

Article

The ecological niche of SARS-CoV-2-like viruses in bats, as inferred from phylogeographic analyses of *Rhinolophus* species

Alexandre Hassanin ^{1*}, Vuong Tan Tu ², Manon Curaudeau ³, and Gabor Csorba ⁴

- 1 Institut de Systématique, Évolution, Biodiversité (ISYEB), Sorbonne Université, MNHN, CNRS, EPHE, UA, Paris; alexandre.hassanin@mnhn.fr
- 2 Institute of Ecology and Biological Resources, Vietnam Academy of Science and Technology, No. 18, Hoang Quoc Viet Road, Cau Giay District, Hanoi, Vietnam; tuvuongtan@gmail.com
- 3 Institut de Systématique, Évolution, Biodiversité (ISYEB), Sorbonne Université, MNHN, CNRS, EPHE, UA, Paris; manon.curaudeau1@mnhn.fr
- 4 Department of Zoology, Hungarian Natural History Museum, Baross u. 13, Budapest H-1088, Hungary; csorba.gabor@nhmus.hu

[°] co-first authors

^{*} correspondence: alexandre.hassanin@mnhn.fr; Tel.: +33 1 40 79 56 93.

Abstract: To date, viruses closely related to SARS-CoV-2 have been reported in four bat species: *Rhinolophus acuminatus*, *Rhinolophus affinis*, *Rhinolophus malayanus*, and *Rhinolophus shameli*. Here, we analysed 343 sequences of the mitochondrial cytochrome c oxidase subunit 1 gene (*COI*) from georeferenced bats of the four *Rhinolophus* species identified as reservoirs of SARS-CoV-2-like viruses. Haplotype networks were constructed in order to investigate patterns of genetic diversity among bat populations of Southeast Asia. No strong geographic structure was found for the four *Rhinolophus* species, suggesting high dispersal capacity. The ecological niche of SARS-CoV-2 like viruses was predicted using the four localities of bat SARS-CoV-2-like viruses and the localities where bats showed identical or very similar *COI* haplotypes than virus-positive bats. The ecological niche of SARS-CoV-like viruses was deduced from the localities where bat SARS-CoV-like viruses were previously detected. The results show that the ecological niche of SARS-CoV2-like viruses includes several regions of mainland Southeast Asia whereas that of SARS-CoV-like viruses is mainly restricted to China. In agreement with these results, human populations in Laos, Vietnam, Cambodia, and Thailand appear to be much less affected by the Covid-19 pandemic than other countries of Southeast Asia. In the climatic transitional zone between the two ecological niches (southern Yunnan, northern Laos, northern Vietnam, and possibly Hainan and Taiwan), genomic recombination between highly divergent viruses is more likely to occur. Since recombinant viruses can threaten the benefit of vaccination campaigns, these regions should be under surveillance.

Keywords: coronavirus; horseshoe bats; reservoir hosts; Indochina; China; herd immunity.

1. Introduction

The Severe Acute Respiratory Syndrome coronavirus 2 (SARS-CoV-2) emerged in December 2019 in Wuhan (China) [1]. After 15 months, the coronavirus disease 2019 (Covid-19) pandemic has affected more than 112 million of people around the world, claiming over 2.49 million lives [2]. The origin of SARS-CoV-2 remains enigmatic and many hypotheses have been advanced to explain the first animal-to-human transmission [3].

