

1 Coronavirus surveillance in Congo basin wildlife detects RNA of
2 multiple species circulating in bats and rodents

3
4 Charles Kumakamba¹, Fabien R. Niama², Francisca Muyembe¹, Jean-Vivien Mombouli², Placide
5 Mbala Kingebeni¹, Rock Aime Nina³, Ipos Ngay Lukusa¹, Gerard Bounga⁴, Frida N’Kawa¹,
6 Cynthia Goma Nkoua², Joseph Atibu Losoma¹, Prime Mulembakani¹, Maria Makuwa^{1,5}, Ubald
7 Tamufe⁶, Amethyst Gillis^{7,8}, Matthew LeBreton⁹, Sarah H. Olson⁴, Kenneth Cameron^{4,10}, Patricia
8 Reed⁴, Alain Ondzie⁴, Alex Tremeau-Bravard¹¹, Brett R. Smith¹¹, Jasmine Pante¹¹, Bradley S.
9 Schneider^{7,12,13}, David J. McIver^{14,15}, James A. Ayukekbong^{14,16}, Nicole A. Hoff¹⁵, Anne W.
10 Rimoin¹⁵, Anne Laudisoit¹⁷, Corina Monagin^{7,11}, Tracey Goldstein¹¹, Damien O. Joly^{4,14,18}, Karen
11 Saylor^{5,7}, Nathan D. Wolfe⁷, Edward M. Rubin⁷, Romain Bagamboula MPassi¹⁹, Jean J.
12 Muyembe Tamfum²⁰, Christian E. Lange^{5,14}

13
14 **1** Metabiota Inc, Kinshasa, Democratic Republic of the Congo, **2** National Laboratory of Public
15 Health, Brazzaville, Republic of the Congo, **3** Ministry of Agriculture and Livestock, Brazzaville,
16 Republic of the Congo, **4** Wildlife Conversation Society, Bronx NY, USA, **5** Labyrinth Global
17 Health St. Petersburg, USA, **6** Metabiota Cameroon Ltd, Yaoundé, Cameroon, **7** Metabiota Inc,
18 San Francisco, USA, **8** Development Alternatives, Inc., Washington DC, USA, **9** Mosaic,
19 Yaoundé, Cameroon, **10** Unites States Fish and Wildlife Service, Bailey’s Crossroads VA, USA,
20 **11** One Health Institute, School of Veterinary Medicine, University of California, Davis CA, USA,
21 **12** Etiologic, Oakland, USA, **13** Pinpoint Science, San Francisco, USA, **14** Metabiota Inc,
22 Nanaimo, Canada, **15** University of California, Los Angeles, USA, **16** Epitech Consulting,
23 Nanaimo, Canada, **17** EcoHealth Alliance, New York, USA, **18** British Columbia Ministry of
24 Environment and Climate Change Strategy, Victoria, Canada, **19** Ministry of National Defense,

25 Republic of Congo, **20** Institut National de Recherche Biomédicale, Kinshasa, Democratic
26 Republic of the Congo

27

28 **Abstract**

29 Coronaviruses play an important role as pathogens of humans and animals, and the emergence
30 of epidemics like SARS, MERS and COVID-19 is closely linked to zoonotic transmission events
31 primarily from wild animals. Bats have been found to be an important source of coronaviruses
32 with some of them having the potential to infect humans, with other animals serving as
33 intermediate or alternate hosts or reservoirs. Host diversity may be an important contributor to
34 viral diversity and thus the potential for zoonotic events. To date, limited research has been done
35 in Africa on this topic, in particular in the Congo Basin despite frequent contact between humans
36 and wildlife in this region. We sampled and, using consensus coronavirus PCR-primers, tested
37 3,561 wild animals for coronavirus RNA. The focus was on bats (38%), rodents (38%), and
38 primates (23%) that posed an elevated risk for contact with people, and we found coronavirus
39 RNA in 121 animals, of which all but two were bats. Depending on the taxonomic family, bats
40 were significantly more likely to be coronavirus RNA-positive when sampled either in the wet
41 (*Pteropodidae* and *Rhinolophidae*) or dry season (*Hipposideridae*, *Miniopteridae*, *Molossidae*,
42 and *Vespertilionidae*). The detected RNA sequences correspond to 15 Alpha- and 6 Beta-
43 coronaviruses, with some of them being very similar (>95% nucleotide identities) to known
44 coronaviruses and others being more unique and potentially representing novel viruses. In seven
45 of the bats, we detected RNA most closely related to sequences of the human common cold
46 coronaviruses 229E or NL63 (>80% nucleotide identities). The findings highlight the potential for
47 coronavirus spillover, especially in regions with a high diversity of bats and close human contact,
48 and reinforces the need for ongoing surveillance.

49

50 **Introduction**

51 Coronaviruses are relatively large enveloped viruses with a single-stranded positive-sense RNA
52 genome of 26-32 kilobases that form their own taxonomic family within the *Nidovirales* order of
53 viruses [1]. There are two *Coronaviridae* subfamilies, *Letovirinae* and *Orthocoronavirinae*, and the
54 latter contains the genera *Alpha-* and *Betacoronavirus*, with viruses infecting mammalian species
55 as well as the genera *Gamma-* and *Deltacoronavirus* that primarily contain viruses found in birds
56 [2]. Although known for decades as important enteric and respiratory pathogens in domestic
57 animals, and as causative agent of mild respiratory infections in humans, it was only the
58 emergence of severe acute respiratory syndrome coronavirus (SARS-CoV) in humans in 2002
59 that brought coronaviruses broader attention [3]. The emergence and sporadic re-emergence of
60 Middle East respiratory syndrome coronavirus (MERS-CoV) since 2012 and the global COVID-
61 19 pandemic caused by SARS-CoV-2 have highlighted the enormous importance of this viral
62 family in the context of global public health [4-6].

63 Coronaviruses identical or closely related to SARS-CoV-1, MERS-CoV and SARS-CoV-2 have
64 been found in civets, camels, and bats, supporting zoonotic events as the most likely source of
65 the respective outbreaks in humans [5, 7-11]. Studies to identify the origin of these zoonotic
66 viruses also led to the discovery of many other, related or completely novel, animal coronaviruses
67 in the process; in particular they have detected an astonishing diversity of alpha- and beta
68 coronaviruses in bats, including relatives of coronaviruses previously identified in other hosts [12-
69 15]. This led to the hypothesis that bats are a reservoir for coronaviruses and that these viruses
70 are crossing into other non-bat species on a somewhat regular basis. As a result, they may
71 establish a novel permanent virus-host relationship, as in the case of MERS-CoV and camels, or
72 a transient relationship as in the case of SARS-CoV-1 and civets. However most interspecies
73 transmissions are likely dead ends for the virus and remain undetected [15-17]. The human
74 common cold viruses HCoV-229E and HCoV-NL63 are most likely animal origin viruses that
75 succeeded in establishing a permanent relationship with humans after crossing species barriers

76 directly or indirectly from bats [12, 13, 18, 19]. Coronaviruses OC43 and HKU-1, which also cause
77 common cold in humans, are likewise expected to have originated in animals, though likely in
78 rodents rather than bats [6, 20].

