bioRxiv preprint doi: https://doi.org/10.1101/2020.05.14.056085; this version posted September 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

EVOLUTIONARY DYNAMICS AND GEOGRAPHIC DISPERSAL OF BETA CORONAVIRUSES IN AFRICAN BATS

Babatunde Olanrewaju Motayo[1,2], Olukunle Oluwapamilerin Oluwasemowo[2], Paul Akiniyi Akinduti[3]

1. Department of Medical Microbiology and Parasitology, Federal Medical Center, Abeokuta, Nigeria.
2. Department of Virology, College of Medicine, University of Ibadan, Ibadan, Nigeria.
3. Department of of Biological Science, Covenant University, Otta, Nigeria.

Address for correspondence, babatundemotayo@yahoo.com; niyiakinduti@gmail.com.

Babatunde O Motayo e-mail babatundemotayo@yahoo.com

Paul A Akinduti e-mail niyiakinduti@gmail.com

Olukunle O Oluwasemowo e-mail olukunlesemowo@yahoo.com

RUNNING TITLE: EVOLUTION OF CORONAVIRUSES IN AFRICAN BATS

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.14.056085; this version posted September 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

ABSTRACT

Bats have been shown to serve as reservoir host of various viral agents including coronaviruses. They have also been associated with the novel coronavirus SARS-CoV-2. This has made them an all important agent for CoV evolution and transmission. Our objective in this study was to investigate the dispersal, phylogenomics and evolution of *betacoronavirus* (βCoV) among African bats. We retrieved sequence data from established databases such as GenBank and Virus Pathogen Resource, covering the partial RNA dependent RNA polymerase (RdRP) gene of Bat coronaviruses from eight African, three Asian, five European, two South American countries and Australia. We analyzed for Phylogeographic information relating to genetic diversity and evolutionary dynamics. Our study revealed that majority of the African strains fell within *Norbecovirus* subgenera, with an Evolutionary rate of 1.301×10^{-3}, HPD ($1.064 \times 10^{-3} – 1.434 \times 10^{-3}$) subs/site/year. The African strains diversified into three main subgenera, *Norbecovirus*, *Hibecovirus* and *Marbecovirus*. The time to most common recent ancestor for *Norbecovirus* strains was 1968, and 2010, for the African *Marbecovirus* strains. There was evidence of inter species transmission of *Norbecovirus* among bats in Cameroun and DRC. Phlylogeography showed that there were inter-continental spread of Bt-CoV from Europe, China and Hong Kong into Central and Southern Africa, highlighting the possibility of long distance transmission. Our study has elucidated the possible evolutionary origins of βCoV among African bats, we therefore advocate for broader studies of whole genome sequences of BtCoV to further understand the drivers for their emergence and zoonotic spillovers into human population.

Key words: Phylogeography, Evolutionary dynamics, BtCoV, Africa

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.14.056085; this version posted September 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

INTRODUCTION

Coranaviruses are a large group of enveloped, positive sense, single stranded RNA viruses belonging to the order Nidovirales and family Coronaviridea [1]. The subfamily coronavirinae contains four genera: *Alphacoronavirus*, *Betacoronavirus*, *Gammacoronavirus* and *Deltacoronavirus* [1]. The international committee on the Taxonomy of viruses recently adopted addition changes to the nomenclature of Coronaviruses to include the inclusion of subgenera replacing the elsewhile lineage classification system [2]. Under this new system the genus *betacoronavirus* was further classified into five subgenera *Sarbecovirus, Marbecovirus, Norbecovirus, Embecovirus and Hibecovirus* [2]. *Betacoronaviruses* generally infect animals such as mammals and birds, they are the causative agents of many pathogenic diseases such as transmissible gastroenteritis of swine (TGEV), infectious bronchitis virus (IBD), mouse hepatitis virus (MHV), and bovine coronavirus (BCoV) [3]. Coronavirus have been reported as early as 1930 (1Masters and Pearlman, 2013) but the earliest report of human coronavirus was in the year 1960, where two strains namely hCoV229-E and hCoVOC43 were described [4,5]. Generally *betacoranvirus* have been observed to cause paucisymptomatic disease in man and are largely known to be zoonotic. It was not until after the advent of the severe acute respiratory syndrome (SARS) outbreak in Hong Kong and parts of China in 2003, that studies into the zoonotic origin of the incriminated pathogen SARS CoV revealed that the Chinese Rinolophid bats maintained a genetically related strain of the SARS CoV, [6,7]. This finding sparked up interest in bat CoV research. Another coronavirus termed Middle East severe respiratory syndrome virus (MERS CoV) was reported in the Arabian peninsula in 2012 [8]. Genetically similar strains to the MERS CoV were also isolated from *Pipistellus, Tyloncteris* and *Neoromica* bats [9].

In Africa large scale surveillance studies have identified diverse strains of coronavirus (CoV) circulating among African bats from Kenya, Ghana and Nigeria [10]. Studies have also provided evidence that the human coronavirus hCoV229E originated from African bat CoV (AfrBtCoV) [11]. Also the rich fauna and biodiversity in Africa has made it a hotspot for emerging viral diseases. It is also inhabited by a diversity of bats which have been identified to serve as a reservoir of high consequence zoonotic diseases such as Marburg hemorrhagic fever and Rabies [9]. Recently a novel coronavirus SARS CoV-2 was identified to be the cause of a gobal pandemic which originally broke out in Wuhan, Hubei province, China [12]. Some studies have also postulated that the SARS CoV-2 probably spilled over into human population through a zoonotic event involving Chinese SARS-related BtCoV [12,13].

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.14.056085; this version posted September 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

The recent evidence of African bats as a potential reservoir host for several *betacoronaviruses* (βCoV) gave rise to the conceptualization of this study which aims to investigate the spatial dispersal, phylogenomics and evolution of βCoV among African bats.