Within the family Coronaviridae, the subgenus *Sarbecovirus* includes viruses related to SARS-CoV-2 and those related to SARS-CoV, which was responsible for the SARS epidemic in 2002-2004 [4]. The great majority of sarbecoviruses related to SARS-CoV were found in several horseshoe bat species of the genus *Rhinolophus* (family Rhinolophidae),

suggesting that this taxon is the natural reservoir host of these viruses [5]. More recently, several viruses showing between 96.2 and 91.8 % of genome identity with SARS-CoV-2 were identified from fecal samples of four horseshoe bat species: the RaTG13 virus was isolated from a *Rhinolophus affinis* bat collected in Yunnan in 2013 [1]; the RmYN02 virus was found in a *Rhinolophus malayanus* bat sampled in Yunnan in 2019 [6]; two variants of the same virus RshSTT200 (other variant: RshSTT182) were detected in two *Rhinolophus shameli* bats caught in northern Cambodia in 2010 [7]; and five variants of the same virus RacCS203 (other variants: RacCS224, RacCS253, RacCS264, and RacCS271) were sequenced from five *Rhinolophus acuminatus* bats collected in eastern Thailand in 2020 [8]. These data confirmed that the genus *Rhinolophus* is the natural reservoir host of all viruses of the subgenus *Sarbecovirus* [3,5]. The bat species *Rhinolophus acuminatus* and *Rhinolophus shameli* are endemic to Southeast Asia whereas the two other bat species, *Rhinolophus affinis* and *Rhinolophus malayanus*, are distributed in both Southeast Asia and China (Figure 1), suggesting that Southeast Asia is the main region where bat SARS-CoV-2-like viruses diversified.

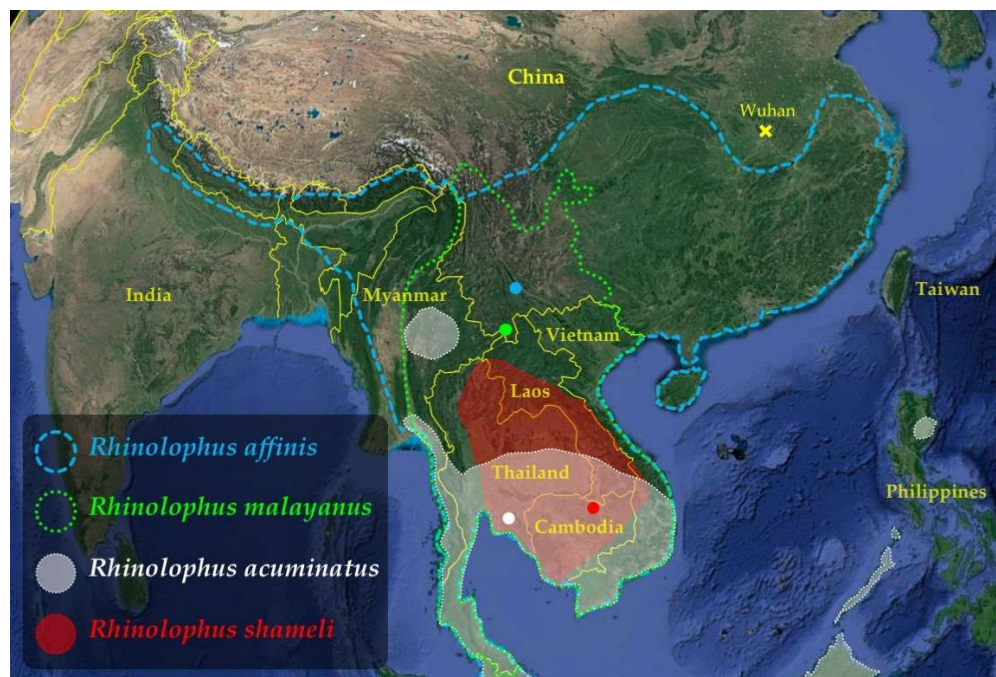


Figure 1. Geographic distribution of the four *Rhinolophus* species found positive for SARS-CoV-2-like viruses in southern China and Southeast Asia. Map from Google Earth US Dept of State Geographer © 2020 Google – Image Landsat / Copernicus - Data SIO, NOAA, U.S. Navy, NGA, GEBCO. For each of the four *Rhinolophus* species, the geographic distribution was extracted from the IUCN website [9]. The coloured dots show the four geographic locations where bats found positive for SARS-CoV-2-like viruses were collected.

