \mathbf{Y}), where species are rows and sites
94 are columns, and a value of 1 indicates occurrence. We extracted the \mathbf{Y} matrix assuming that every pixel
95 represents a unique location, and following best practices (Legendre and Condit 2019) transformed it
96 using Hellinger's distance to account for unequal bat richness at different pixels. The correction of raw
97 community data is particularly important for two reasons: first, it prevents the artifact of richer sites
98 having higher importance; second, it removes the effect of overall species richness, which is already
99 incorporated in the phylogenetic diversity component. High values of LCBD indicate that the pixel has a
100 community that is on average more dissimilar in species composition than what is expected knowing the
101 entire matrix, i.e. a more unique community. Recent results by Dansereau et al. (2022) shows that LCBD
102 measures are robust with regards to spatial scale, and are therefore applicable at the global scale.

103 **Viral sharing between hosts**

104 For all bat hosts of *Betacoronavirus*, we extracted their predicted viral sharing network, generated from a
105 previously published generalized additive mixed model of virus sharing by a tensor function of
106 phylogenetic distance and geographic range overlap (Albery et al. 2020). This network stores pairwise
107 values of viral community similarity. To project viral sharing values into a single value for every pixel, we
108 averaged the pairwise scores. High values of the average sharing propensity means that this specific extant
109 bat assemblage is likely to be proficient at exchanging viruses.

110 **Composite risk map**

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

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

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

128 **Viral phylogeography and evolutionary diversification**

129 To next represent phylogeography of betacoronaviruses in bats, we aggregated and analyzed
130 betacoronavirus sequence data. We used the following query to pull all *Betacoronavirus* sequence data
131 from the GenBank Nucleotide database except SARS-CoV-2; (“Betacoronavirus”[Organism] OR
132 betacoronavirus[All Fields]) NOT (“Severe acute respiratory syndrome coronavirus 2”[Organism] OR
133 sars-cov-2[All Fields]). We added a single representative sequence for SARS-CoV-2 and manually curated
134 to remove sequences without the RNA-dependent RNA polymerase (RdRp) sequence or that contained

135 words indicating recombinant or laboratory strains including “patent,” “mutant,” “GFP,” and
136 “recombinant.” We filtered over-represented taxa including betacoronavirus 1, hCoV-OC43, Middle East
137 respiratory syndrome coronavirus, Murine hepatitis virus, and hCoV-HKU1. Curated betacoronavirus
138 RdRp sequences were then aligned using MAFFT v 1.4.0 [Katoh and Standley (2013); Algorithm
139 FFT-NS-2, Scoring matrix 200PAM / k=2, gap open penalty 1.53m offset value 0.123] and a maximum
140 likelihood tree reconstructed in IQ-TREE v 1.6.12 (Nguyen et al. 2015) with ModelFinder
141 (Kalyaanamoorthy et al. 2017) ultrafast bootstrap approximation (Hoang et al. 2018) with a general time
142 reversible model with empirical base frequencies and the 5-discrete-rate-category FreeRaye model of
143 nucleotide substitution (GTR+F+R5).

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

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

153 Subsequently, we tested the hypothesis that the biogeography of bat betacoronaviruses should track the
154 biogeography of their hosts. To test this idea, we loosely adapted a method from (Kreft and Jetz 2007,
155 2010), who proposed a phylogenetic method for the delineation of animal biogeographic regions. In their
156 original method, a distance matrix - where each row or column represents a geographic raster’s grid cell,
157 and the dissimilarity values are the “beta diversity similarity” of their community assemble - undergoes
158 non-metric multidimensional scaling (NMDS); the first two axes of the NMDS are projected
159 geographically using a four-color bivariate map. Here, we build on this idea with an entirely novel
160 methodology. First, we measure the phylogenetic distance between the different viruses in the
161 betacoronaviruses tree by using the cophenetic function in ape (Paradis and Schliep 2019); subsequently,
162 we take a principal components analysis of that distance matrix (readily interchangeable for NMDS in this
163 case) to project the viral tree into an n-dimensional space. We then take the first two principal

164 components and, as with the evolutionary distinctiveness analysis, aggregated these to a mean host value
165 and projected them using a four-color bivariate map.

166 Results and discussion

167 Host richness does not predict virus distinctiveness

168 Bats, the second most diverse groups of mammals, are found worldwide and serve as the main animal
169 reservoir for different strains of betacoronaviruses (Drexler et al. 2014). This has attracted attention to
170 areas where high diversity of bats, and therefore presumably high diversity of betacoronaviruses, can be an
171 important issue for human health (Calisher et al. 2006, Moratelli and Calisher 2015). By overlaying the
172 IUCN range maps for confirmed bat hosts of betacoronaviruses [fig. 1; top], we see that the the main
173 hotspots (both in terms of size and higher values) of host richness are primarily South-Eastern Asia, parts
174 of Southern Europe, and to a lesser extent parts of Africa in the -25-0 range of latitudes. The description of
175 host richness is an important first step towards understanding risk, as previous research (Anthony et al.
176 2017, Mollentze and Streicker 2020) states that locally diverse bat communities could maintain more
177 viruses and hence, a higher probability of having a pathogen that could represent a risk for human health.

178 [Figure 1 about here.]

179 Nevertheless, locally diverse and virus-rich bat communities could represent an increased risk of spillover
180 under climate change through the creation of novel interactions (Carlson et al. 2022), and therefore the
181 diversity of *Betacoronavirus* strains should similarly be ccounted for. In fig. 1 (bottom), we contrast the
182 evolutionary distinctiveness of bats and viruses – this reveals a slightly different portrait than bat richness
183 alone. Chiropterans can be classified, from a macro-evolutionary standpoint, as Yangochiroptera and
184 Yinpterochiroptera elsewhere (Teeling et al. 2005, Springer 2013). Specifically, we would expect that the
185 so-called “New World” group of bats, being more evolutionary distinct, would also have evolutionary
186 distinct viruses. Indeed fig. 1 (bottom) reveals it to be the case, and this region harbors a distinct
187 bat-betacoronaviruses complex. This can be explained by the fact that Yangochiroptera, although not
188 limited to the western hemisphere, contain the highly diverse adaptive radiation in the Phyllostomidae
189 (Villalobos and Arita 2010), which is restricted to the western hemisphere. By contrast, South-Eastern
190 Asia has a lot of non-evolutionary distinct bats, who nevertheless hosted evolutionary-distinct viruses.