79 In sum, there is thus mounting biological evidence that spillover has happened repeatedly in the
80 past, continues to happen today, and will likely continue to occur in the future. Hence it is important
81 to study animal coronaviruses to characterize the risks posed by these potentially emerging
82 viruses, to understand the dynamics of the emergence of these pathogens, and to make informed
83 decisions concerning prevention and risk mitigation [15, 21].

84 However, the virus' biology is only one piece of the puzzle. We know that the emergence and
85 epidemic spread of SARS-CoV-1 and likely SARS-COV-2 are linked to human behavioral factors,
86 such as close contact with wild animals, and with factors such as biodiversity and wildlife
87 abundance, important prerequisites for virus diversity. Hotspots for zoonotic disease emergence
88 generally exist where humans are actively encroaching on such animal habitats [22, 23], as is
89 happening in Southeast Asia and Central Africa. While potential sources of zoonotic
90 coronaviruses are increasingly being explored, a great deal remains to be documented in most
91 parts of the biodiverse African continent, especially in Central Africa. Findings from countries such
92 as Kenya, Madagascar, Rwanda, South Africa and others suggest there are many coronaviruses
93 circulating, primarily in bats, including species related to pathogens such as SARS-CoV-1, MERS-
94 CoV, HCoV-229E and HCoV-NL63 [15, 24-30].

95 In the Democratic Republic of the Congo (DRC) and the Republic of Congo (ROC) contact with
96 wildlife is common for large parts of the population via the value chain (food or otherwise), as
97 pests in house and fields, at peri-domestic and co-feeding interfaces, or in the context of
98 conservation and tourism [31, 32]. This close contact does not only involve risks for humans, but
99 also potentially for endangered animal species such as great apes [33]. To explore the
100 coronavirus presence in wildlife in this region representing one of the most biodiverse places on
101 the African continent, we launched large scale sampling of primarily bats, rodents and non-human

102 primates (NHPs). Our goal was to determine the degree of coronavirus circulation and diversity,
103 using a consensus Polymerase Chain reaction (PCR) approach, coupled with the collection of,
104 and coupling with, ecological data.

105

106 **Materials and Methods**

107 Sample acquisition differed depending on the species and interface. Animals in peri-domestic
108 settings were captured and released after sampling (bats, rodents and shrews only), while
109 samples from the (bushmeat) value chain were collected from freshly killed animals voluntarily
110 provided by local hunters upon their return to the village following hunting, or by vendors at
111 markets. Fecal samples were collected from free-ranging NHPs. Some NHP samples were also
112 collected during routine veterinary exams in zoos and wildlife sanctuaries. Hunters and vendors
113 were not compensated, to avoid incentivizing hunting. Oral and rectal swab samples were
114 collected into individual 2.0 ml screw-top cryotubes containing 1.5 ml of either Universal Viral
115 Transport Medium (BD), RNA later, lysis buffer, or Trizol® (Invitrogen), while pea-sized tissue
116 samples were placed in 1.5ml screw-top cryotubes containing 500ul of either RNA later or lysis
117 buffer (Qiagen), or without medium. All samples were stored in liquid nitrogen as soon as
118 practical. Sample collection staff wore dedicated clothing: N95 masks, nitrile gloves, and
119 protective eyewear during animal capture, handling and sampling.

120 RNA was extracted either manually using Trizol®, with an Qiagen AllPrep kit (tissue), Qiagen
121 Viral RNA Mini Kit (swabs collected prior to 2014), or with a Zymo Direct-zol RNA kit (swabs
122 collected after 2014) and stored at -80°C. Afterwards RNA was converted into cDNA using a
123 Maxima H Minus First Strand cDNA Synthesis Kit (Thermo Scientific) or GoScript™ Reverse
124 Transcription kit (Promega) and stored at -20°C until analysis. Two conventional nested broad
125 range PCR assays, both targeting conserved regions within the RNA-Dependent RNA
126 Polymerase gene (RdRp) were used to test the samples for coronavirus RNA. The first PCR
127 amplifies a product of approximately 286nt between the primer binding sites. The first round (CoV-

128 FWD1: CGT TGG IAC WAA YBT VCC WYT ICA RBT RGG and CoV-RVS1:
129 GGTCATKATAGCRTCAVMASWWGCNACATG) and second round (CoV-FWD2: GGC WCC
130 WCC HGG NGA RCA ATT and CoV-RVS2: GGW AWC CCC AYT GYT GWA YRT C) primers of
131 this PCR were specifically designed for the detection of a broad range of coronaviruses [34]. The
132 second PCR was used in two modified versions: one of them specifically targeting a broad range
133 of coronaviruses in bats, the second one broadly targeting coronaviruses of other hosts. In both
134 cases, the first round of the semi nested PCR utilized the primers CoV-FWD3 (GGT TGG GAY
135 TAY CCH AAR TGT GA) and CoV-RVS3 (CCA TCA TCA SWY RAA TCA TCA TA). In the second
136 round, either CoV-FWD4/Bat (GAY TAY CCH AAR TGT GAY AGA GC) or CoV-FWD4/Other
137 (GAY TAY CCH AAR TGT GAU MGW GC) were used as forward primers, while the reverse
138 primer was again CoV-RVS3 [35]. Both versions amplify 387nt between the primer binding sites.
139 PCR products were subjected to gel electrophoresis on a 1.5% agarose gel and products of the
140 expected amplicon sizes were excised. DNA was extracted using the Qiagen QIAquick Gel
141 Extraction Kit and either sequenced by Sanger sequencing at the UC Davis DNA sequencing
142 facility or was sent for commercial Sanger sequencing (GATC or Macrogen). Extracts with low
143 DNA concentrations were cloned prior to sequencing. All results from sequencing were analyzed
144 in the Geneious 7.1 software, and primer trimmed consensus sequences compared to the
145 GenBank database (BLAST N, NCBI).

146 Viral sequences were deposited in the GenBank database under submission numbers KX284927-
147 KX284930, KX285070-KX285095, KX285097-KX285105, KX285499-KX285513, KX286248-
148 KX286258, KX286264-KX286286, KX286295-KX286296, KX286298-KX286322, MT064119-
149 MT064126, MT064226, MT064272, MT081973, MT081997-MT082004, MT082032, MT082059-
150 MT082060, MT082072, MT082123-MT082136, MT082145, MT082299, MT222036- MT222037.

151 Maximum likelihood phylogenetic trees were constructed including different genera (Alpha, Beta
152 and Gamma) and species of known coronaviruses, as well as species/sub-species detected in
153 DRC and ROC during the PREDICT project. Only a single sequence was included representing

154 sequences with nucleotide identities of more than 95%. Multiple sequence alignments were made
155 in Geneious (version 11.1.3, ClustalW Alignment). Bayesian phylogeny of the polymerase gene
156 fragment was inferred using MrBayes (version 3.2) with the following parameters: Datatype=DNA,
157 Nucmodel=4by4, Nst=1, Coavion=No, # States=4, Rates=Equal, 2 runs, 4 chains of 5,000,000
158 generations. The sequence of an avian Gamma Coronavirus (NC_001451) served as outgroup
159 to root the trees, and trees were sampled after every 1,000 steps during the process to monitor
160 phylogenetic convergence [36]. The average standard deviation of split frequencies was below
161 0.006 for the Watanabe PCR amplicon based analysis and below 0.0029 for the Quan PCR
162 amplicon based analysis (MrBayes recommended final average <0.01). The first 10% of the trees
163 were discarded and the remaining ones combined using TreeAnnotator (version 2.5.1;
164 <http://beast.bio.ed.ac.uk>) and displayed with FIGTREE (1.4.4; <http://tree.bio.ed.ac.uk/>) [37].
165 Ecological data related to the locality and the host animals was compiled and analyzed with
166 respect to a correlation with the frequency of virus detection. The data included sex, human
167 interface at which the animals were collected and sampled (value chain or other), and local
168 calendric season (wet/dry) based on the climate-data.org data set, and were evaluated using two-
169 tailed Chi-square tests with Yates correction.