In this report, we sequenced the mitochondrial cytochrome c oxidase subunit 1 gene (CO1) from 144 georeferenced bats of the four species *Rhinolophus acuminatus*, *Rhinolophus affinis*, *Rhinolophus malayanus* and *Rhinolophus shameli* collected in several localities of Southeast Asia. The new sequences were aligned to all CO1 sequences available in GenBank for these four species, and haplotype networks were constructed to investigate geographic patterns of genetic diversity among bat populations. The results of these analyses were used to select specific location coordinates to predict the ecological niche of bat SARS-CoV-2-like viruses.

2. Materials and Methods

2.1. DNA extraction and sequencing

One hundred and forty-four tissue samples of morphologically identified bats of *Rhinolophus acuminatus* (n = 10), *Rhinolophus affinis* (n = 57), *Rhinolophus malayanus* (n = 14), and *Rhinolophus shameli* (n = 63) were specially analysed for this study. In the field, bats were captured with mist nets and harp-traps and were measured, photographed and identified by the authors.

Tissue samples were taken from the chest muscles of voucher specimens or from the patagium (biopsy punches; 3 mm diameter) of released bats. Samples were preserved in 95% ethanol.

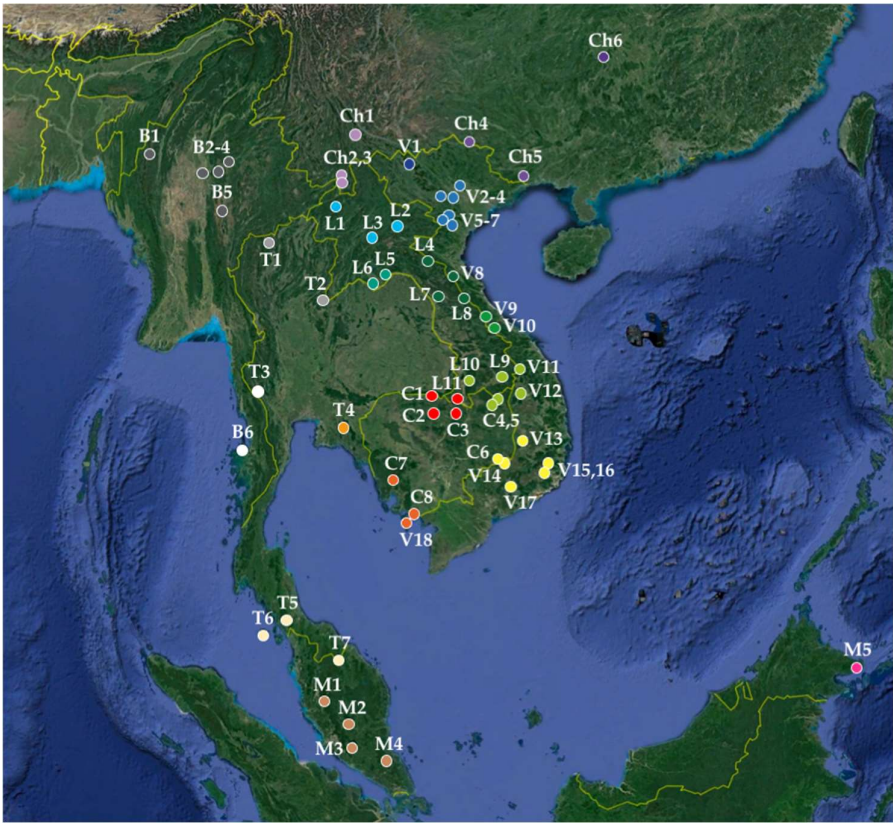


Figure 2. Geographic localities analysed for CO1 haplotype networks. The codes used for the countries are the following: B (Myanmar), C (Cambodia), Ch (China), I (Indonesia), L (Laos), M (Malaysia), T (Thailand), and V (Vietnam). Different regions were highlighted using colours. Map from Google Earth US Dept of State Geographer © 2020 Google – Image Landsat / Copernicus - Data SIO, NOAA, U.S. Navy, NGA, GEBCO. The locality names and GPS are provided in Table S1.