MATERIALS AND METHODS

Data Collection

We searched for and downloaded partial or complete gene coding regions for the RNA dependent RNA polymerase sequences (RdRP) of Afr-BtCoV from GeneBank and the Virus Pathogen resource database http://www.viprbic.org , and Genbank. The data set generated contained African bat βCoV from eight countries, namely Nigeria, Kenya, Ghana, Cameroun, Democratic Republic of Congo (DRC), Rwanda, Madagascar and South Africa (n=94), BtCoV from China, Hong Kong and Philippines in Asia; and France, Spain, Netherland, Italy and Luxemburg in Europe (n=95). Mexico and Brazil (n=3), Australia (n=1) and reference African CoV OC43, CoVHKU1, MERSCoV, and alpha coronavirus from Africa (n=35). Information such as country of origin, Host species, and date of collection were combined with the sequence information for the purpose of accurate phylogenetic determination. The final data stets had information from seven African countries, four European countries and three Asian countries. All the data used in this study can be assessed in Supplementary Table 1. Majority of the African BtCoV sequences were generated by nRT-PCR using primers targeting the 440bp partial RdRP gene region [14] and Sanger sequencing. Full genome sequences of ZBCoV were generated by both Sanger and ultra high throughput sequencing (UHTP) sequencing [15].

Phylogenetic analysis

Sequences were aligned using clustal W version 2.1 using default settings, the final alignment was 400bp in length. Phylogenetic trees were constructed in MEGA 7.0 software www.megasoftwre.net using the maximum likelihood method with a general time reversible GTR with a gama distributed rate variation (T_4) and a p- distance model with 1000 bootstrap resampling. The final trees were then visualized in fig tree http://tree.bio.ed.ac.uk/software/figtree/.

Discrete Phylogeographic analysis

Aligned sequences were analyzed for evidence of sufficient temporal clock signal using TempEst version 1.5 [16]. The relationship between root-to-tip divergence and sampling dates supported the use of molecular clock analysis in this study. Phylogenetic trees were generated by Bayasian inference through Markov chain Monte Carlo (MCMC), implemented in BEAST version 1.10.4 [17]. We partitioned the coding genes into first+second and third codon positions and applied a separate Hasegawa-Kishino-Yano (HKY+G) substitution model with gamma-distributed rate

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.14.056085; this version posted September 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

heterogeneity among sites to each partition [18]. Two clock models were initially evaluated strict and relaxed molecular clock, with four different tree priors, constant population size, exponential population size, Bayesian Skyride plot and Gausian Markov Random Field Skyride plot. Each selected model was run for an initial 30, 000, 000 states. Models were compared using Bayes factor with marginal likelihood estimated using the path sampling and stepping stone methods implemented in BEAST version 1.10.4 [17]. The relaxed clock with Gausian Markov Random Field Skyride plot (GMRF) coalescent prior was selected for the final analysis. The MCMC chain was set at 100, 000, 000 states with10% as burn in. Results were visualized using Tracer version 1.8. (http://tree.bio.ed.ac.uk/software/tracer/), all effective sampling size ESS values were >200 indicating sufficient sampling. Bayesian skyride analysis was carried out to visualize the epidemic evolutionary history using Tracer v 1.8. (http://tree.bio.ed.ac.uk/software/tracer/).To reconstruct the ancestral-state phylogeographic transmission across countries and hosts, we used the discrete-trait model implemented in BEAST version 1.10.4 [17].The Bayesian stochastic search variable selection (BSSVS) approach [19] was used to explore the most important historical dispersal routes for the spread BtCoV across their countries of origin, as well as the most probable host-species transition. The spatiotemporal viral diffusion was then visualized using the Spatial Phylogenetic Reconstruction of Evolutionary Dynamics SPREAD3 software [20].

RESULTS AND DISCUSION

PHYLOGENY AND EVOLUTIONARY DYNAMICS

We analyzed βCoV sequences from seven African countries distributed among eight bat species as shown in Table 1. The most abundant bat species sampled in this study was *Micropteropus. pussilus*, and Cameroon had the highest distribution of bat species sampled in this study. This result does not necessarily represent the true picture of bat species diversity in Africa, as some countries lack sequence information for bats due lack of surveillance. Few studies have been carried in Africa on CoV among bats leaving a huge gap in epidemiologic information regarding BtβCoV in Africa.

Phylogenetic analysis of Bt-βCoV sequences revealed a significant proportion of the African strains, isolated from fruit bats fell within the sub-genera Norbecovirus formerly known as lineage D consisting of strains from Cameroon, DRC, Kenya, Madagascar and Nigeria. This observation is in agreement with a previous report which identified the widespread circulation of Norbecovirus (Lineage D) among fruit bats in certain African regions [21]. However, it was identified that isolates consisting largely of strains isolated among *Neomoricia* South African bats clustering within the sub-genus *Marbecovirus* (formerly Lineage C) together with strains isolated from Italy and Spain (Figure 1). The phylogenetic classifications utilized in this study is based on the partial RdrP group unit (RGU), utilized for the rapid classification of field isolates of βCoV [22]. The species-specific phylogenetic clustering observed among the *Neomoricia* bats suggests limited inter-species βCoV transmission and host specific evolution among these species of bats in Africa as previously reported for BtCoV [23]. Larger epidemiological studies are needed among these species of bats to shed more light as to the cause of this observed trend in Africa. Studies have shown that βCoV of subgenera *Marbecovirus* such as MERSCoV are

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.14.056085; this version posted September 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

capable of both intra-species transmission and inter host transmission [24]. In this study there seemed to be inter-species transmission among the *Norbercovirus* (lineage D βCoV), evidenced by circulation of same *Norbecovirus* clade within different species of bats from same country around with the same year of isolation. For instance from Figure 2 it can be seen that Cameroonian bat species *Micropteropus pussilus, Epomophorus gambianus* and *Epomophorus franquenti*, were infected by the same *Norbecovirus* clade, isolated around the year 2013. This is also observed among bat species from DRC (Figure 2).This type of event allows for potential recombination and rapid evolution of this lineage, as previously reported [25]. This type of observation was also reported in an earlier study of an inter-species transmission event of alpha CoV HKU10 in bat species of different orders [26]. Results for root to tip divergence showed the data set had a positive temporal signal (Supplementary Figure 1) with the correlation coefficient = 0.0286 and R^2 = 0.0818.