191 It is noteworthy that outside of South America, viral evolutionary distinctiveness does not accurately track
192 host diversity, with some areas having over-distinct viruses (eastern China but, oddly, not the rest of
193 southeast Asia). There are a number of likely explanations. First, given the richness of bats in southeast
194 Asia, many betacoronaviruses likely remain to be discovered in this region. Indeed, global predictions by
195 Becker et al. (2022) highlight that southeast Asia is a likely hotspot of unconfirmed hosts of
196 betacoronaviruses, which would likely result in additional viral discoveries. This idea is unsurprising
197 given the growing realization, especially since the emergence of SARS-CoV-2, that unique lineages of
198 similar viruses are widespread in bats but still mostly undescribed. The most distinct
199 bats-betacoronaviruses complex is found in South America, a region with a comparatively lower number
200 of hosts; this matches with the isolation through variance of the host group, and may highlight a different
201 co-evolutionary dynamic. Alternatively, this distinctiveness hotspot may be a product of under-sampling:
202 South-America is one of the places where the fewest *Betacoronavirus* sequences have been discovered
203 (Anthony et al. 2017, Olival et al. 2017, Allen et al. 2017), resulting in sparser phylogenetic tree, thereby
204 artificially inflating distinctiveness. Adding more viruses would bring the distinctiveness of known
205 sequences down. Finally, South-America is the range of Phyllostomidae, a group of bats that underwent
206 explosive diversification events (Villalobos and Arita 2010), which may drive the emergence of multiple
207 viral lineages.

208 **The phylogeographic regions of hosts and their viruses overlap**

209 Despite the difference in evolutionary distinctiveness globally, there are reasons to expect that the
210 phylogeography of bats and betacoronaviruses should show some degree of congruence (Letko et al. 2020,
211 Van Brussel and Holmes 2022). In particular, this should be the case if viruses can circulate among hosts
212 and co-evolve with local hosts communities, making their evolutionary process more than a byproduct of
213 host evolution. High density of hosts sharing the same virus (albeit possibly different strains) can drive or
214 result from evolution of the bat antiviral immune system, resulting in spatially distinct immunological
215 responses, as evidenced in several bat species (Banerjee et al. 2020). Immune characteristics that allow
216 bats to be better adapted to infection by emerging viruses (Gorbunova et al. 2020, Irving et al. 2021), in
217 addition to being hardcoded in their genome (Jebb et al. 2020), may be related to a wide variety of diets
218 (Banerjee et al. 2020, Moreno Santillán et al. 2021), themselves likely to be driven by spatial effects,
219 especially at the local scale – bats, indeed, occupy a variety of environments, and therefore display a

220 variety of adaptations to these environments (Muylaert et al. 2022).

221 [Figure 2 about here.]

222 In fig. 2, we show a projection of the phylogeographic signal of bats (top) and viruses (bottom) in space;
223 the distinct groupings (represented by different colors symbolizing positions in the subspace formed by
224 the first two axes of the PCoA) are essentially equivalent between the two groups, and can be coarsely
225 delineated as southeast Asia, Eurasia above a northing of 25, and Africa and south America. These results
226 suggest that, although the evolutionary distinctiveness of the bat-betacoronaviruses complex varies
227 spatially, the system shows an important degree of spatial consistency, with a reduced number of
228 bioregions. Available information describing the spillover of zoonotic betacoronaviruses of bat origin
229 where data was available before and up through the COVID-19 pandemic puts spillover events of
230 SARS-CoV-2 in Wuhan, China; SARS-CoV in Guangdong, China based on the presence of closest known
231 viruses circulating in nature, and a nearby location where serological (antibody) evidence has indicated
232 human exposure to SARS-like viruses (Wang et al. 2018); MERS-CoV in Saudi Arabia based on index cases
233 available from a recently-published compendium of cases (Ramshaw et al. 2019). For the latest event,
234 most if not all index cases are presumed to be camel-to-human transmission, and the precise origin point
235 (if it exists) of MERS-CoV in bats is uncertain. Recent recombinant canine coronavirus spillover events in
236 Haiti (Lednicky et al. 2021) and Europe (Vlasova et al. 2022) are not relevant here, as bats' involvement in
237 these cycles of transmission have been supposed to be non-existent. These index cases fall within different
238 phylogeographic bioregions (fig. 2), which further highlight the issue that different host-virus sub-systems
239 may lead to widespread emergence.

240 **Coevolution-informed spillover risk is different in space**

241 As host richness, joint distinctiveness, or phylogeographic structure suggest that the bat-betacoronaviruses
242 complex is globally fragmented enough to give rise to both different levels of risk (as evidenced by the
243 spatial location of spillover events) and different types of co-evolutionary dynamics, we turn to the
244 Geographic Mosaic Theory of Coevolution (Thompson 2005) to provide a measure of risk accounting for
245 multiple processes. In fig. 3, we overlapped three components of spillover risk: viral sharing, *i.e.* the
246 chance that two bats will share viruses overall; Local Contribution to Beta Diversity, *i.e.* the fact that a bat
247 community is compositionally unique compared to the average compositional similarity across the entire

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

252 [Figure 3 about here.]

253 From the perspective of spillover risk, the most important combination of factors is a high phylogenetic
254 diversity of hosts with low viral sharing; this, essentially, means that very different betacoronaviruses
255 could co-exist within the same place. This is particularly the case given that betacoronaviruses often
256 evolve and even achieve host shifts through recombination, which requires the co-occurrence of
257 sufficiently distinct viruses to be a major driver of emergence. In fig. 3, this corresponds to yellow to pale
258 green areas, which are essentially limited to South-Eastern Asia, and to some part of Sub-Saharan Africa.
259 Adopting a geographic mosaic theory perspective on risk, other regions of the world are of lesser concern
260 fig. 4. Our risk decomposition does not account for viral diversity or distinctiveness. The simple rationale
261 behind it is that the acquisition of viral data is rarely disconnected from the acquisition of host data. There
262 are more sources of information on hosts than on viruses, allowing to develop a host-centric perspective
263 on risk (although this estimate would be more accurate with viral traits related to *e.g.* ability to switch hosts
264 or pathogenic potential). Areas with high bat diversity and high turnover *may* facilitate the evolutionary
265 radiation of viruses, matching previous findings that the diversification of bat coronaviruses is driven
266 largely by host shifts (inter-genus or higher levels of cross-species transmission) and, to a lesser degree,
267 cospeciation and sharing (intra-genus cross-species transmission; Anthony et al. 2017). This
268 diversification is not an actual risk factor for spillover itself, but acts downstream of a spillover event by
269 increasing the random chance of the emergence of a virus with the raw genomic components required for
270 the potential to infect humans.

271 [Figure 4 about here.]