170

171 **Results**

172 Between 2006 and 2018 a total of 3,561 animals (2,630 from DRC and 931 from RoC) were
173 sampled and tested, of which 1,356 were bats (24 genera), 1,347 rodents (33 genera), 836 NHPs
174 (14 genera), and 22 shrews, (Figure 1, Supplements 1 and 2). The majority of the 5,586 collected
175 samples were oral (2,258) or rectal (2,238) swabs, with others being tissue samples, including
176 liver and spleen (385), lung (187) or intestinal tract (175), as well as feces (167), blood, serum or
177 plasma (140) and others (36). Coronavirus RNA was detected in oral (23) and/or rectal (102)
178 swabs and one pooled liver and spleen sample from a total of 121 animals. Viral RNA was
179 amplified in 83 samples using the Watanabe PCR assay and in 73 samples using the Quan PCR

180 assay (Supplement 3). Two of the animals with detected coronavirus RNA were rodents (<1% of
181 sampled rodents), while 119 were bats (8.8% of sampled bats). Coronavirus RNA positive animals
182 were found in 25% (27/106) of bat sampling events (same location and same day) and <1%
183 (2/235) of rodent sampling events (Supplements 2-4). In 10 of the bat sampling events, a single
184 coronavirus RNA positive bat was among the tested animals, while in 17 events the number of
185 bats positive for coronavirus ranged from 2 to 16 (Supplement 3). RNA was detected in two
186 species of rodents, one *Deomys ferrugineus* (1/1) and one *Malacomys longipes* (1/38), and in at
187 least 14 different bat species, namely *Chaerephon pumilus*, *Eidolon helvum*, *Epomops franqueti*,
188 *Hipposideros caffer*, *Hipposideros gigas*, *Hipposideros ruber*, *Megaloglossus woermanni*,
189 *Micropteropus pusillus*, *Miniopterus inflatus*, *Mops condylurus*, *Myonycteris sp.*, *Rhinolophus sp.*,
190 *Scotophilus dinganii* and *Triaenops persicus* (Table 1, Supplement 3). Among the five bat species
191 from which more than 100 individuals were sampled and tested, *Eidolon helvum* had the highest
192 rate of coronavirus RNA positives (22.3%), followed by *Epomops franqueti* (15.8%),
193 *Megaloglossus woermanni* (8.5%), *Mops condylurus* (7.6%), and *Micropteropus pusillus* (7%)
194 (Table 1). With 10.2% Yinpterchiroptera bats had a significantly ($N=0.015$ C²Y) higher rate of
195 coronavirus RNA positive animals than Yangochiroptera bats with 5.0% (Table 1). No coronavirus
196 RNA positive animals were detected among the sampled NHPs or shrews.

197 Significant seasonal differences for the rate of coronavirus RNA positive animals were detected
198 across the bats with a 10.5% PCR positive rate in the wet season and a 6.6% rate in the dry
199 season ($p = 0.0176$) (Table 1, Supplement 4). Bats that were associated with the (bushmeat)
200 value chain were more frequently positive for coronavirus RNA (25.4%) than bats sampled at
201 other human animal (peri-domestic) interfaces (5%), and this difference was highly significant (p
202 < 0.0001). Male bats were significantly overrepresented among the Coronavirus RNA positives p
203 $= 0.0183$), while there was insufficient data to analyze an influence of age (Supplement 2).

204 Upon phylogenetic analysis, the sequences fall into 13 separate clusters based on the Quan PCR
205 amplicon and 13 separate clusters based on the Watanabe PCR amplicon. Based on amplicons

206 obtained with both PCRs from the same sample or animal, the respective Quan and Watanabe
207 sequence clusters Alpha 5, 6, and 7 (Q7=W2), as well as Beta 1, 2, and 3 correspond to each
208 other. In one bat, RNA corresponding to two different alphacoronaviruses was detected in the oral
209 and the rectal sample by the same PCR assay (ZB12030), while in another bat one PCR assay
210 amplified RNA indicating an Alpha- and the other assay an RNA indicating a betacoronavirus
211 (GVF-RC-1006) (Supplement 3). Given the overall results, RNA of 15 different Alpha- and 6
212 betacoronaviruses was detected in the study population. In 22 of the sampling events, only a
213 single type/strain of these coronaviruses was detected, two in two events, and three, five or eight
214 in one event each. Identical or very similar coronavirus sequences were found with a spatial
215 distance of up to 1975 km apart and a temporal distance of up to 1708 days (Supplement 5).
216 Although the two coronavirus sequences we detected in rodents were clustering with known
217 sequences from bat alphacoronaviruses, there were no sequences in GenBank that shared more
218 than 80% identities with either of them (Figure 2, Supplement 3). The detected bat coronavirus
219 sequences on the contrary mostly clustered closely with known ones, that to a large part were
220 detected in hosts from the same genus (Figures 2 and 3). The majority of the detected sequences
221 were closely related to only two known viruses. Sequences with nucleotide identities of 97% or
222 higher to Kenya bat coronavirus BtKY56 were found in 53 individual bats of 9 different species
223 sampled on 14 occasions (Q-/W-Beta 2), while sequences with identities of 99% or higher to
224 Eidolon bat coronavirus/Kenya/KY24 were detected in 30 individual bats of 3 different species
225 sampled on 8 occasions (Q-/W-Beta 3) (Supplements 3 & 5). Bat coronavirus sequences in
226 clusters Q-Alpha 1, 2, 7, and 8, W-Alpha 2 and 7, Q-/W-Beta 1, and Q-Beta 4 and 5 had identities
227 of below 95% with known coronaviruses.
228 In three cases (Q-Alpha 1, W-Alpha 7 and 8), sequences were closest to coronaviruses found in
229 bats and camels with a high similarity (>90% nucleotide identities) to human coronavirus 229E
230 (Figures 2 and 3). Similarly, the viral sequences in clusters Q-Alpha 2 and Q-/W-Alpha 6 were

231 most closely related to bat coronaviruses with some similarity (>80% nucleotide identities) to
232 human coronavirus NL63 (Figures 2 and 3).

233

234 **Discussion**

235 We detected coronavirus RNA in a significant proportion of the sampled bats (8.8%), but only in
236 a small proportion of rodents (<1%) and none in NHPs or shrews. Finding relatively high numbers
237 of coronavirus RNA positive bats is consistent with what has been previously reported; continuous
238 high circulation of coronaviruses seems to be common especially in bats in tropical and
239 subtropical climates [15]. The specific PCR positive rates need to be approached with caution
240 though, since factors such as species, season, location and others could play a role, as well as
241 sample material and assays used for detection in comparison to other studies.