The MCC tree of the Afr-βBtCoV strains shows clearly the two major sub-genera *Norbercovirus*, and *Marbecovirus* (Figure 2). We also observed that majority of the *Sarbecovirus* (formerly lineage B) BtβCoV were isolated in Europe, precisely France and Spain and also China (Figure 3). These consisted of SARSr viruses for instance EP11 strains which have been reported to be widely distributed across Europe and parts of Asia [27]. The absence of *Sarbecovirus* subgenera among the AfrBtCoV in our study seems to support the hypothesis that highly pathogenic CoV's such as SARS evolved outside the African continent. Although the lack of *Sarbecovirus* in this study does not imply that this group of viruses is not currently circulating among African bats, as the closest subgenra *Hibecovirus* which was also formerly classified under Lineage B was identified in African Bats from Nigeria and Ghana [15], clustering with an Australian isolate ascension no: EU834950. This simply reflects the information gap in molecular data of BtCoV owing to poor surveillance in Africa. This also extends to information human coronaviruses such as hCoVOC43, hCoV229E, in which sequence data is limited to just a few countries such as Kenya and South Africa [28,29]. Phylogenetically the genus *Rhinolophus* (Horse shoe bat) exhibited highest potential for intra-host diversity for BtCoV with the genus co-circulating both *Sarbecovirus* and *Norbecovirus* strains Figure 3. Our observation was similar to that of a study from Thailand [30] and supports the theory of diverse intra/inter-host transmission among different bat species which has been reported in previous studies [31,32]. Although we did not find this type of intra host genetic diversification among the African bat species in this study, it is believed that *Rhinolophus* bats are well distributed in Africa and are capable of zoonotic transmission of pathogenic hCoV such as SARS, as evidenced by a study that identified SARSrCoV antibodies among *Rhinolophus* bats in Africa [33].

The TMRCA for African *Norbecovirus* dating back to 1973, 95% HPD(1951-1990), and the TMRCA for *Marbecovirus* strains 2007 95% HPD (2003-2012). This shows that *Marbecovirus* is relatively recent and probably evolved from the existing *Norbecovirus* strains. Evolutionary rate of the African BtβCoV was set at 1.301×10^{-3}, HPD ($1.064 \times 10^{-3} – 1.434 \times 10^{-3}$), this is slightly higher that the evolutionary rate for the ongoing SARS CoV-2 which has been estimated to have an evolutionary rate of 8.0×10^{-4} (www.nextstrain.org/ncov/global). This is also slightly higher than the evolutionary rate reported for the partial RdRP gene of HuCoV OC43 of 1.06×10^{-4} [34]. A similar topology was also observed for the MCC tree which included βCoV from Asia and Europe, with a MRCA of 1915 (HPD 1880-1950) for all the strains, with the

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.14.056085; this version posted September 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

African strains showing the consistent TMRCA as described above (Figure 3). The African *Norbecovirus* strains seemed to emerge from their parental strain at around the year 1960 (HPD, 1930-1970) this observation supports the hypothesis that *Norbecoviruses* could have been circulating in Africa long before they were first isolated. Whereas for the South African *Marbecovirus* (lineage C) strains seemed to emerge from their parental lineage around the year 1987 (Figure 3), indicating a more recent introduction into Africa, however studies have dated their origin based on partial RdrP gene to as back as 1859 [35].

PHYLOGEOGRAPHY OF Bat β-CoV IN AFRICA

Phylogeographic dispersal of the Bat β-CoV revealed numerous inter-continental spread events from China and Hong Kong into Central Africa (DRC and Kenya), Cameroun in West Africa, and South Africa, and also Mexico and Argentina in the Americas into West Africa Figure 4. These long distance spread events may not necessarily represent actual transmission events, such as inter/intra-host transmission by migrating bat species, as bats have not been known to migrate across the Atlantic Ocean. However reports have shown the possibility of African bats to cover long distances during migration [37]. These observations simply represent genetic similarity and gene flow pathways of BtCoVs, which may due to other factors such as international trade in exotic and wild animals serving as intermediate host of these viruses. There were also dispersal of these viruses between France and some African countries such as Cameroun, South Africa and the DRC. There was also dispersal from Spain into Kenya, South Africa and Madagascar. We also observed dispersal across the Atlantic from Europe to Brazil, as well as across the Indian Ocean from Australia into East and Southern Africa. One limitation to this study is that we were unable to collect consistent data on the bat species from the reference isolates from other continents. Hence the data presented serves as a hypothetical model reflecting genetic dispersal of BtCoV and not species specific movements. The only dispersal event from Italy into Africa was into Nigeria. Studies have shown the potential for African fruit bats to migrate long distances covering thousands of Kilometers, for instance a study using satellite telemetry in Zambia showed that *Elodiun hevium* is capable of covering thousands of kilometers during migration [36]. Another study showed that African bats were capable of migration exceeding 2000km [37]. Intra-continental dispersal events were observed between Cameroun, DRC and South Africa, as well as direct dispersal from Cameroon into Madagascar. The AfrBtCoV strains displayed steady state population demography as depicted by their Bayesian Skyline plot (Figure 5). The population demography reported in this study might not represent the true picture of the virus population, as the dataset utilized in this study is limited by its size and might not represent the true demographic population of BtCoV in Africa.