Total DNA was extracted using QIAGEN DNeasy Tissue Kit (Qiagen, Germany) in accordance with the manufacturer’s instructions. The barcode fragment of the CO1 gene (657 bp) was amplified and sequenced using the primers UTyr and C1L705 [10]. PCR amplifications of the CO1 gene were performed as previously published [11]. PCR products were purified using ExoSAP Kit (GE Healthcare, UK) and sequenced using the Sanger method on an ABI 3730 automatic sequencer at the Centre National de Séquençage (Genoscope) in Evry (France). Haplotypes were assembled with forward and reverse electropherograms using Sequencer 5.1 (Gene Codes Corporation, Ann Arbor, MI, USA). Sequences generated for this study were deposited in the GenBank database (accession numbers XXXXXX-XXXXXX; They will be provided during review) (see details in Table S1).

2.2. Analyses of CO1 sequences

Our sequences were aligned with 199 additional *CO1* sequences downloaded from GenBank. Note that the *CO1* sequences of seven bats found positive for SARS-CoV-2-like viruses [1,6,8] were assembled on Geneious® 10.2.2 (Biomatters Ltd., Auckland, New Zealand) by mapping available SRA data to a *CO1* reference. Sequences were aligned using AliView 1.22 [12]. No gaps and stop codons were found in the sequences after translation into amino-acids. Our final alignments contain 37 sequences for *Rhinolophus acuminatus*, 44 sequences for *Rhinolophus malayanus*, 82 sequences for *Rhinolophus shameli*, and 180 sequences for *Rhinolophus affinis*. These four alignments were analysed in PopART 1.5 [13] to construct haplotype networks using the median joining method with equal weights for all mutations. The localities where bats were sampled are shown in the map of Figure 2 and their geographic coordinates are detailed in Table S1.

2.3. Prediction of ecological niches

Van Proosdij *et al.* [14] have estimated that a minimum of 13 records is required to develop accurate distribution models for widespread taxa. However, bat viruses related to SARS-CoV-2 were only detected in four geographic localities (two in Yunnan, one in northern Cambodia, and one in eastern Thailand), which is not enough to predict accurately the ecological niche of the virus. For that reason, we used a genetic approach to increase the number of geographic records. Since the detection of identical haplotypes between different bat populations is indicative of recent gene flow, we also selected the 16 geographic records where bats showed the same *CO1* haplotypes than virus-positive bats (dataset A: 20 points). As an alternative approach, we also included the 10 geographic records where bats showed *CO1* haplotypes different at only one nucleotide position than virus-positive bats (dataset B: 30 points). The *CO1* haplotypes used in datasets A and B are listed in Table S1. For bat SARS-CoV-like viruses, the ecological niche was inferred using GPS data collected for virus published during the last two decades. The list of the 19 available geographic records is provided in Table S2.

For each of the three datasets (datasets A and B for SARS-CoV-2-like viruses; SARS-CoV-like dataset), the 19 bioclimatic variables available in the WorldClim database [15] were studied for an area corresponding to the minimum and maximum latitudes and longitudes of the selected points (20, 30, and 19 points, respectively for the datasets A, B and SARS-CoV-like) and the caret R package [16] was used to determine the least correlated variables ($|r| < 0.7$) [17]. For the two SARS-CoV-2-like datasets, the following seven predictor bioclimatic variables were selected: the mean diurnal range (bio2), the isothermality (bio3), the temperature annual range (bio7), the mean temperature of the warmest quarter (bio10), the precipitation of the wettest month (bio13), the precipitation seasonality (bio15), the precipitation of the driest quarter (bio17), and the precipitation of the warmest quarter (bio18). For the SARS-CoV dataset, the following five predictor bioclimatic variables were retained: the isothermality (bio3), the temperature seasonality (bio4), the maximum temperature of the warmest month (bio5), the precipitation seasonality (bio15), and the precipitation of the warmest quarter (bio18). Ecological niche modelling was performed with the MaxEnt algorithm using ENMTools in R [18]. The MaxEnt approach was chosen for its ability to work with presence-only datasets and to produce results with a low sample size [19]. To test for sampling bias, the distribution model using all selected localities was tested against a null model developed by 1000 times drawing an equal number of random points from the entire study area [20].