242 Our data suggest that coronavirus circulation in bats, at least in the Congo Basin, may indeed
243 depend to some extent on species and seasonality (Supplement 4). We observed a significant
244 difference in the number of bats testing positive depending on the local calendric season ($p =$
245 0.0176), with 10.5% of coronavirus RNA positive bats in the wet season but only 6.6% in the dry
246 season at similar sample sizes for both seasons (Table 1). Interestingly, when looking at the family
247 and species level, this holds true only for the Pteropodidae and Rhinolophidae species ($p <$
248 0.0001) while Hipposideridae, Miniopteridae, Molossidae, and Vespertilionidae species are more
249 likely to be positive for coronavirus RNA in the dry season ($p < 0.0001$) (Table 1). The latter,
250 though not for those specific families but for bats in general, has been proposed to be the
251 correlation on a global scale [15]. We can only speculate as to the reasons of the apparent
252 seasonality, but family and species seem to be important determinants. Due to the diverse set of
253 species in our sample set, individual sample numbers for most species are too small to draw
254 definite conclusions, however the significant seasonal difference between *Yinpterchiroptera* and
255 *Yangochiroptera* are largely supported by respective trends in the individual species. We tested
256 if the results from any particular species might be responsible for the observed correlation

257 between season and the rate of positive coronavirus RNA animals. The only species that turned
258 out to have a strong influence on the outcome was *Eidolon helvum*. However, the effect of
259 dropping it from the analysis did only influence the outcome for bats in total, while it was not strong
260 enough to negate the observed statistical significances for season within the *Pteropodidae* family
261 or the *Yinpterchiroptera* suborder.

262 We did find *Eidolon helvum*, a bat usually roosting in large colonies, to be significantly
263 overrepresented among the coronavirus positive bats ($p = 0.0005$), and higher detection rates in
264 this species have been reported before [15, 39]. However, samples from *Eidolon helvum* in this
265 study were collected from animals sold at two different markets on seven different days, and
266 although we detected coronavirus RNA in some *Eidolon helvum* bats obtained at each of those
267 occasions, it is possible that many of these bats came from the same roosts. The fact that all but
268 one of the *Eidolon helvum* bats were found to be positive for the same coronavirus type (Q-W-
269 Beta-3) supports the assertion that there may be a connection between those bats. Our dataset
270 does contain evidence that bat coronaviruses are readily shared within local bat populations. In
271 fact, 109 out of 119 coronavirus positive bats were from sampling events with at least one other
272 coronavirus RNA positive bat, and in all but six of these cases there was another bat with the
273 same coronavirus type in the event-cohort (Supplements 3 and 5). Even though we cannot
274 pinpoint the exact roosting relationship between all of these bats, this does confirm that
275 coronaviruses are readily shared among the bats in an area, even across species boundaries. It
276 also highlights that several different coronaviruses can circulate in parallel, including occasional
277 double infections (Supplement 3).

278 We found a much higher percentage of bats that were part of the bushmeat value chain to be
279 positive for coronavirus RNA, which could have significant implications for the risk of coronavirus
280 spillover from bats into humans. In our data set, 81% of the value chain samples were collected
281 in the wet season, and this group also contained all of the *Eidolon helvum* samples. This suggests
282 that seasonality and preferentially hunted species (80% *Pteropodidae*) are likely responsible for

283 the higher rate of coronavirus RNA positive animals in the value chain. According to our data, it
284 also seems that male animals are overrepresented among bats with coronavirus positive
285 samples. It is possible that behavioral differences between males and females play a role, such
286 as reduced activity of females during the time of birthing and breastfeeding or higher stress levels
287 among males during the breeding season [40]. Further investigation is required to confirm and
288 assess the reasons for this observation.

289 It appears clear from our findings, that bats rather than rodents or primates are sustaining a
290 significant circulation of coronaviruses in the Congo Basin. Evidence for coronavirus circulation
291 in wild animals other than bats is generally much scarcer, even though civets, raccoon, dogs, and
292 camels have been shown to be involved in outbreaks of SARS and MERS [8, 11, 15].

293 We estimate that the 121 detected sequences correspond to 21 different coronaviruses based on
294 the differences between the amplified sequences, considering the conserved nature of the
295 amplified fragments within the RdRp open reading frame (ORF). These 21 coronaviruses include
296 some that appear to only be distantly related to already described coronaviruses, and others that
297 have already been found elsewhere, such as Kenya bat coronavirus BtKY56 and Eidolon bat
298 coronavirus/Kenya/KY24 (Figures 2 and 3, Supplement 3).

299 RNA of either Kenya bat coronavirus BtKY56 or Eidolon bat coronavirus/Kenya/KY24 was
300 detected in ~70% (83) of the positive bats in this study and in several hundred bats reported
301 previously (GenBank). Interestingly Kenya bat coronavirus BtKY56 appears to be a common virus
302 species in the Congo Basin, while elsewhere it appears to be Eidolon bat
303 coronavirus/Kenya/KY24 that is more common. These observations are undoubtedly susceptible
304 to a sampling bias, for example due to the species composition of sample sets, particularly with
305 *Eidolon helvum*, which can be sampled in large numbers when colonies are present or when they
306 are present in markets [41]. However, we do find evidence of these two viruses in a relative wide
307 array of bat hosts, indicating that species barriers may not be a limiting factor for sharing these
308 specific Beta coronaviruses (Figures 2 and 3, Supplement 3). In contrast, most of the other

309 sequences that we detected with related sequences in GenBank were detected in bats of the
310 same genus by us and previously by others, supporting some degree of general species
311 specificity and virus host co-evolution despite the latent ability of at least some coronaviruses to
312 jump species barriers within and outside of the taxonomic order of hosts [14, 15, 17]. How often
313 these events occur is not fully understood, but it is generally assumed that bats serve as a
314 reservoir for coronaviruses [16]. With SARS-CoV-1 and SARS-CoV-2, the available evidence
315 suggests that they were successfully transmitted from bats into humans, either directly or
316 indirectly [20]. When we add to these two coronaviruses MERS that originated in bats and
317 established a sustained reservoir in camels with occasional spillover into humans, we have
318 witnessed three coronavirus spillover events with a bat origin in less than two decades.
319 Considering our increased awareness and abilities to detect the emergence of novel viruses, it
320 can be assumed that there may have been multiple coronavirus zoonotic events in the past that
321 either led to some degree of either self-limiting outbreaks, or may have established a permanent
322 virus host relationship with a new host [42]. MERS-CoV and SARS-CoV-1 may represent
323 examples for the former, while the latter may be represented by human coronaviruses 229E and
324 NL63; the ultimate outcome with regards to SARS-CoV-2 remains undetermined however. In our
325 study we also detected viral RNA related to human coronaviruses 229E and NL63 in eight bats.
326 Whether or not these relatives of human pathogens or other strains of the coronaviruses currently
327 circulating in bats can and will jump into humans in the future is difficult to predict at present.
328 Progress in the understanding of molecular processes such as RNA polymerase proofreading
329 capability, receptor usage, as well as in the field of human behavior are however certainly helping
330 our understanding of risk [43]. The close contact of humans with wildlife including bats in the
331 Congo Basin, especially in the context of hunting and wild animal trade, are certainly factors
332 contributing to a higher risk for zoonotic events involving coronaviruses or other infectious agents.
333 The two sequences we detected in rodents (*Deomys ferrugineus* and *Malacomys longipes*) likely
334 correspond to novel Alpha coronaviruses. The lack of sequences closely related to the two

335 indicates that rodent coronaviruses may be an understudied field, especially considering that
336 rodents are the largest family of mammals.