272 From another perspective, areas of high host uniqueness and virus sharing (red-to-pink) could provide
273 hotspots of *Betacoronavirus* risk through mixing of unique viruses (via codivergence) and in turn
274 recombination. Under our framework, such a hotspot was identified in Madagascar, where most bat
275 species are endemic following evolutionary divergence from sister species in both African and Asian

276 continents (e.g. Shi et al. 2014). Recent surveillance (Kettenburg et al. 2022) has identified a novel
277 *Betacoronavirus* (in the subgenus *Nobecovirus*) in Madagascar-endemic pteropid bat species (*Pteropus*
278 *rufus*, *Rousettus madagascariensis*), emphasizing strong proof of principle in model predictions.

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

280 Based on the previous result, we extracted the risk component from the composite map (see Methods), to
281 provide a single measure of risk varying between 0 and 1. This measure is presented in fig. 4. As this map
282 represents the potential risk, it must be weighed by the potential for contacts with humans. As a proxy for
283 this measure, we used the proportion of build/urban land from the EarthEnv dataset: this is a reasonable
284 proxy for the density of humans per unit area, which increases the probability of pathogen spread more
285 widely (Hazarie et al. 2021). Since human activity is required to amplify the frequency of virus encounters
286 and thus create areas of viral amplification, mapping the potential risk against measures of land use is
287 required to generate a more actionable assessment of risk. This map is presented in fig. 5. Most of South
288 America and Europe are at comparatively lower risk, as although densely populated, settlements tend to
289 be in areas with lower potential risk. Regions like Malaysia and the North coast of Australia have a high
290 risk component, but should represent a relatively lower effective risk due to low human density. However,
291 this mapping reveals that South-East Asia, the Indian subcontinent, and parts of sub-Saharan Africa, are
292 at high risk due to the overlap between built areas and bat communities representing more opportunities
293 for cross-species transmission of betacoronaviruses. In looking for the origins of SARS in China, Xu et al.
294 (2004) present serological evidence that strongest human-animal contact results in higher risk of virus
295 exposure, regardless of the animal species, but that different types of contact had different impacts. Ideally,
296 finer-grained information about human activity (rather than human presence through anthropisation)
297 could allow to partition this risk further, albeit at the cost of more hypotheses required to estimate the
298 amount of risk represented by each activity. Our map of risk overlays with recent results from Cohen et al.
299 (2022) – areas of purported high risk/diversitification potential (Madagascar, South-America) overlay with
300 sampling gaps for *Betacoronavirus*, stressing the need for spatially targeted monitoring and discovery.

301

[Figure 5 about here.]

302 **Conclusion**

303 Our study focuses largely on the biogeography of hosts. Specifically, we identify the Amazon and
304 South-Eastern Asia as hotspots where the phylogenetic distinctiveness of *Betacoronavirus* is the highest
305 (Anthony et al. 2017); surprisingly, current data suggest that viral sharing between hosts is high in the
306 Amazon and low in South-Eastern Asia, which has the potential to result in different evolutionary
307 dynamics between these two regions, hinting at different futures for their viral communities. This work is
308 important both as a description of the bats-betacoronaviruses complex, but also because more broadly,
309 bats are known reservoirs for a variety of emerging viruses and other pathogens (Calisher et al. 2006,
310 Melaun et al. 2014), making balancing the needs for bat conservation and disease prevention most likely
311 very difficult and a source of human-wildlife conflicts, especially in more densely populated areas (Stone
312 et al. 2015, Rego et al. 2015).

313 Yet, we know that viruses with high host plasticity, that is, the ability of a given virus to adapt to various
314 taxonomic orders and ecological groups (Kreuder Johnson et al. 2015), are more likely to amplify viral
315 spillover, followed by secondary human-to-human transmission, and geographical spread (Hazarie et al.
316 2021). High viral host plasticity is an especially important trait for RNA viruses like betacoronaviruses
317 (Haddad et al. 2021). Indeed, our analysis of viral sequences reveals that Latin America is a hotspot of
318 viral distinctiveness, suggesting that this part of the bats-betacoronaviruses complex may be undergoing
319 independent evolutionary dynamics (related species sharing viruses that are different from the rest of the
320 global pool). The other hotspot of viral distinctiveness is S.E. Asia, in which richness is high but sharing is
321 low; this suggests a different type of evolutionary dynamics (unrelated viruses coevolving with
322 evolutionarily distinct hosts, generating high diversity locally, see e.g. Latinne et al. 2020). Both of these
323 areas should be priority areas for sampling, especially since Becker et al. (2022) advance that they harbor
324 undiscovered hosts of beta-coronaviruses. This diversity of hosts, and the mechanisms by which the
325 exchange of viruses occurs between species, is largely affected by the local environmental conditions and
326 environmental change.

327 Bats are important reservoir hosts for different classes of microorganisms, many of which a threat to
328 human health (Letko et al. 2020, Van Brussel and Holmes 2022). Chiropterans emerged around 64 million
329 years ago and are one of the most diverse mammalian orders, with an estimated richness of more than
330 1400 species (Peixoto et al. 2018, Simmons and Cirranello 2020). They exhibit a broad variety of habitat

331 use, behaviour, and feeding strategies, putting them at key positions in the delivery and provisioning of
332 several ecosystem services, tied to important ecosystem-derived benefits to human (Kasso and
333 Balakrishnan 2013). For example, bats are an essential component of many seed-dispersal networks
334 (Mello et al. 2011). Over two-thirds of bats are known to be either obligate or facultative insectivores,
335 therefore actively contributing for agricultural pest control (Williams-Guillén et al. 2008, Voigt and
336 Kingston 2016), and vectors of pathogens that put a risk on human health (Gonsalves et al. 2013a, b).
337 Because bats are globally distributed and have a long evolutionary history, phylogeographic and
338 biogeographic approaches are required to shed light on the contemporary distribution of coevolutionary
339 processes between bats and the pathogens they host. Not all areas in which bats, viruses, and human are
340 co-occurring are facing a risk of spillover towards human populations, and the areas in which this risk exist
341 may not be facing risks of the same nature and magnitude.

342 There are several factors that drive changes in the diversity of bats (Alves et al. 2018), but human
343 activities' effects on the ecosystem (like modifications of land use) could significantly decrease it.
344 Therefore, it can be suggested that changes in the diversity of betacovs in bats are linked to their
345 biogeographic variation, and human population density and other anthropogenic factors are decisive
346 moderators for its implications in public health. With the increase of contact between humans and
347 potential hosts, we also increase the risk of emergence of novel diseases (Johnson et al. 2020), as previous
348 studies on RNA viruses suggest the importance of host phylogeography at the time of virus dispersal
349 (Gryseels et al. 2017). One of these scenarios where interaction between bats and humans can occur can
350 be seed dispersal in tropical agroecosystems. It opens the discussion of whether the fruits thrown by bats
351 not only disperse seeds but could also be a source of indirect interaction between viruses of bat origin and
352 humans (Deshpande et al. 2022). This represents a challenge for conservation strategies and disease
353 ecology since some areas can have both potential for the acquisition of zoonotic viruses and bat-human
354 interactions; in particular, the challenge lies in the fact that actual exposure must then be quantified
355 accounting for several transmission scenarios, including both direct and indirect bat - human interaction.