CONCLUSIONS

We have presented data on the phylodiversity and evolutionary dynamics of Afr-βCoV and their possible dispersal across the continent. Mutiple dispersal pathways were identified between Europe and East/Southern Africa; there were also evidence of spread of BtCoV strains from Asia into Africa. We also identified three CoV sub-genera *Norbecovirus, Hibecovirus* and *Marbecovirus* circulating among African bat species with the probability of inter-species transmission among bats. We also identified multiple corona virus sub-genera co-circulating in

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.14.056085; this version posted September 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

China among the bat specie *Rhinolophus sinicus*, with the capability of zoonotic transmission [32, 33]. Study limitations include the lack of sufficient sequence data in Genebank covering AfrBtCoV, the relatively short genomic fragment analyzed and our inability to analyze spike protein sequence data of these viruses, as a result of paucity of African BtCoV spike protein sequences in established databases; this would have shed more light on their evolution in relation to infectivity and transmission. We have shown the importance of molecular surveillance of viruses with zoonotic potential such as coronaviruses. We advocate for broader trans-continental studies involving full genome sequences of BtCoV to further understand the drivers for their emergence and zoonotic spillovers into human population.

FIGURE LEGEND

Figure 1. Maximum Likelihood Phylogenetic tree, of Bat βCoV from this study (n=219). Grey colored clade indicate alpha CoV from bats, green colored lines indicate lineage D(*Norbecovirus*) βCoV. The clade in red represents lineage B(*Sarbecovirus*) Bat βCoV, while the blue colored clade represents lineage C(*Marbecovirus*) Bat βCoV. The orange cloured labels represents Lineage E(*Embecovirus*),while the violet colured labels represent *Hibercovirus* sub-genus. Bootstrap values are indicated as circles on the branches of the tree, the legend represents the value of the bootstrap where 1 is equivalent to 100%, 0.75 represents 75% 0.25 represents 25% and so on.

Figure 2. Time scaled Maximum clade credibility tree (MCC) with Bayesian phylogeographic reconstruction of partial RdrP gene region of African Bt-βCoV. Branch lengths are proportional to time in years, while branch colors are indicative of country of origin as shown in the legend. The blue horizontal bar represents Lineage D/*Norbecovirus* strains, while red represents Lineage C/*Marbecovirus*. The black arrow indicates *Hibercovirus* (formally under lineage B) strains.

Figure 3. Time scaled Maximum clade credibility tree (MCC) with Bayesian phylogeographic reconstruction of partial RdrP gene region of Bt-βCoV. Branch lengths are proportional to time in years, while branch colors are indicative of country of origin as shown in the legend. Names of bat species of CoV isolation are shown in the tip labels. The blue horizontal bar represents Lineage D/*Norbecovirus* strains, green represents Lineage B/*Sarbecovirus*, while red represents Lineage C/*Marbecovirus*.

Figure 4. Spaciotemporal diffusion of Bt-βCoV, the names of the countires of isolation are written in map. The diamerter of the pink circles is proportional to the number of maximum clade credibility branches (MCC). Concave lines (upward curving) show a clockwise spread movement between two points, while the convex bending (downward curving) lines depict anticlockwise movement between the points.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.14.056085; this version posted September 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Figure 5. Bayesian Skygrid plot, showing effective virus population size estimates over time of Afr Bt-βCoV. The solid blue line represents the median population size, while the upper and lower blue lines represent the 95% high posterior density intervals (95% HPD) for the population.

AUTHORS CONTRIBUTION

B.O.M and O.O.O conceptualized the study, B.OM, and O.O.O, collected data, B.O.M conducted the formal analysis, B.O.M, and O.O.O software, B.O.M wrote the first draft, P.A.A contributed in intellectual discuss all authors read and approved the manuscript.

REFERENCES

1. Masters P.S and Perlman S. Chapter 28, Coronaviridea; In Knipe D. Howley P, Fields Virology 6th edition. Lippincot Williams and Wilkins. 2013.

2. Siddell, S.G., Walker, P.J., Lefkowitz, E.J. *et al.* Additional changes to taxonomy ratified in a special vote by the International Committee on Taxonomy of Viruses (October 2018). *Arch Virol*. 2019.**164,** 943–946.

3. Saif, L.J., 2004. Animal coronaviruses: what can they teach us about the severe acute respiratory syndrome? Rev. Sci. Tech. 23, 643–660.

4. Hamre, D., Procknow, J.J., 1966. A new virus isolated from the human respiratory tract. Proc. Soc. Exp. Biol. Med. 121, 190–193.

5. McIntosh, K., Dees, J.H., Becker, W.B., Kapikian, A.Z., Chanock, R.M., 1967. Recovery in tracheal organ cultures of novel viruses from patients with respiratory disease. Proc. Nat. Acad. Sci. USA 57, 933–940.

6. Lau, S.K., Woo, P.C., Li, K.S., Huang, Y., Tsoi, H.W., Wong, B.H., Wong, S.S., Leung, S.Y., Chan, K.H., Yuen, K.Y.2005. Severe acute respiratory syndrome coronaviruslike virus in Chinese horseshoe bats. Proc. Nat. Acad. Sci. USA 102, 14040–14045.

7. Li, W., Shi, Z., Yu, M., Ren, W., Smith, C., Epstein, J.H., Wang, H., Crameri, G., Hu, Z., Zhang, H., Zhang, J., McEachern, J., Field, H., Daszak, P., Eaton, B.T., Zhang, S.,Wang, L.F., 2005. Bats are natural reservoirs of SARS-like coronaviruses. Science(New York, NY) 310, 676–679.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.14.056085; this version posted September 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

8. Zaki, A.M., van Boheemen, S., Bestebroer, T.M., Osterhaus, A.D., Fouchier, R.A., 2012. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. N.Engl. J. Med. 367, 1814–1820.

9. Annan A, Baldwin HJ, Corman VM, Klose SM, Owusu M, Nkrumah EE,Badu EK, Anti P, Agbenyega O, Meyer B, Oppong S, Sarkodie YA, KalkoEK, Lina PH, Godlevska EV, Reusken C, Seebens A, Gloza-Rausch F, ValloP, Tschapka M, Drosten C, Drexler JF. 2013. Human betacoronavirus 2cEMC/2012-related viruses in bats, Ghana and Europe. Emerg Infect Dis. 19:456–459.

10. Drexler JF, Corman VM, Drosten C. 2014. Ecology, evolution and classification of bat coronaviruses in the aftermath of SARS. Antiviral Res101:45–56.