337 We conclude overall, that bats and to a much smaller degree rodents in the Congo Basin harbor
338 diverse coronaviruses, of which some might have the molecular potential for spillover into
339 humans. Considering the close contact between wildlife and humans in the region, as part of the
340 value chain or in peri-domestic settings, there is an elevated and potentially increasing risk for
341 zoonotic events involving coronaviruses. Thus, continued work to understand the diversity,
342 distribution, molecular mechanisms, host ecology, as well as consistent surveillance of
343 coronaviruses at likely hotspots, are critical to help prevent future global pandemics.

344

345 **Acknowledgements**

346 The authors would like to thank: The governments of the Democratic Republic of the Congo and
347 of the Republic of Congo for the permission to conduct this study; the staff of Metabiota, Wildlife
348 Conservation Society, and EcoHealth Alliance who assisted in sample collection and testing and
349 other members of the PREDICT-1 and PREDICT-2 consortium
350 (<https://ohi.vetmed.ucdavis.edu/programs-projects/predict-project/authorship>). This study was
351 made possible by the generous support of the American people through the United States Agency
352 for International Development (USAID) Emerging Pandemic Threats PREDICT program
353 (cooperative agreement numbers GHN-A-OO-09-00010-00 and AID-OAA-A-14-00102). The
354 contents are the responsibility of the authors and do not necessarily reflect the views of USAID
355 or the United States Government.

356

357 **Compliance with Ethical Standards**

358 Animal capture and specimen collection was approved by the Institutional Animal Care and Use
359 Committee (UCDavis IACUC, Protocol #s 16067 and 17803), the Institute Congolais pour la
360 Conservation de la Nature (0374/ICCN/DG/ADG/ADG/KV/2011) in DRC, and the Ministry of

361 Forest Economy and Sustainable Development (1102/MEFDD/DGEFDFAP-SPR) and the
362 Ministry of Scientific Research and Technical Innovation (permit number
363 014/MRS/DGRST/DMAST, 018/MRSIT/DGRST/DMAST) in the RoC. All authors declare that
364 they have no conflicts of interest.

365

366

367 **References**

- 368 1. Masters PS, Perlmann S. Coronaviridae. In: Knipe DM, Howley PM (eds) *Fields Virology*,
369 6th edn. Lippincott Williams & Wilkins, Philadelphia pp 825-858; 2013.
- 370 2. King AMQ, Adams MJ, Carstens EB, Lefkowitz EJ. *Virus taxonomy. Classification and*
371 *nomenclature of viruses, ninth report of the international committee on taxonomy of*
372 *viruses*. International Union of Microbiological Societies, Virology Division, Elsevier
373 Academic Press; 2012.
- 374 3. Drosten C, Günther S, Preiser W, et al. Identification of a novel coronavirus in patients
375 with severe acute respiratory syndrome. *N Engl J Med*. 2003 May 15;348(20):1967-76.
- 376 4. Zaki AM, van Boheemen S, Bestebroer TM, Osterhaus AD, Fouchier RA. Isolation of a
377 novel coronavirus from a man with pneumonia in Saudi Arabia. *N Engl J Med*
378 2012;367:1814-20.
- 379 5. Zhou P, Yang XL, Wang XG, et al. A pneumonia outbreak associated with a new
380 coronavirus of probable bat origin. *Nature* 2020;579:270-273.
- 381 6. Ayukekbong JA, Ntemgwa ML, Ayukekbong SA, et al. COVID-19 compared to other
382 epidemic coronavirus diseases and the flu. *World J Clin Infect Dis* 2020;10:1-13.
- 383 7. Li W, Shi Z, Yu M, et al. Bats are natural reservoirs of SARS-like coronaviruses. *Science*
384 2005;310:676-9.
- 385 8. Kan B, Wang M, Jing H, et al. Molecular evolution analysis and geographic investigation
386 of severe acute respiratory syndrome coronavirus-like virus in palm civets at an animal
387 market and on farms. *J Virol* 2005;79:11892–11900.
- 388 9. Yip CW, Hon CC, Shi M, et al. Phylogenetic perspectives on the epidemiology and origins
389 of SARS and SARS-like coronaviruses. *Infect Genet Evol* 2009;9:1185-96.
- 390 10. Corman VM, Ithete NL, Richards LR, et al. Rooting the phylogenetic tree of middle East
391 respiratory syndrome coronavirus by characterization of a conspecific virus from an
392 African bat. *J Virol* 2014;88:11297-303.

- 393 11. Sabir JS, Lam TT, Ahmed MM, et al. Co-circulation of three camel coronavirus species
394 and recombination of MERS-CoVs in Saudi Arabia. *Science* 2016;351:81-4.
- 395 12. Pfefferle S, Oppong S, Drexler JF, et al. Distant relatives of severe acute respiratory
396 syndrome coronavirus and close relatives of human coronavirus 229E in bats, Ghana.
397 *Emerg Infect Dis* 2009;15:1377–1384.
- 398 13. Huynh J, Li S, Yount B, et al. Evidence supporting a zoonotic origin of human coronavirus
399 strain NL63. *J Virol* 2012;86:12816-25.
- 400 14. Lau SK, Li KS, Tsang AK, et al. Recent transmission of a novel alphacoronavirus, bat
401 coronavirus HKU10, from Leschenault's rousettes to pomona leaf-nosed bats: first
402 evidence of interspecies transmission of coronavirus between bats of different suborders.
403 *J Virol* 2012;86:11906-18.
- 404 15. Anthony SJ, Johnson CK, Greig DJ, et al. Global patterns in coronavirus diversity. *Virus*
405 *Evol* 2017;3:vex012.
- 406 16. Woo PC, Lau SK, Huang Y, et al. Coronavirus diversity, phylogeny and interspecies
407 jumping. *Exp Biol Med* 2009;234:1117–1127.
- 408 17. Leopardi S, Holmes EC, Gastaldelli M, et al. Interplay between co-divergence and cross-
409 species transmission in the evolutionary history of bat coronaviruses. *Infect Genet Evol*
410 2018;58:279-289.
- 411 18. Crossley BM, Mock RE, Callison SA, et al. Identification and characterization of a novel
412 alpaca respiratory coronavirus most closely related to the human coronavirus 229E.
413 *Viruses* 2012; 4:3689 –3700.
- 414 19. Corman VM, Baldwin HJ, Tateno AF, et al. Evidence for an Ancestral Association of
415 Human Coronavirus 229E with Bats. *J Virol* 2015;89:11858-70.
- 416 20. Cui J, Li F, Shi ZL. Origin and evolution of pathogenic coronaviruses. *Nat Rev Microbiol*
417 2019;17:181-192.