3. Results and Discussion

3.1. Genetic analyses of bat species identified as reservoirs of SARS-CoV-2-like viruses

Until now, viruses closely related to SARS-CoV-2 have been found in four bat species of the genus *Rhinolophus*: *R. acuminatus*, *R. affinis*, *R. malayanus*, and *R. shameli*. The haplotype networks constructed using CO1 sequences of these four species are shown in Figure 3. A star-like genetic pattern, characterized by one dominant haplotype and several satellite haplotypes was found for the two bat species endemic to Southeast Asia, i.e. *Rhinolophus acuminatus* and *Rhinolophus shameli*.

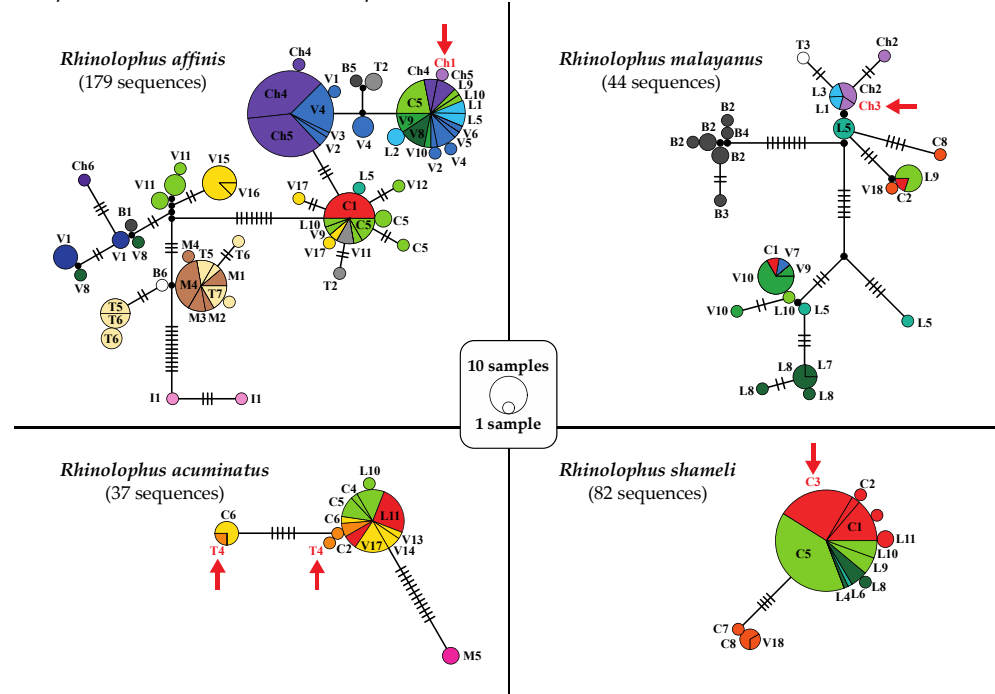


Figure 3. Haplotype networks based on CO1 sequences of the four *Rhinolophus* species found positive for SARS-CoV-2-like viruses. The networks were constructed with the median joining method available in PopART 1.5 [13]. The codes used for the countries are the following: B (Myanmar), C (Cambodia), Ch (China), I (Indonesia), L (Laos), M (Malaysia), T (Thailand), and V (Vietnam). Colours indicate the geographic origin of haplotypes, according to Figure 2 (see Table S1 for details). The circles indicate haplotypes separated by at least one mutation. The black lines on the branches show the number of mutations ≥ 2 . Black circles represent missing haplotypes. Circle size is proportional to the number of haplotypes. Haplogroups separated by more than seven mutations (pairwise nucleotide distances $> 1\%$) are highlighted by dotted lines. The red arrows show the positions of bats found positive for a SARS-CoV-2-like virus.