356 **Acknowledgements:** We acknowledge that this study was conducted on land within the traditional
357 unceded territory of the Saint Lawrence Iroquoian, Anishinabewaki, Mohawk, Huron-Wendat, and
358 Omàmiwininiwak nations. This work was supported by funding to the Viral Emergence Research
359 Initiative (VERENA) consortium including NSF BII 2021909 and a grant from Institut de Valorisation des
360 Données (IVADO). This research was enabled in part by support provided by Calcul Québec

361 (www.calculquebec.ca) and Compute Canada (www.computecanada.ca). NF is funded by the NSERC
362 BIOS² CREATE program. TP and NF are funded by the Courtois Foundation. RLM was supported by
363 Bryce Carmine and Anne Carmine (née Percival), through the Massey University Foundation.

364 **References**

- 365 Agosta, S. J. et al. 2010. How specialists can be generalists: Resolving the “parasite paradox” and
366 implications for emerging infectious disease. - *Zoologia (Curitiba)* 27: 151–162.
- 367 Albery, G. F. et al. 2020. Predicting the global mammalian viral sharing network using phylogeography. -
368 *Nature Communications* 11: 2260.
- 369 Albery, G. F. et al. 2022. Urban-adapted mammal species have more known pathogens. - *Nature Ecology*
370 & Evolution
- 371 Allen, T. et al. 2017. Global hotspots and correlates of emerging zoonotic diseases. - *Nature*
372 *Communications in press*.
- 373 Alves, D. M. C. C. et al. 2018. Geographic variation in the relationship between large-scale environmental
374 determinants and bat species richness. - *Basic and Applied Ecology* 27: 1–8.
- 375 Anthony, S. J. et al. 2017. Global patterns in coronavirus diversity. - *Virus Evolution* in press.
- 376 Banerjee, A. et al. 2020. Novel Insights Into Immune Systems of Bats. - *Frontiers in Immunology* 11: 26.
- 377 Becker, D. J. et al. 2020. Beyond Infection: Integrating Competence into Reservoir Host Prediction. -
378 *Trends in Ecology & Evolution* 35: 1062–1065.
- 379 Becker, D. J. et al. 2022. Optimising predictive models to prioritise viral discovery in zoonotic reservoirs. -
380 *The Lancet Microbe* in press.
- 381 Calisher, C. H. et al. 2006. Bats: Important Reservoir Hosts of Emerging Viruses. - *Clinical Microbiology*
382 *Reviews* 19: 531–545.
- 383 Carlson, C. J. et al. 2022. Climate change increases cross-species viral transmission risk. - *Nature*: 1–1.
- 384 Cavender-Bares, J. et al. 2009. The merging of community ecology and phylogenetic biology. - *Ecol. Lett.*
385 12: 693–715.

- 386 Cohen, L. E. et al. 2022. Sampling strategies and pre-pandemic surveillance gaps for bat coronaviruses.:
387 2022.06.15.496296.
- 388 Dansereau, G. et al. 2022. Evaluating ecological uniqueness over broad spatial extents using species
389 distribution modelling. - Oikos n/a: e09063.
- 390 Deshpande, K. et al. 2022. Forbidden fruits? Ecosystem services from seed dispersal by fruit bats in the
391 context of latent zoonotic risk. - Oikos (Copenhagen, Denmark): oik.08359.
- 392 Drexler, J. F. et al. 2014. Ecology, evolution and classification of bat coronaviruses in the aftermath of
393 SARS. - Antiviral Research 101: 45–56.
- 394 Engering, A. et al. 2013. Pathogen–host–environment interplay and disease emergence. - Emerging
395 Microbes & Infections 2: e5.
- 396 Faith, D. P. 1992. Conservation evaluation and phylogenetic diversity. - Biological Conservation 61: 1–10.
- 397 Gomulkiewicz, R. et al. 2000. Hot Spots, Cold Spots, and the Geographic Mosaic Theory of Coevolution. -
398 The American Naturalist 156: 156–174.
- 399 Gomulkiewicz, R. et al. 2007. Dos and don'ts of testing the geographic mosaic theory of coevolution. -
400 Heredity 98: 249–258.
- 401 Gonsalves, L. et al. 2013a. Do mosquitoes influence bat activity in coastal habitats? - Wildlife Research 40:
402 10–24.
- 403 Gonsalves, L. et al. 2013b. Mosquito Consumption by Insectivorous Bats: Does Size Matter? - PLOS ONE
404 8: e77183.
- 405 Gorbunova, V. et al. 2020. The World Goes Bats: Living Longer and Tolerating Viruses. - Cell Metabolism
406 32: 31–43.
- 407 Gryseels, S. et al. 2017. When Viruses Don't Go Viral: The Importance of Host Phylogeographic Structure
408 in the Spatial Spread of Arenaviruses (JH Kuhn, Ed.). - PLOS Pathogens 13: e1006073.
- 409 Haddad, D. et al. 2021. SARS-CoV-2: Possible recombination and emergence of potentially more virulent
410 strains (H Attoui, Ed.). - PLOS ONE 16: e0251368.
- 411 Hazarie, S. et al. 2021. Interplay between population density and mobility in determining the spread of
412 epidemics in cities. - Communications Physics 4: 191.