- 418 21. Jones KE, Patel NG, Levy MA, et al. Global trends in emerging infectious diseases. *Nature*
419 2008;451:990-993.
- 420 22. Wolfe ND, Daszak P, Kilpatrick AM, et al. Bushmeat hunting, deforestation, and
421 prediction of zoonotic disease. *Emerg Infect Dis* 2005;11:1822.
- 422 23. Allen T, Murray KA, Zambrana-Torrel C, et al. Global hotspots and correlates of
423 emerging zoonotic diseases. *Nat Commun* 2017;8:1124.
- 424 24. Tong S, Conrardy C, Ruone S, et al. Detection of Novel SARS-like and Other
425 Coronaviruses in Bats from Kenya. *Emerg Infect Dis* 2009;15:482-5.
- 426 25. Tao Y, Tang K, Shi M, et al. Genomic characterization of seven distinct bat coronaviruses
427 in Kenya. *Virus Res* 2012;167:67-73.
- 428 26. Geldenhuys M, Weyer J, Nel LH, et al. Coronaviruses in South African Bats. *Vector Borne*
429 *Zoonotic Dis* 2013;13:516-9.
- 430 27. Razanajatovo NH, Nomenjanahary LA, Wilkinson DA, et al. Detection of new genetic
431 variants of Betacoronaviruses in Endemic Frugivorous Bats of Madagascar. *Virol J*
432 2015;12:42.
- 433 28. Geldenhuys M, Mortlock M, Weyer J, et al. A metagenomic viral discovery approach
434 identifies potential zoonotic and novel mammalian viruses in *Neoromicia* bats within South
435 Africa. *PLoS One* 2018;13:e0194527.
- 436 29. Markotter W, Geldenhuys M, Jansen van Vuren P, et al. Paramyxo- and Coronaviruses in
437 Rwandan Bats. *Trop Med Infect Dis* 2019;4pii:E99.
- 438 30. Nziza J, Goldstein T, Cranfield M, et al. Coronaviruses Detected in Bats in Close Contact
439 with Humans in Rwanda. *Ecohealth* 2020;17:152-159.
- 440 31. Van Vliet N, Nebesse C, Gambalemoke S, et al. The bushmeat market in Kisangani,
441 Democratic Republic of Congo: Implications for conservation and food security. *Oryx*
442 2012;46:196-203.

- 443 32. Van Vliet N, Schulte-Herbrüggen B, Muhindo J, et al. Trends in bushmeat trade in a
444 postconflict forest town: implications for food security. *Ecology and Society* 2017;22:35.
- 445 33. Patrono LV, Samuni L, Corman VM, et al. Human coronavirus OC43 outbreak in wild
446 chimpanzees, Côte d'Ivoire, 2016. *Emerg Microbes Infect* 2018;7:118.
- 447 34. Quan PL, Firth C, Street C, et al. Identification of a severe acute respiratory syndrome
448 coronavirus-like virus in a leaf-nosed bat in Nigeria. *mBio* 2010;1 pii: e00208-10.
- 449 35. Watanabe S, Masangkay JS, Nagata N, et al. Bat coronaviruses and experimental
450 infection of bats, the Philippines. *Emerg Infect Dis* 2010;16:1217-1223.
- 451 36. Ronquist F, Teslenko M, van der Mark P, et al. MRBAYES 3.2: Efficient Bayesian
452 phylogenetic inference and model selection across a large model space. *Syst Biol*
453 2012;61:539-542
- 454 37. Bouckaert R, Vaughan TG, Barido-Sottani J, et al. BEAST 2.5: An advanced software
455 platform for Bayesian evolutionary analysis. *PLoS Comput Biol* 2019;15:e1006650.
- 456 38. R Core Team. R: A Language and Environment for Statistical Computing (2017). R
457 Foundation for Statistical Computing, Vienna, Austria.
- 458 39. Montecino-Latorre D, Goldstein T, Gilardi K, et al. Reproduction of East-African bats may
459 guide risk mitigation for coronavirus spillover. *One Health Outlook* 2020;2:2.
- 460 40. Fayenuwo JO, Halstead LB. Breeding cycle of straw-colored fruit bat, *Eidolon helvum*, at
461 Ile-Ife, Nigeria. *J Mammal* 1974;55:453-454.
- 462 41. Kamins AO, Restif O, Ntiamoah-Baidu Y, et al. Uncovering the fruit bat bushmeat
463 commodity chain and the true extent of fruit bat hunting in Ghana, West Africa. *Biological*
464 *Conservation* 2011;144:3000-08.
- 465 42. Wang N, Li SY, Yang XL, et al. Serological Evidence of Bat SARS-Related Coronavirus
466 Infection in Humans, China *Virol Sin* 2018;33:104-107.
- 467 43. Akem ES, Pemunta NV. The bat meat chain and perceptions of the risk of contracting
468 Ebola in the Mount Cameroon region. *BMC Public Health* 2020;20:593.

469

470 **Figure legends**

471 **Figure 1**

472 Geographical map indicating all sampling sites within the Republic of Congo and the Democratic
473 Republic of the Congo. Locations where coronaviruses were detected are highlighted with blue
474 triangles for bats and red circles for rodents. Sampling sites without viral RNA detection are
475 marked by black dots (see also Supplement 1).

476

477 **Figure 2**

478 Maximum likelihood phylogenetic tree of coronavirus sequences presented as a proportional
479 cladogram, based on the RdRp region targeted by the PCR by Watanabe et. al. [35]. The tree
480 includes the sequences detected during the project (red boxes) and indicates the number of
481 sequences sharing more than 95% nucleotide identities in brackets. GenBank accession numbers
482 are listed for previously published sequences, while sequences obtained during the project are
483 identified by cluster names (compare Supplement 3). Black font indicates coronavirus sequences
484 obtained from bats, brown font indicates rodents, blue humans and gray other hosts. The host
485 species and country of sequence origin are indicated for bats and rodents if applicable. In case
486 of clusters W-Alpha-1 sequences were detected in *Mops condylurus* and *Chaerephon sp.*, host
487 species in cluster W-Beta-1 were *Megaloglossus woermanni* and *Epomops franqueti* and in case
488 of cluster W-Beta-2 *Micropteropus pusillus*, *Epomops franqueti*, *Rhinolophus sp.*, *Myonycteris*
489 *sp.*, *Mops condylurus*, *Megaloglossus woermanni*, and *Eidolon helvum* (compare Supplement 3).
490 Numbers at nodes indicate bootstrap support.