In the network of *Rhinolophus acuminatus*, the most common haplotype (Rac1) was found in northern Cambodia, southern Laos, eastern Thailand and southern Vietnam, indicating recent gene flow among these populations. Since a virus related to SARS-CoV-2 (91.8% of genome identity), named RacCS203, was detected in five *Rhinolophus acuminatus* bats caught in eastern Thailand in June 2020 [8], the genetic pattern obtained for this species suggests that viruses closely related to RacCS203 may have circulated in most southern regions of mainland Southeast Asia. In contrast, *Rhinolophus acuminatus* bats collected in Borneo (M5) showed a divergent haplotype (separated by 12 mutations), suggesting that the South China Sea between mainland Southeast Asia and Borneo constitutes a barrier to gene flow. Isolated populations of *Rhinolophus acuminatus* described in northern Myanmar, Indonesia (Java and Sumatra) and the Philippines [21] should be further studied.

The network of *Rhinolophus shameli* shows a typical star-like pattern, the most common haplotype (Rsh1) being detected in northern Cambodia and Laos. Since a virus related to SARS-CoV-2 (93.1% of genome identity), named RshSTT200, was recently discovered in two *Rhinolophus shameli* bats collected in northern Cambodia in December 2010 [7],

the genetic pattern obtained for this species suggests that viruses closely related to RshSTT200 may have circulated, at least in the zone between northern Cambodia and central Laos. The bats sampled south to the Tonle Sap lake ($n = 4$; southern Cambodia and Vietnamese island of Phu Quoc) were found to be genetically isolated from northern populations (four mutations). However, further sampling in the south is required to confirm this result, as it may reveal haplotypes identical to those detected in the north.

For the two species distributed in both China and Southeast Asia, i.e. *Rhinolophus affinis* and *Rhinolophus malayanus*, the genetic patterns are more complex with three different haplogroups showing more than 1% of nucleotide divergence. In the network of *Rhinolophus affinis*, there are three major haplogroups (named I, II and III in Figure 3) separated by a minimum of eight mutations. The results are therefore in agreement with those previously published using CO1 and D-loop mitochondrial sequences [22]. The haplotypes detected in the localities sampled in southern China (ch1, ch4, ch5) are distantly related to the single haplotype available for central China (ch6), but they are identical or similar to the haplotypes from Laos, northern and central Vietnam, northern Thailand and northeastern Myanmar. This result suggests recent gene flow between populations from southern Yunnan and those from northern mainland Southeast Asia. Since a virus related to SARS-CoV-2 (96.2% of genome identity), named RaTG13, was detected in one *Rhinolophus affinis* bat captured in southern Yunnan in 2013 [1], the genetic pattern obtained for this species suggests that viruses closely related to RaTG13 may have circulated in the zone comprising southern Yunnan and northern mainland Southeast Asia.

In the network of *Rhinolophus malayanus*, there are three major haplogroups (named I, II and III in Figure 3) separated by a minimum of eight mutations. The haplotypes detected in the localities sampled in southern China (ch2 and ch3) are identical to those found in northern Laos (L1 and L3), suggesting recent gene flow between populations from these two countries. Since a virus related to SARS-CoV-2 (93.7% of genome identity), named RmYN02, was recently isolated from one *Rhinolophus malayanus* bat collected in southern Yunnan in June 2019 [6], the genetic pattern obtained for this species suggests that viruses closely related to RmYN02 may have circulated, at least between southern Yunnan and northern Laos. In contrast, the bats sampled in Myanmar were found to be genetically isolated from other geographic populations (haplogroup II in Figure 3).