- 413 Hoang, D. T. et al. 2018. UFBoot2: Improving the Ultrafast Bootstrap Approximation. - Molecular Biology
414 and Evolution 35: 518–522.
- 415 Hochberg, M. E. and Baalen, M. 1998. Antagonistic coevolution over productivity gradients. - The
416 American Naturalist 152: 620–634.
- 417 Irving, A. T. et al. 2021. Lessons from the host defences of bats, a unique viral reservoir. - Nature 589:
418 363–370.
- 419 Isaac, N. J. B. et al. 2007. Mammals on the EDGE: Conservation Priorities Based on Threat and Phylogeny.
420 - PLOS ONE 2: e296.
- 421 IUCN 2021. The IUCN Red List of Threatened Species.
- 422 Janzen, D. H. 1980. When is it Coevolution? - Evolution 34: 611–612.
- 423 Jebb, D. et al. 2020. Six reference-quality genomes reveal evolution of bat adaptations. - Nature 583:
424 578–584.
- 425 Johnson, C. K. et al. 2020. Global shifts in mammalian population trends reveal key predictors of virus
426 spillover risk. - Proceedings of the Royal Society B: Biological Sciences 287: 20192736.
- 427 Kalyaanamoorthy, S. et al. 2017. ModelFinder: Fast model selection for accurate phylogenetic estimates. -
428 Nature Methods 14: 587–589.
- 429 Kasso, M. and Balakrishnan, M. 2013. Ecological and Economic Importance of Bats (Order Chiroptera). -
430 ISRN Biodiversity 2013: e187415.
- 431 Katoh, K. and Standley, D. M. 2013. MAFFT Multiple Sequence Alignment Software Version 7:
432 Improvements in Performance and Usability. - Molecular Biology and Evolution 30: 772–780.
- 433 Keke, S. et al. 2010. Study on skin color image segmentation used by Fuzzy-c-means arithmetic. - 2010
434 Seventh International Conference on Fuzzy Systems and Knowledge Discovery 2: 612–615.
- 435 Kettenburg, G. et al. 2022. Full Genome Nobecovirus Sequences From Malagasy Fruit Bats Define a
436 Unique Evolutionary History for This Coronavirus Clade. - Frontiers in Public Health in press.
- 437 Kreft, H. and Jetz, W. 2007. Global patterns and determinants of vascular plant diversity. - Proceedings of
438 the National Academy of Sciences 104: 5925–5930.

- 439 Kreft, H. and Jetz, W. 2010. A framework for delineating biogeographical regions based on species
440 distributions. - *Journal of Biogeography* 37: 2029–2053.
- 441 Kreuder Johnson, C. et al. 2015. Spillover and pandemic properties of zoonotic viruses with high host
442 plasticity. - *Scientific Reports* 5: 14830.
- 443 Laine, A.-L. 2005. Spatial scale of local adaptation in a plant-pathogen metapopulation. - *Journal of*
444 *Evolutionary Biology* 18: 930–938.
- 445 Latinne, A. et al. 2020. Origin and cross-species transmission of bat coronaviruses in China. - *bioRxiv*:
446 The Preprint Server for Biology: 2020.05.31.116061.
- 447 Lednicky, J. A. et al. 2021. Isolation of a Novel Recombinant Canine Coronavirus from a Visitor to Haiti:
448 Further Evidence of Transmission of Coronaviruses of Zoonotic Origin to Humans. - *Clinical*
449 *Infectious Diseases: An Official Publication of the Infectious Diseases Society of America*: ciab924.
- 450 Legendre, P. and De Cáceres, M. 2013. Beta diversity as the variance of community data: Dissimilarity
451 coefficients and partitioning (H Morlon, Ed.). - *Ecology Letters* 16: 951–963.
- 452 Legendre, P. and Condit, R. 2019. Spatial and temporal analysis of beta diversity in the Barro Colorado
453 Island forest dynamics plot, Panama. - *Forest Ecosystems* 6: 7.
- 454 Letko, M. et al. 2020. Bat-borne virus diversity, spillover and emergence. - *Nature Reviews Microbiology*
455 18: 461–471.
- 456 MacLean, O. A. et al. 2021. Natural selection in the evolution of SARS-CoV-2 in bats created a generalist
457 virus and highly capable human pathogen. - *PLOS Biology* 19: e3001115.
- 458 Melaun, C. et al. 2014. Bats as Potential Reservoir Hosts for Vector-Borne Diseases. - In: Kliment, S. and
459 Mehlhorn, H. (eds), *Bats (Chiroptera) as Vectors of Diseases and Parasites: Facts and Myths*.
460 *Parasitology Research Monographs*. Springer, pp. 25–61.
- 461 Mello, M. A. R. et al. 2011. The Missing Part of Seed Dispersal Networks: Structure and Robustness of
462 Bat-Fruit Interactions. - *PLOS ONE* 6: e17395.
- 463 Mollentze, N. and Streicker, D. G. 2020. Viral zoonotic risk is homogenous among taxonomic orders of
464 mammalian and avian reservoir hosts. - *Proceedings of the National Academy of Sciences* 117:
465 9423–9430.

- 466 Moratelli, R. and Calisher, C. H. 2015. Bats and zoonotic viruses: Can we confidently link bats with
467 emerging deadly viruses? - *Memórias do Instituto Oswaldo Cruz* 110: 1–22.
- 468 Moreno Santillán, D. D. et al. 2021. Large-scale genome sampling reveals unique immunity and metabolic
469 adaptations in bats. - *Molecular Ecology*: mec.16027.
- 470 Mull, N. et al. 2022. Virus isolation data improve host predictions for New World rodent
471 orthohantaviruses. - *Journal of Animal Ecology* in press.
- 472 Muylaert, R. L. et al. 2022. Present and future distribution of bat hosts of sarbecoviruses: Implications for
473 conservation and public health. - *Proceedings of the Royal Society B: Biological Sciences* 289:
474 20220397.
- 475 Nguyen, L.-T. et al. 2015. IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating
476 Maximum-Likelihood Phylogenies. - *Molecular Biology and Evolution* 32: 268–274.
- 477 Nuismer, S. L. et al. 2003. Coevolution between hosts and parasites with partially overlapping geographic
478 ranges. - *Journal of Evolutionary Biology* 16: 1337–1345.
- 479 Olival, K. J. et al. 2017. Host and viral traits predict zoonotic spillover from mammals. - *Nature* 546:
480 646–650.
- 481 Paradis, E. and Schliep, K. 2019. Ape 5.0: An environment for modern phylogenetics and evolutionary
482 analyses in R. - *Bioinformatics* 35: 526–528.
- 483 Parratt, S. R. et al. 2016. Infectious Disease Dynamics in Heterogeneous Landscapes. - *Annual Review of
484 Ecology, Evolution, and Systematics* 47: 283–306.
- 485 Peixoto, F. P. et al. 2018. A synthesis of ecological and evolutionary determinants of bat diversity across
486 spatial scales. - *BMC ecology* 18: 18.
- 487 Plowright, R. K. et al. 2015. Ecological dynamics of emerging bat virus spillover. - *Proceedings of the
488 Royal Society B: Biological Sciences* 282: 20142124.
- 489 Plowright, R. K. et al. 2017. Pathways to zoonotic spillover. - *Nature Reviews Microbiology* 15: 502–510.
- 490 Power, A. G. and Mitchell, C. E. 2004. Pathogen Spillover in Disease Epidemics. - *The American
491 Naturalist* 164: S79–S89.
- 492 Price, P. W. 2002. Macroevolutionary Theory on Macroecological Patterns. - Cambridge University Press.