491

492 **Figure 3**

493 Maximum likelihood phylogenetic tree of coronavirus sequences presented as a proportional
494 cladogram, based on the RdRp region targeted by the PCR by Quan et. al. [34]. The tree includes

495 the sequences detected during the project (red boxes) and indicates the number of sequences
496 sharing more than 95% nucleotide identities in brackets. GenBank accession numbers are listed
497 for previously published sequences, while sequences obtained during the project are identified
498 by cluster names (compare Supplement 2). Black font indicates coronavirus sequences obtained
499 from bats, brown font indicates rodents, blue humans and gray other hosts. The host species and
500 country of sequence origin are indicated for bats and rodents if applicable. In case of clusters Q-
501 Alpha-4 sequences were detected in *Mops condylurus* and *Chaerephon sp.*, host species in
502 cluster Q-Alpha-7 were *Epomops franqueti* and *Chaerephon pumilus*, in case of cluster Q-Beta-
503 2 *Micropteropus pusillus* and *Epomops franqueti*, and for cluster Q-Beta-3 *Megaloglossus*
504 *woermanni*, *Eidolon helvum*, and *Epomops franqueti* (compare Supplement 3). Numbers at nodes
505 indicate bootstrap support.
506

507 Table 1: PCR results by species and season (Bats)

508

Suborder, family and species (>10 sampled individuals)	Wet Season PCR positives	Dry Season PCR positives	Total PCR positives
<i>Yinpterchiroptera</i> total**	13.3% (78/586)	5.6% (23/408)	10.2% (101/994)
<i>Pteropodidae</i> total**	13.6% (77/567)	4.0% (12/303)	10.2% (89/870)
<i>Micropteropus pusillus**</i>	10.3% (27/263)	1.3% (2/153)	7% (29/416)
<i>Epomops franqueti</i>	16.5% (18/109)	13.5% (5/37)	15.8% (23/146)
<i>Megaloglossus woermanni</i>	11.9% (5/42)	6.6% (5/76)	8.5% (10/118)
<i>Eidolon helvum</i>	22.3% (23/103)	- (0/0)	22.3% (23/103)
<i>Myonycteris torquata</i>	0% (0/11)	0% (0/11)	0% (0/22)
<i>Rhinolophidae</i> total**	100% (1/1)	0% (0/61)	1.6% (1/62)
<i>Hipposideridae</i> total*	0% (0/18)	25% (11/44)	17.7% (11/62)
<i>Hipposideros ruber</i>	0% (0/12)	33.3% (3/9)	14.3% (3/21)
<i>Triaenops persicus</i>	- (0/0)	13.8% (4/29)	13.8% (4/29)
<i>Yangochiroptera</i> total**	0.6% (1/167)	8.7% (17/194)	5.0% (18/361)
<i>Miniopteridae</i> total	0% (0/1)	20% (3/15)	18.8% (3/16)
<i>Pipistrellus nanus</i>	0% (0/1)	0% (0/10)	0% (0/11)
<i>Molossidae</i> total**	1.1% (1/92)	14.8% (13/88)	7.8% (14/180)
<i>Chaerephon pumilus</i>	0% (0/33)	12.5% (4/32)	6.2% (4/65)
<i>Mops condylurus</i>	1.9% (1/52)	13.2% (7/53)	7.6% (8/105)
<i>Vespertilionidae</i> total	0% (0/74)	1.1% (1/91)	0.6% (1/165)
<i>Scotophilus dinganii</i>	0% (0/18)	9% (1/11)	3.4% (1/29)
Total**	10.5% (79/754)	6.6% (40/602)	8.8% (119/1356)

509

510 * Significant difference between calendric seasons $P < 0.05$ (Chi-square with Yates correction)

511 ** Highly significant difference between calendric seasons $P < 0.01$ (Chi-square with Yates

512 correction)

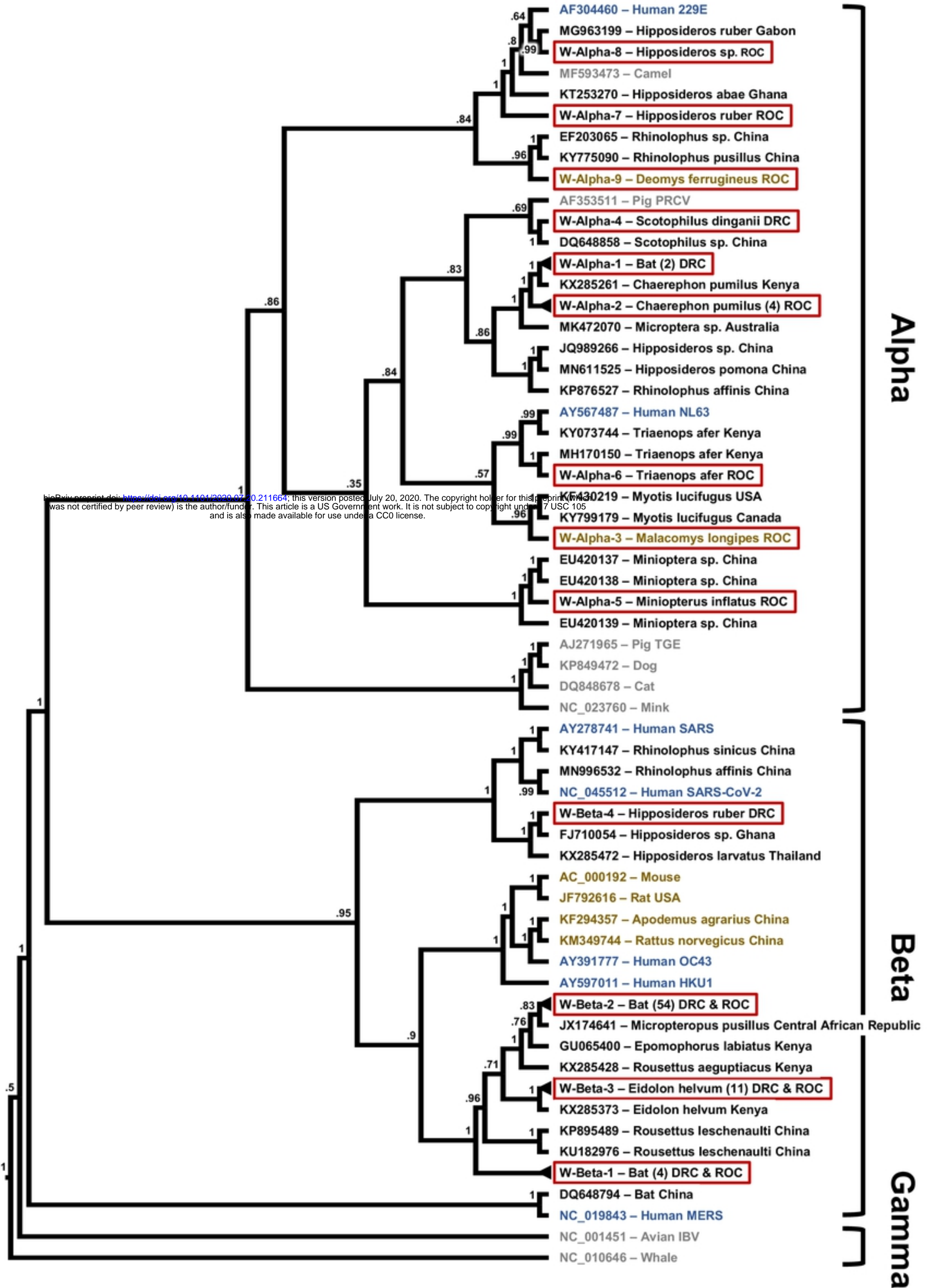
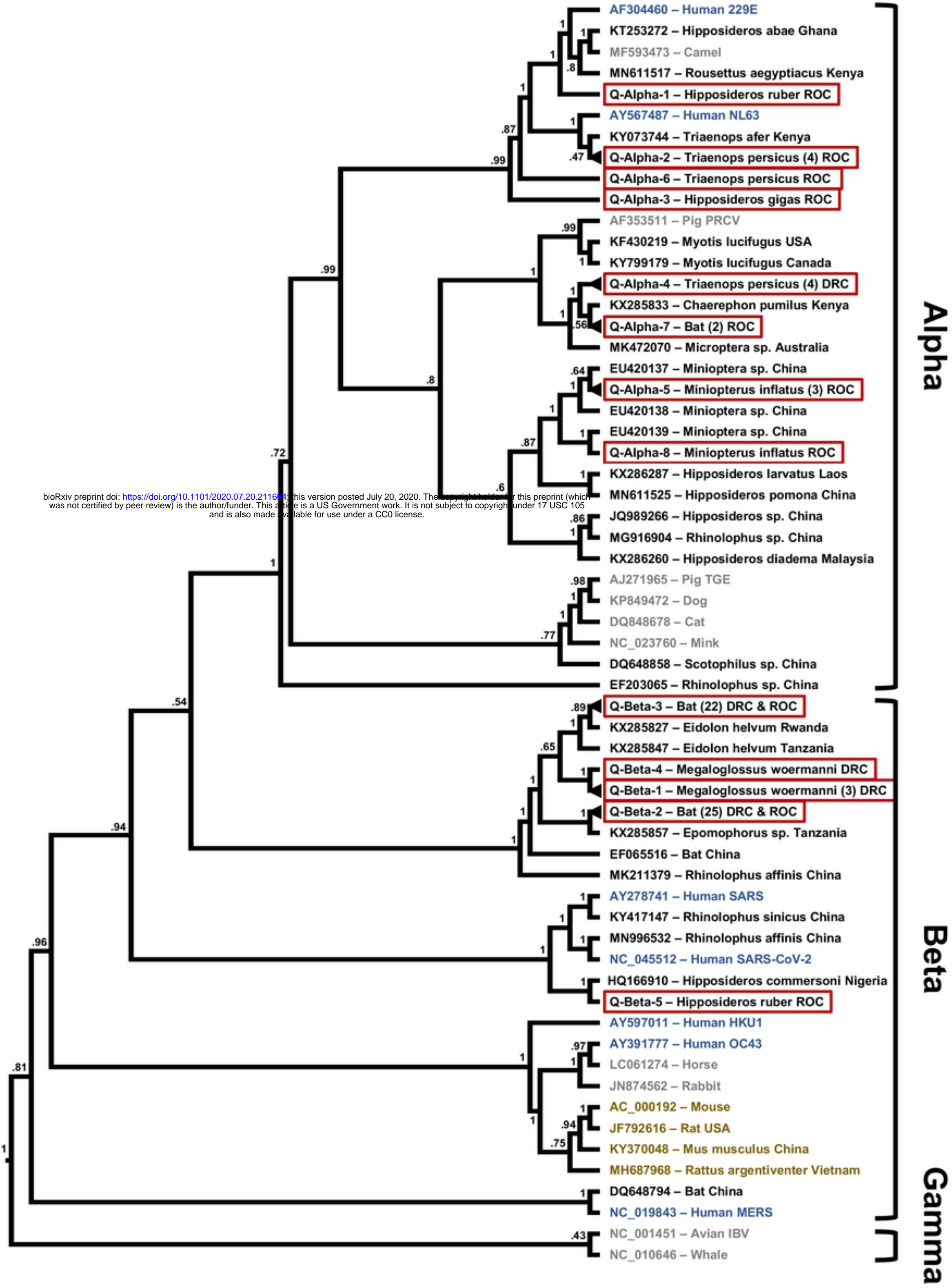