3.2. Two different ecological niches for bat viruses related to either SARS-CoV-2 or SARS-CoV

The ecological niche of bat SARS-CoV-2 like viruses was predicted using the four localities where bat SARS-CoV-2-like viruses were previously detected [1,6-8] and the 16 localities where bats showed the same CO1 haplotype than virus-positive bats (dataset A: 20 points; see Table S1 for details). The areas showing the highest probabilities of occurrence (highlighted in green in Figure 4a) include northern and southern Laos, northeastern and southwestern Cambodia, northwestern and southern Vietnam, eastern, northern and western Thailand, southern Myanmar, and western Philippines. As an alternative approach, the ecological niche of bat SARS-CoV-2 like viruses was predicted using 10 additional localities where bats showed similar CO1 haplotypes (differing by only one mutation) to virus-positive bats (dataset B: 30 points; see Table S1 for details). The predicted distribution is presented in Figure 4b. It is very similar to that deduced from dataset A, but several regions showed higher probabilities of occurrence, such as northeastern India, northwestern Myanmar, northeastern Vietnam, central Laos, Hainan and Taiwan, whereas northern Thailand and the Philippines showed lower probabilities of occurrence. The ecological niche predicted for bat SARS-CoV-like viruses (dataset: 19 points; see Table S2 for details) shows a more northern distribution (Figure 4c), as the highest probabilities of occurrence were found in Nepal, Bhutan, Bangladesh, northeastern India, northern Myanmar, northern Vietnam, most regions of China south of the Yellow River, Taiwan, North and South Korea, and southern Japan.