- 493 Ramshaw, R. E. et al. 2019. A database of geopositioned Middle East Respiratory Syndrome Coronavirus
494 occurrences. - *Scientific Data* 6: 318.
- 495 Rego, K. M. da C. et al. 2015. Assessing human-bat interactions around a protected area in northeastern
496 Brazil. - *Journal of Ethnobiology and Ethnomedicine* 11: 80.
- 497 Rouault, E. et al. 2022. GDAL/OGR Geospatial Data Abstraction software Library. - Zenodo.
- 498 Ruiz-Aravena, M. et al. 2022. Ecology, evolution and spillover of coronaviruses from bats. - *Nature*
499 *Reviews Microbiology* 20: 299–314.
- 500 Seekell, D. A. et al. 2018. A geography of lake carbon cycling. - *Limnology and Oceanography Letters* 3:
501 49–56.
- 502 Shi, J. J. et al. 2014. A Deep Divergence Time between Sister Species of Eidolon (Pteropodidae) with
503 Evidence for Widespread Panmixia. - *Acta Chiropterologica* 16: 279–292.
- 504 Simmons, N. B. and Cirranello, A. L. 2020. Bat Species of the World: A taxonomic and geographic
505 database.
- 506 Springer, M. S. 2013. Phylogenetics: Bats United, Microbats Divided. - *Current Biology* 23: R999–R1001.
- 507 Stone, E. et al. 2015. Managing Conflict between Bats and Humans: The Response of Soprano Pipistrelles
508 (*Pipistrellus pygmaeus*) to Exclusion from Roosts in Houses. - *PLoS ONE* 10: e0131825.
- 509 Teeling, E. C. et al. 2005. A Molecular Phylogeny for Bats Illuminates Biogeography and the Fossil Record.
510 - *Science (New York, N.Y.)* 307: 580–584.
- 511 Thompson, J. N. 1994. The Coevolutionary Process. - University of Chicago Press.
- 512 Thompson, J. N. 2005. The Geographic Mosaic of Coevolution. - University Of Chicago Press.
- 513 Thrall, P. H. et al. 2007. Coevolution of symbiotic mutualists and parasites in a community context. -
514 *Trends in Ecology & Evolution* 22: 120–126.
- 515 Turner, W. C. et al. 2021. The roles of environmental variation and parasite survival in
516 virulence–transmission relationships. - *Royal Society open science* 8: 210088.
- 517 Upham, N. S. et al. 2019. Inferring the mammal tree: Species-level sets of phylogenies for questions in
518 ecology, evolution, and conservation. - *PLOS Biology* 17: e3000494.

- 519 Van Brussel, K. and Holmes, E. C. 2022. Zoonotic disease and virome diversity in bats. - Current Opinion
520 in Virology 52: 192–202.
- 521 Villalobos, F. and Arita, H. T. 2010. The diversity field of New World leaf-nosed bats (Phyllostomidae). -
522 Global Ecology and Biogeography 19: 200–211.
- 523 Vlasova, A. N. et al. 2022. Animal alphacoronaviruses found in human patients with acute respiratory
524 illness in different countries. - Emerging Microbes & Infections 11: 699–702.
- 525 2016. Bats in the Anthropocene: Conservation of Bats in a Changing World (CC Voigt and T Kingston,
526 Eds.). - Springer International Publishing.
- 527 Wang, N. et al. 2018. Serological Evidence of Bat SARS-Related Coronavirus Infection in Humans, China.
528 - Virologica Sinica 33: 104–107.
- 529 Williams-Guillén, K. et al. 2008. Bats Limit Insects in a Neotropical Agroforestry System. - Science 320:
530 70–70.
- 531 Xu, R.-H. et al. 2004. Epidemiologic Clues to SARS Origin in China. - Emerging Infectious Diseases 10:
532 1030–1037.