11. Corman VM, Baldwin HJ, Tateno AF, Zerbinati RM, Annan A, Owusu M,Nkrumah EE, Maganga GD, Oppong S, Adu-Sarkodie Y, Vallo P, da SilvaFilho LV, Leroy EM, Thiel V, van der Hoek L, Poon LL, Tschapka M,Drosten C, Drexler JF. 2015. Evidence for an ancestral association ofhuman coronavirus 229E with bats. J Virol. 89:11858 –11870.

12. Zhou P, Yang XL, Wang XG, Hu B, Zhang L, Zhang W, Si HR, Zhu Y, Li B, Huang CL, Chen HD, Chen J, Luo Y, Guo H, Jiang RD, Liu MQ, Chen Y, Shen XR, Wang X, Zheng XS, Zhao K, Chen QJ, Deng F, Liu LL, Yan B, Zhan FX, Wang YY, Xiao GF, Shi ZL. A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature. 2020 Mar;579 (7798):270-273.

13. Li X, Zai J, Zhao Q, Nie Q, Li Y, Foley BT, Chaillon A. 2020. Evolutionary history, potential intermediate animal host, and cross-species analyses of SARS-CoV-2. J Med Virol. doi: 10.1002/jmv.25731.

14. DeSouza LN, Hieser V, Regamey N, Panning M, Drexler JF, Mulangu S, Poon L, Baumegarte S, Haijema BJ, Kaiser L, Dorsen C. Genetic detection of Coronavirus and differentiation at the prototype strain level by RT-PCR and Nonflourescent low density Microarray. J. Clin. Microbiol. 2007. 45(3): 1049-1052.

15. Quan, P.L., C. Firth, C. Street, J. A. Henriquez, A. Petrosov, et al. 2010. Identification of a severe acute respiratory syndrome coronavirus-like virus in a leaf-nosed bat in Nigeria. mBio 1(4):e00208-10. doi:10.1128/mBio.00208-10.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.14.056085; this version posted September 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

16. Rambaut, A.; Lam, T.T.; Carvalho, L.M.; Pybus, O.G. 2016**.** Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). Virus Evol., 2, vew007.

17. Suchard MA, Lemey P, Baele G. Ayers D.L, DrummondA.J, Rambault. 2018. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. Virus Evolution. 4, vey016. DOI:10.1093/ve/vey016.

18. Hasegawa M, Kishino H, Yano T. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J Mol Evol. 22:160–74. doi:10.1007/BF02101694.

19. Lemey P, Rambaut A, Drummond AJ, Suchard MA. 2009. Bayesian Phylogeography Finds Its Roots. PLoS Comput Biol. 5:e1000520. doi:10.1371/journal.pcbi.1000520.

20. Bielejec F, Baele G, Vrancken B, Suchard MA, Rambaut A, Lemey P. 2016. SpreaD3: interactive visualization of spatiotemporal history and trait evolutionary processes. Mol Biol Evol 33:2167–2169.

21. Leopardi S, Oluwayelu D, Meseko C, Marciano S, Tassoni L, Bakarey S, Monne I, Cattoli G, De Benedicts P. 2016. The close genetic relationship of lineage D Betacoronavirus from Nigerian and Kenyan straw-colored fruit bats (Eidolon helvum) is consistent with the existence of a single epidemiological unit across sub-Saharan Africa. Virus Genes.

22. Tao Y, Tang K, Shi M, Conrardy C, Li KS, Lau SK, Anderson LJ, Tong S. 2012. Genomic characterization of seven distinct bat coronaviruses in Kenya. Virus Res 167:67–73.

23. Wertheim J.O, Chu D.K.W, Peiris J.S, Pond S.L.K, Poon L.L.M. A case for the Ancient origin of Coronaviruses. Journal of Virology May 2013, 87 (12) 7039-7045; DOI: 10.1128/JVI.03273-12

24. Min, J., Cella, E., Ciccozzi, M. Pelosi A. Salemi M, Prosperi M. 2016. The global spread of Middle East respiratory syndrome: an analysis fusing traditional epidemiological tracing and molecular phylodynamics. *glob health res policy* **1,** 14. https://doi.org/10.1186/s41256-016-0014-7.

25. Tao Y, Shi M, Chommanard C, QueenK, Zhang J, Markotter W, Kuzmin IV, Holmes EC, Tong S. 2017. Surveillance of bat coronaviruses in Kenya identifies relatives of human coronaviruses NL63 and 229E and their recombination history. J Virol. 91:e01953-16.

bioRxiv preprint doi: https://doi.org/10.1101/2020.05.14.056085; this version posted September 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

26. Lau SK, Li KS, Tsang AK, Shek CT, Wang M, Choi GKY, Guo R, WongBH, Poon RW, Lam CS, Wang SY, Fan RY, Chan KH, Zheng BJ, WooPC, Yuen KY. 2012. Recent transmission of a novel alphacoronaivurs, batcoronavirus HKU10, from Leschenault's rousettes to Pomona leaf-nosedbats: first evidence of interspecies transmission of coronavirus between bats of different suborders. J. Virol.86:11906–11918.

27. Ar Goulih M, Puechmaille SJ, Diancourt L, Vandenbogaert M**,** Serra-Cobo M, Lopez Roïg M, Brown P, Moutou F, Caro V, Vabret A, Manuguerra J, EPICOREM consortium. 2018. SARS-CoV Related Betacoronavirus and Diverse Alphacoronavirus Members Found in Western Old-World. Virology. 517: 88-97.

28. Sipulwa, L.A., Ongus, J.R., Coldren, R.L. *et al.* Molecular characterization of human coronaviruses and their circulation dynamics in Kenya, 2009–2012. *Virol J* **13,** 18 (2016). https://doi.org/10.1186/s12985-016-0474-x.

