

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

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

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

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

Abstract

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

Introduction

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

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

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

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

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

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

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

Materials and Methods

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

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

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

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

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

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

Results

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

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

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

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

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

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

Discussion

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

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

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

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

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

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

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

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

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

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

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

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

Acknowledgements

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

Compliance with Ethical Standards

Animal capture and specimen collection was approved by the Institutional Animal Care and Use Committee (UCDavis IACUC, Protocol #s 16067 and 17803), the Institute Congolais pour la 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

References

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

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

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

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

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

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

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)

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

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

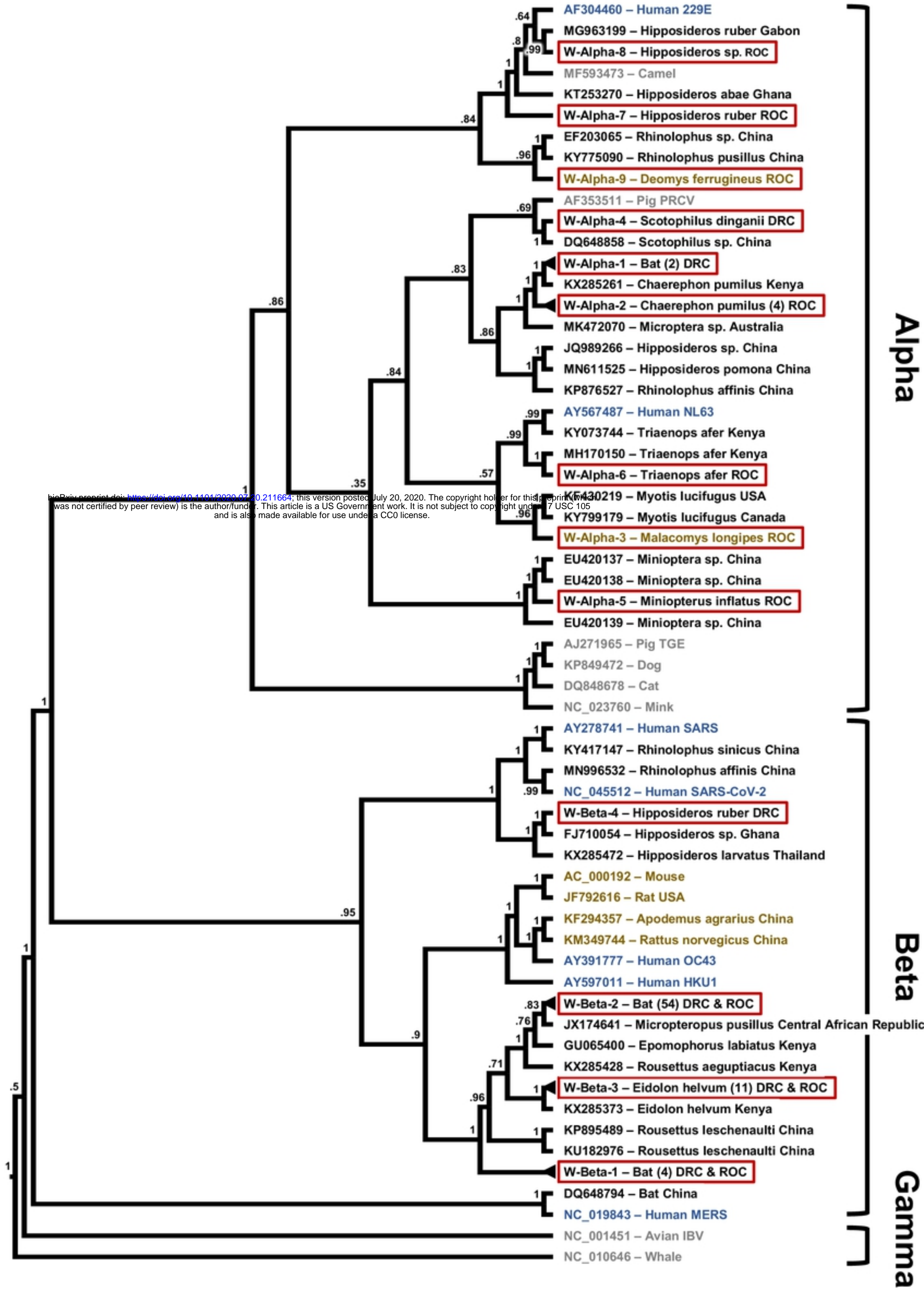


Figure 2

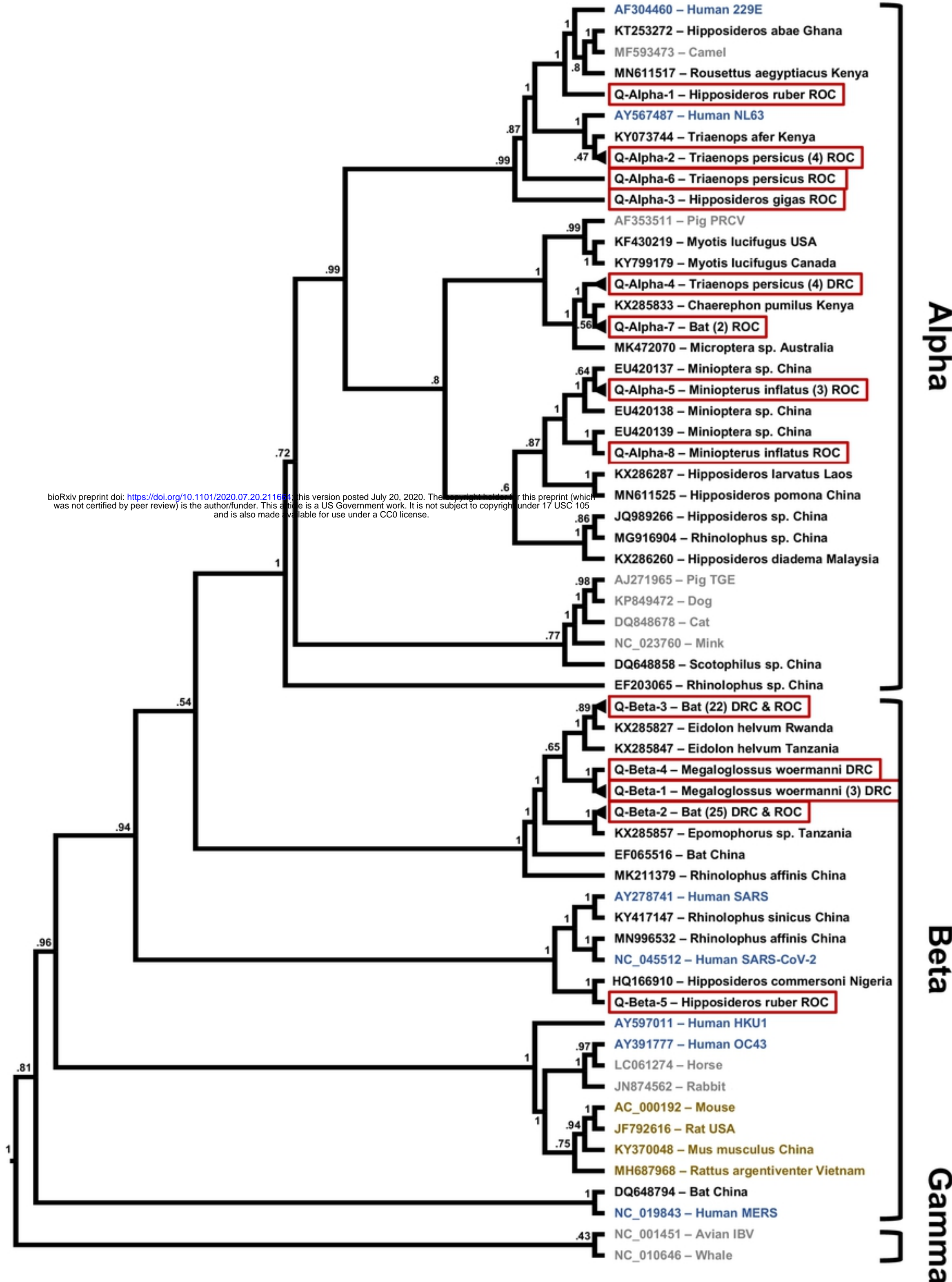


Figure 3

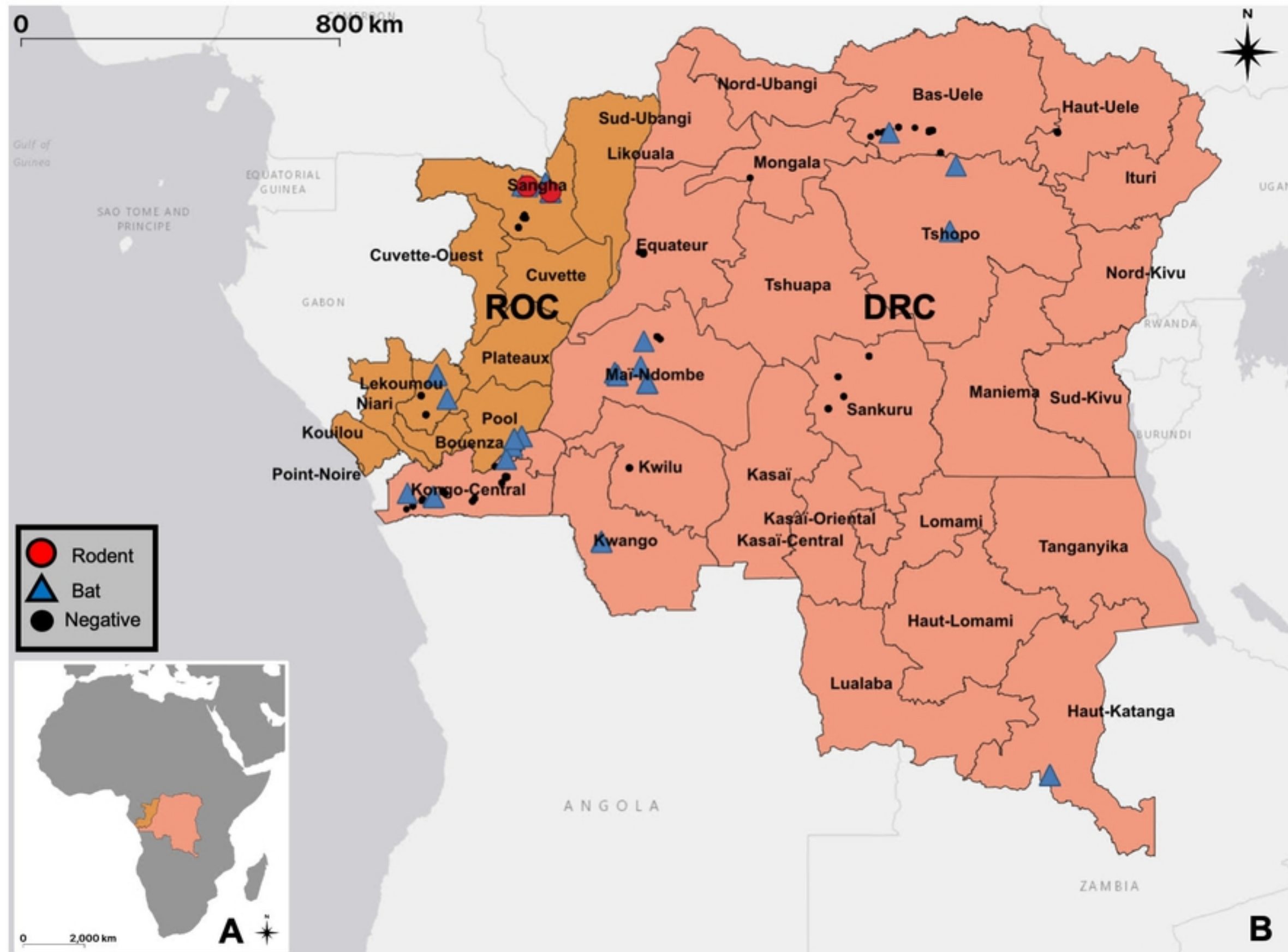


Figure 1