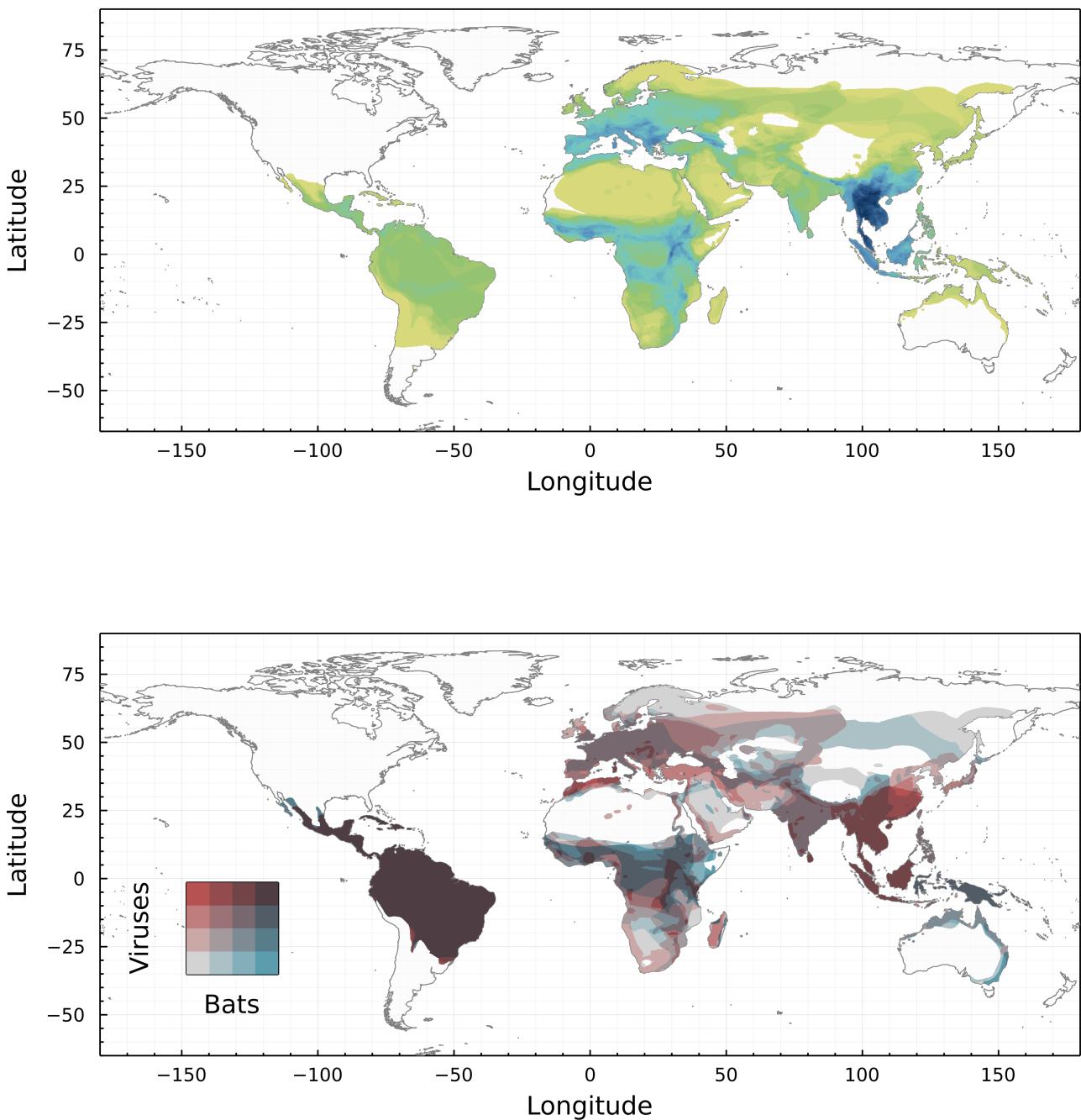


Figure 1: Top panel: relative diversity of known bat hosts of betacoronaviruses. This map shows that the region with the largest number of possible hosts is South-Eastern Asia. Bottom panel: congruence between the evolutionary distinctiveness of the hosts (grey to blue) and the viruses (grey to red). 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.

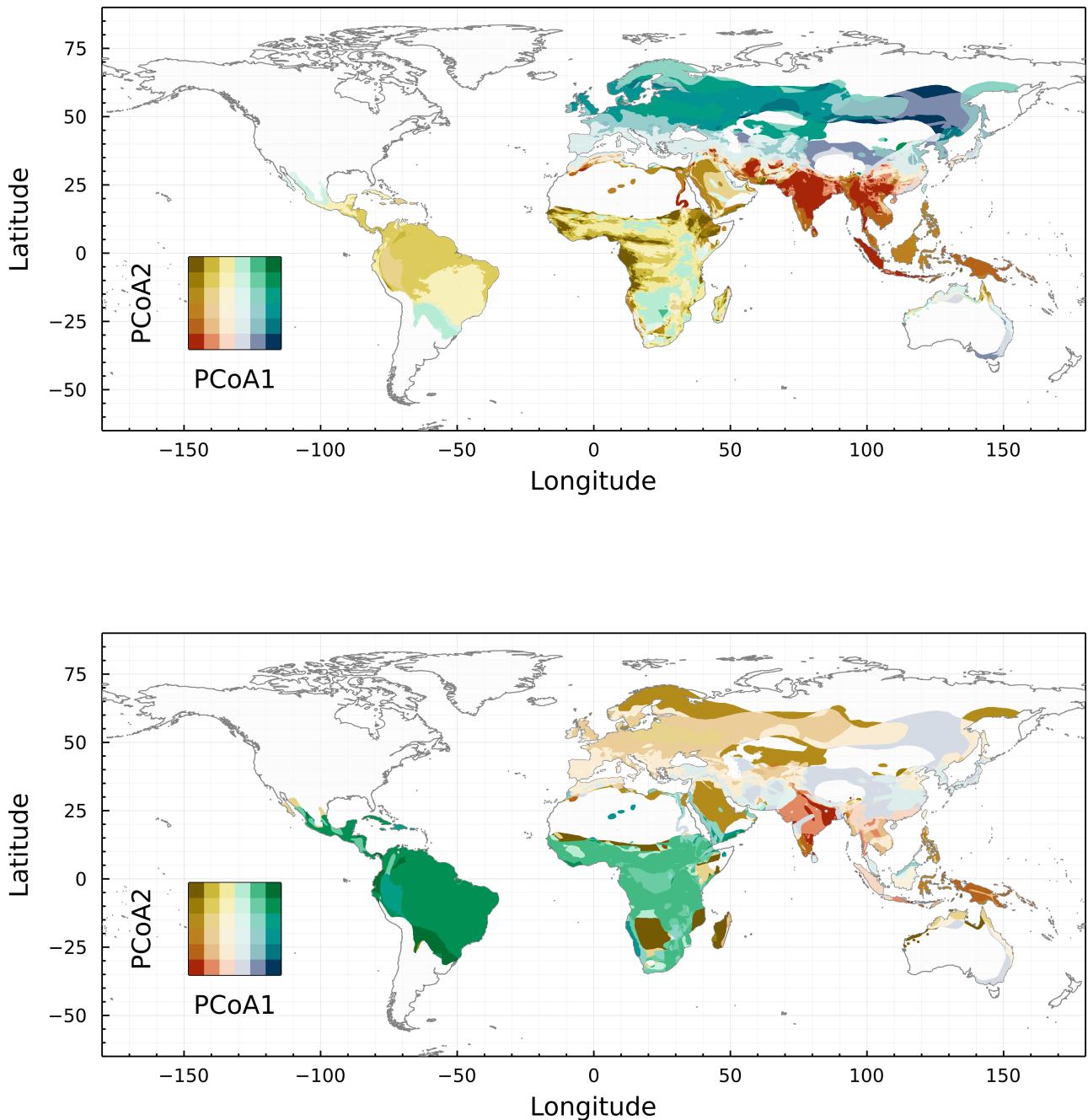


Figure 2: Phylogeographic regions of bats (top) and viruses (bottom) based on the joint analysis of their occurrence and phylogenetic relatedness. The different colors show tendencies to separate alongside the first two components of a PCoA. Note that the PCoA for the bats and viruses are independent,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