29. Abidha CA, Nyiro J, Kamau E, Abulahi O, Kokes DJ Agoti CN. Transmission and evolutionary dynamics of human coronavirus OC43 strains in coastal Kenya investigated by spike protein analysis, 2015-2016. 2020. Virus Evol. 6(1): veaa031. https://doi.org/10.1093/ve/veaa031

30. Wacharapluesadee S., Duengkae P., Rodpan A., Kaewpom T., Maneeorn P., Kanchanasaka B., Yingsakmongkon S., Sittidetboripat N., Chareesaen C., Khlangsap N., Pidthong A., Leadprathom K., Ghai S., Epstein J.H., Daszak P., Olival K.J., Blair P.J., Callahan M.V., Hemachudha T. Diversity of coronavirus in bats from Eastern Thailand. Virol. J. 2015;12:57.

31. Chu DK, Peiris JS, Chen H, Guan Y, Poon LL. Genomic characterizations ofbat coronaviruses (1A, 1B and HKU8) and evidence for co-infections in Miniopterus bats. J Gen Virol. 2008. 89:1282–7.

32. Yuan J, Hon CC, Li Y, Wang D, Xu G, Zhang H, et al. Intraspecies diversity ofSARS-like coronaviruses in Rhinolophus sinicus and its implications for theorigin of SARS coronaviruses in humans. J Gen Virol. 2010;91:1058–62.29.

33. Müller MA, Paweska J, Leman PA, Drosten C, Grywna K, Kemp A, Braack L, Sonnenberg K, Niedrig M, Swanepoel R. Coronaviruis antibodies in African Bat species. Emerg Infec Dis. 2007. 13(9); 1367-70.

34. Vijgen L, Keyaerts E, Lemey P, Maes P, Van Reeth K, Nauwynck H, Pensaert M, Van Ranst M. 2006. Evolutionary history of the closely related group 2 coronaviruses: porcine hemagglutinating encephalomyelitis virus, bovine coronavirus, and human coronavirus OC43. J Virol. 80(14):7270–7274.

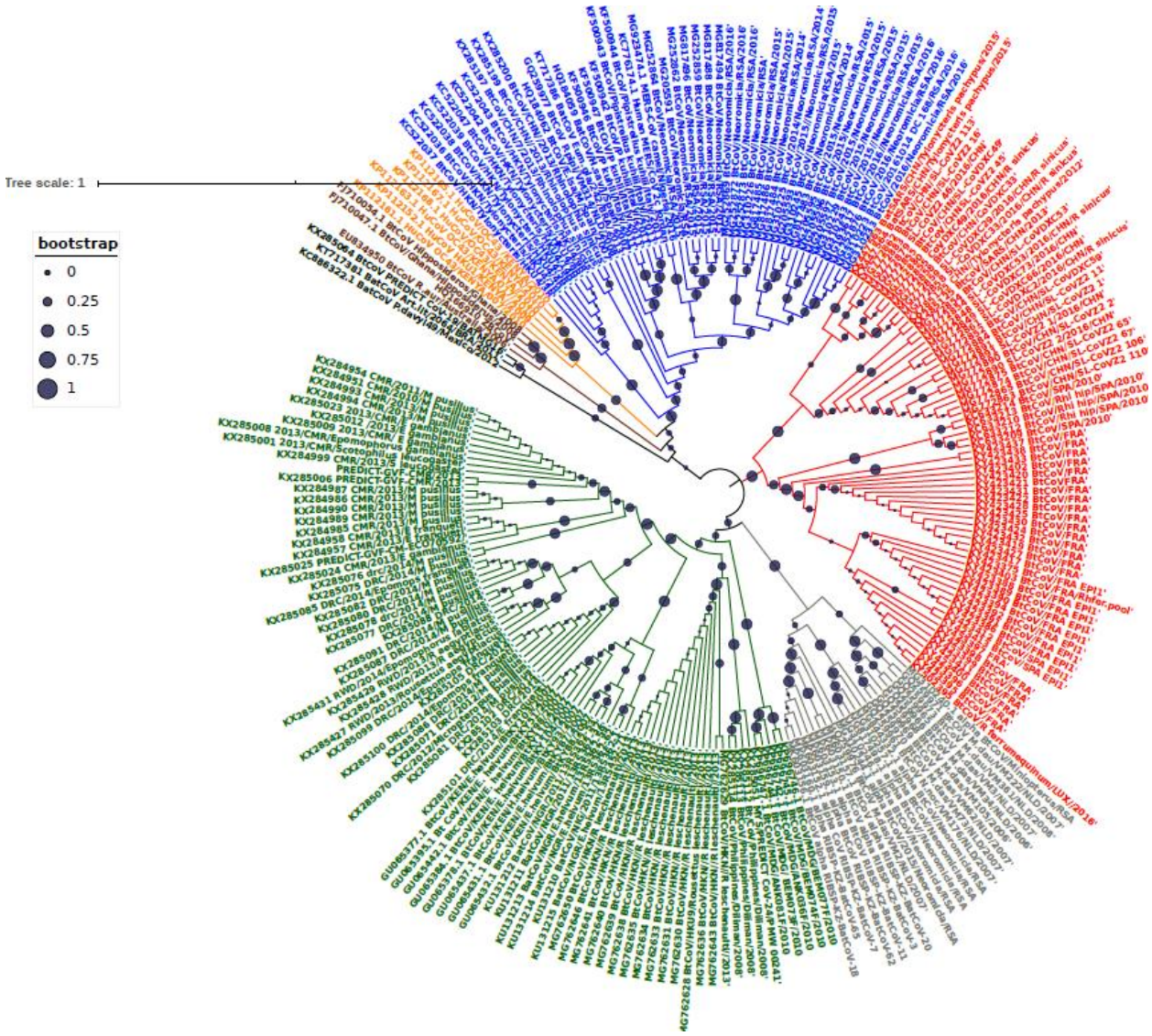
bioRxiv preprint doi: https://doi.org/10.1101/2020.05.14.056085; this version posted September 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

35. Lau S.K, Li K.S, Tsang A.K, Lam C..S, Ahmed S, Chen H, Chan KH, Woo P.C, Yuen K.Y. 2013. Genetic characterization of Betacoronavirus lineage C viruses in bats reveals marked sequence divergence in the spike protein of pipistrellus bat coronavirus HKU5 in Japanese pipistrelle: implications for the origin of the novel Middle East respiratory syndrome coronavirus. J Virol.87(15):8638-50.