Figure 2



bioRxiv preprint doi: <https://doi.org/10.1101/2020.07.20.211684>; this version posted July 20, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.

Figure 3

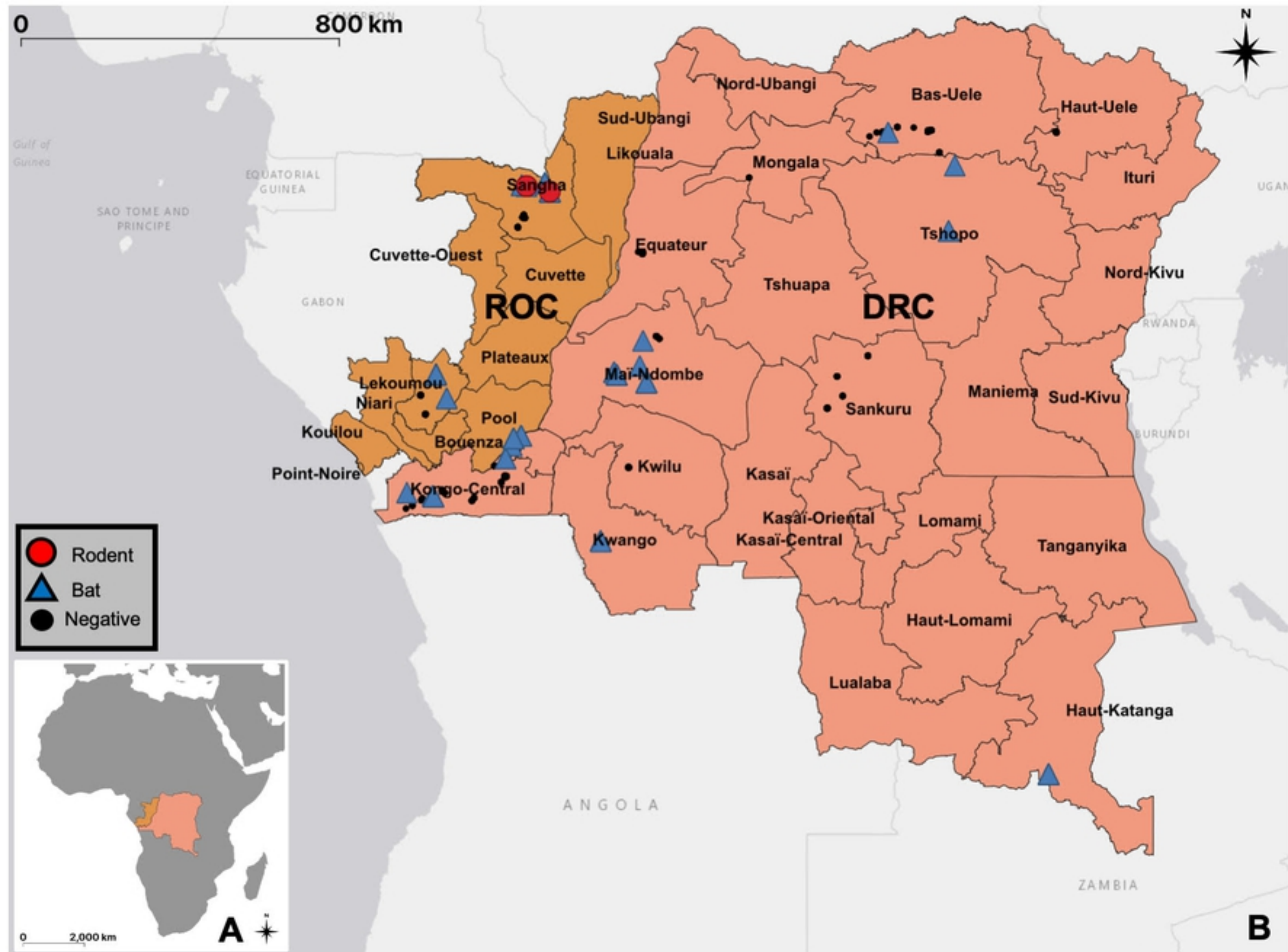


Figure 1