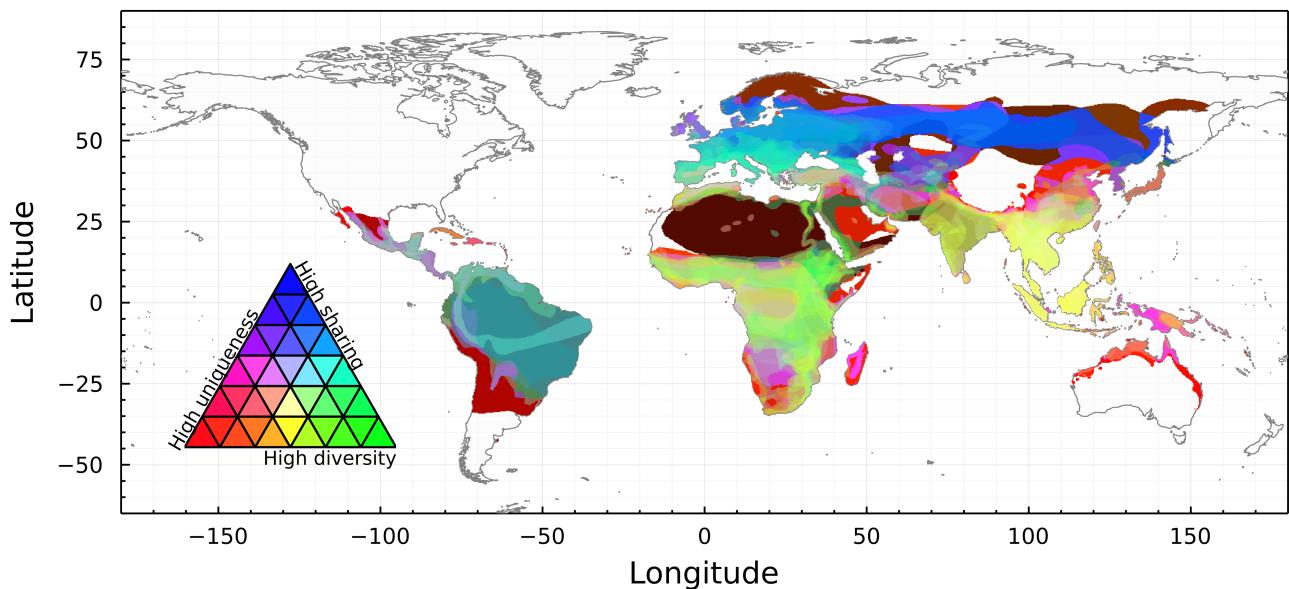


Figure 3: Trivariate additive mapping of the components of risk in the red/green/blue, where high virus sharing is encoded in the blue channel, host phylogenetic diversity in the green channel, and compositional uniqueness in the red channel. A pixel that would maximize all measures (highest possible 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.

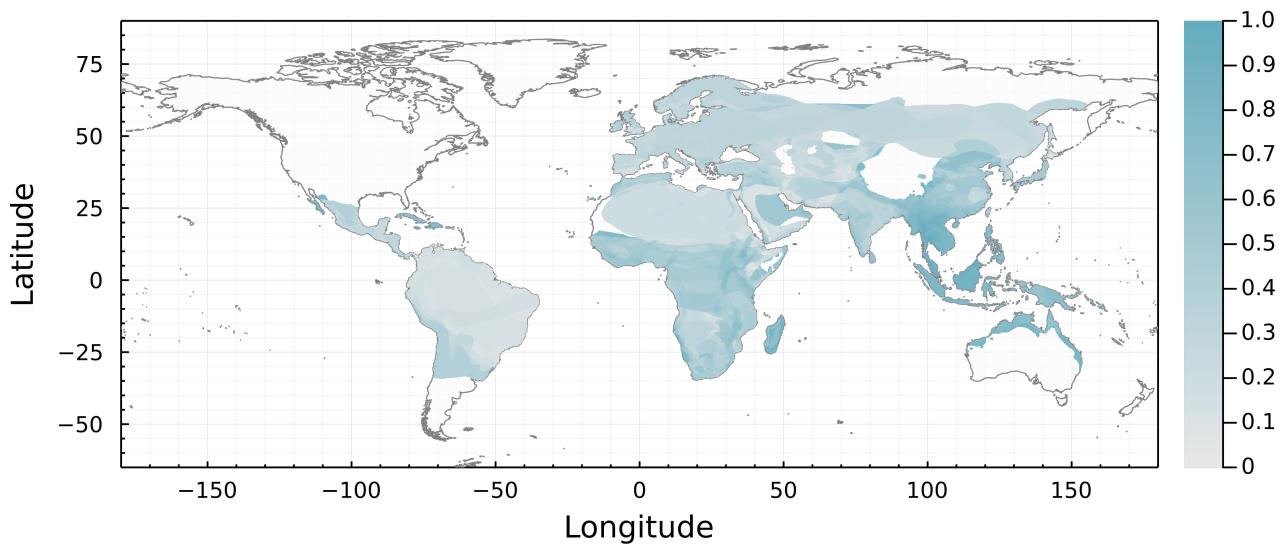


Figure 4: Extraction of a measure of *Betacoronavirus* spillover risk from bat hosts based on the colorimetric space from fig. 3. The risk is a composite measure of the color value and angular distance to the yellow hue, as defined in the methods, ranged in the unit space. Based on this analyses, regions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

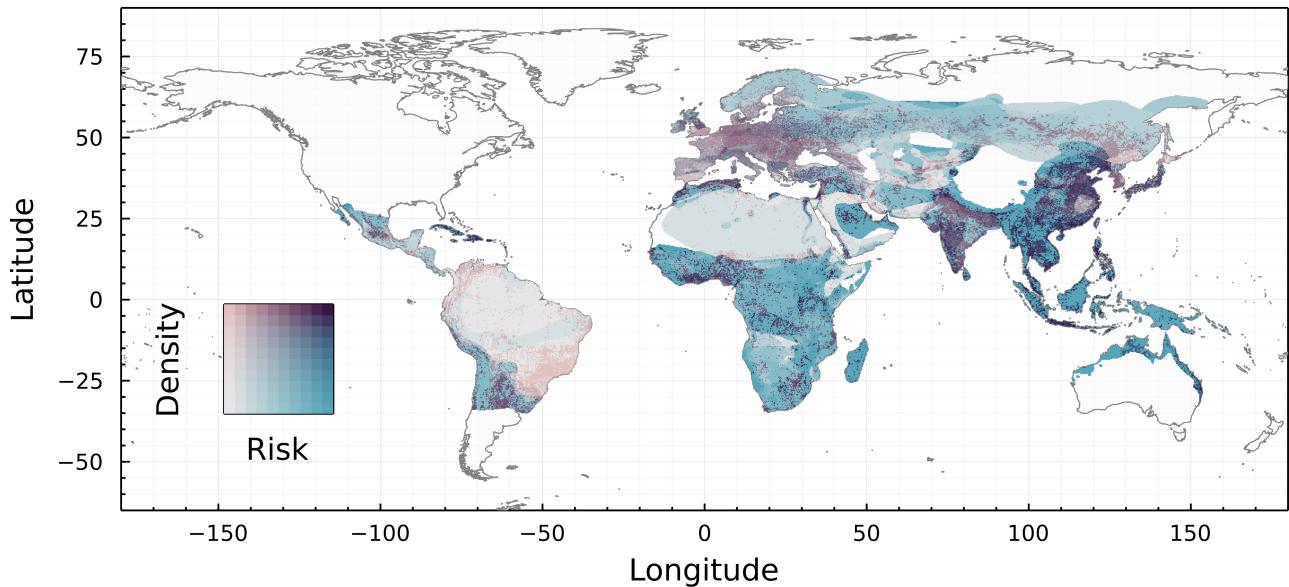


Figure 5: Overlap of the percent of each pixel occupied by urbanized structures, representing the degree of settlement, on the spillover risk map (where the risk comes only from wildlife, and ignores multi-hosts chains of transmissions including non-bats hosts). Darker pixels correspond to more risk, 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.