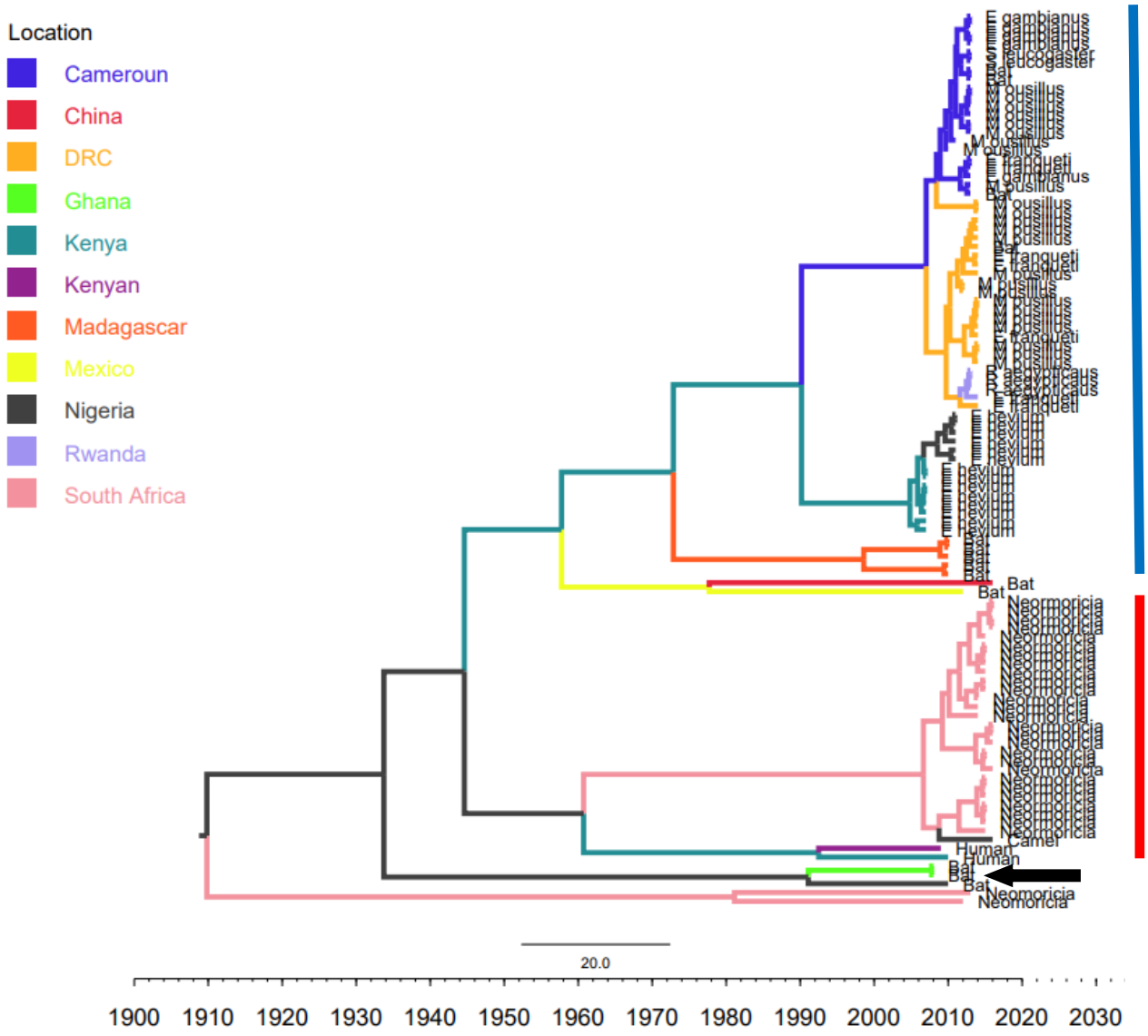
36. Richter H.V, Cumming G.S. 2008. First application of satellite telemetry to track African straw-coloured fruit bat migration. J Zool 275: 172–176.

37. Ossa G, Kramer-Schadt S, Peel A.J, Scharf A.K, Voigt C.C, 2012. The Movement Ecology of the Straw-Colored Fruit Bat, Eidolon helvum, in Sub-Saharan Africa Assessed by Stable Isotope Ratios. PLoS ONE 7(9): e45729. doi:10.1371/journal.pone.0045729.

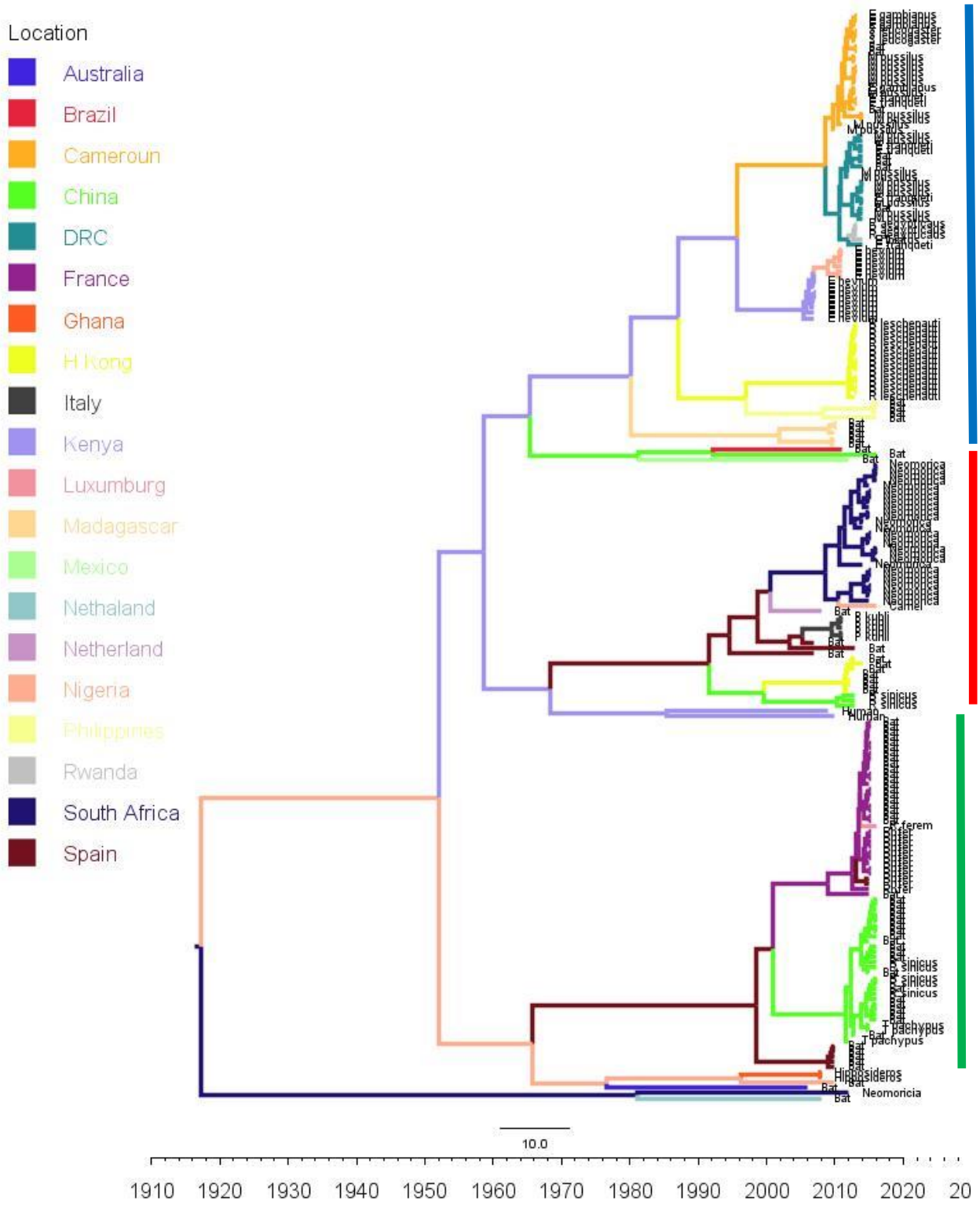
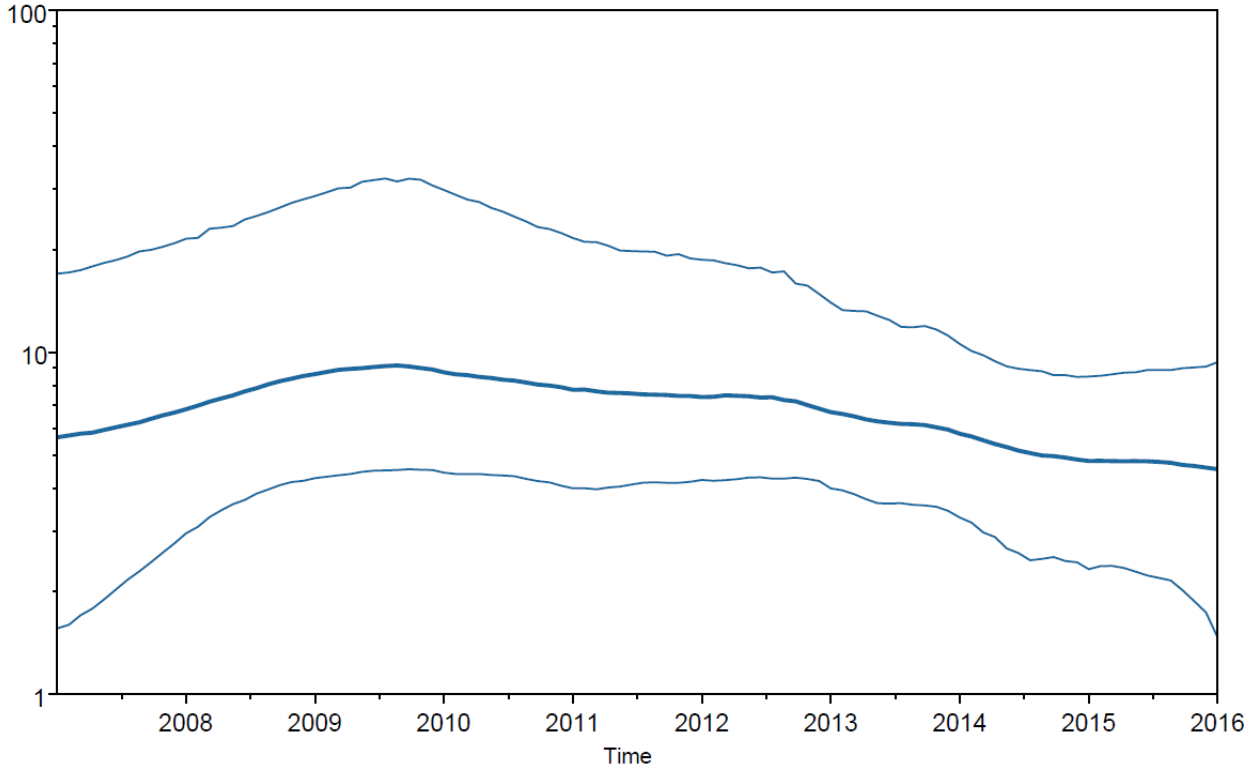
bioRxiv preprint doi: https://doi.org/10.1101/2020.05.14.056085; this version posted September 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.



bioRxiv preprint doi: https://doi.org/10.1101/2020.05.14.056085; this version posted September 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.



bioRxiv preprint doi: https://doi.org/10.1101/2020.05.14.056085; this version posted September 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.



bioRxiv preprint doi: https://doi.org/10.1101/2020.05.14.056085; this version posted September 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.