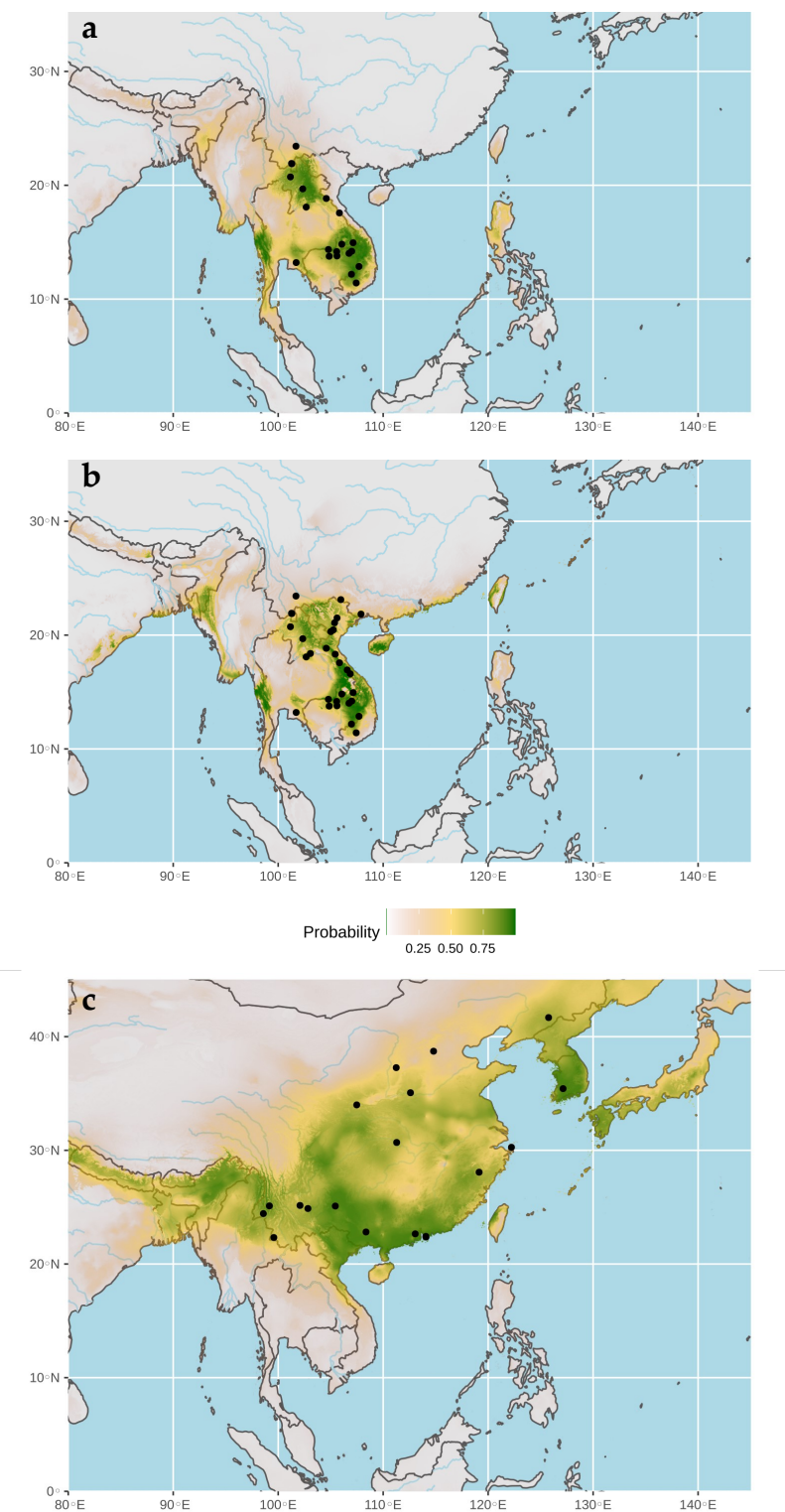


Figure 4. Ecological niches of bat SARS-CoV-2-like viruses (a and b) and bat SARS-CoV-like viruses (c). The geographic distributions of suitable environments were predicted using the Maxent algorithm in ENMTools (see main text for details). Black circles indicate localities used to build the ecological niche model (see geographic coordinates in Tables S1 for SARS-CoV-2-like datasets A and B and Table S2 for SARS-CoV-like dataset).
The ecological niches predicted for bats viruses related to either SARS-CoV-2 or SARS-CoV slightly overlap in the belt zone including southern Yunnan, northern Laos,

and northern Vietnam, as well as possibly the islands of Hainan and Taiwan. This zone corresponds to the northern edge of tropical monsoon climate [23]. Highly divergent sarbecoviruses of the two major lineages are expected to be found in sympatry in this area. This is confirmed by the discovery of viruses related to both SARS-CoV-2 and SARS-CoV lineages in horseshoe bats collected in southern Yunnan [1,6,24]. Collectively, these data suggest that genomic recombination between viruses of the two divergent SARS-CoV and SARS-CoV-2 lineages are more likely to occur in bats roosting, at least seasonally, in the caves of these regions. Since highly recombinant viruses can threaten the benefit of vaccination campaigns, southern Yunnan, northern Laos, and northern Vietnam, as well as possibly the islands of Hainan and Taiwan should be the targets of closer surveillance.

3.3. Mainland Southeast Asia is the cradle of diversification of bat SARS-CoV-2 viruses

Chinese researchers have actively sought sarbecoviruses in all Chinese provinces after the 2002-2004 SARS outbreak. They found many bat viruses related to SARS-CoV [24,25] but only two viruses related to SARS-CoV-2 [1,6] and both of them were discovered in southern Yunnan, the Chinese province bordering Southeast Asia. The ecological niches predicted herein for bat sarbecoviruses suggest that SARS-CoV-2-like viruses are present mostly in Southeast Asia (Figures 4a and 4b) while SARS-CoV-like viruses are dominant in China (Figure 4c). This means that viruses similar to SARS-CoV-2 have been circulating for several decades throughout Southeast Asia, and that different species of bats could have exchanged these viruses in the caves they inhabit.

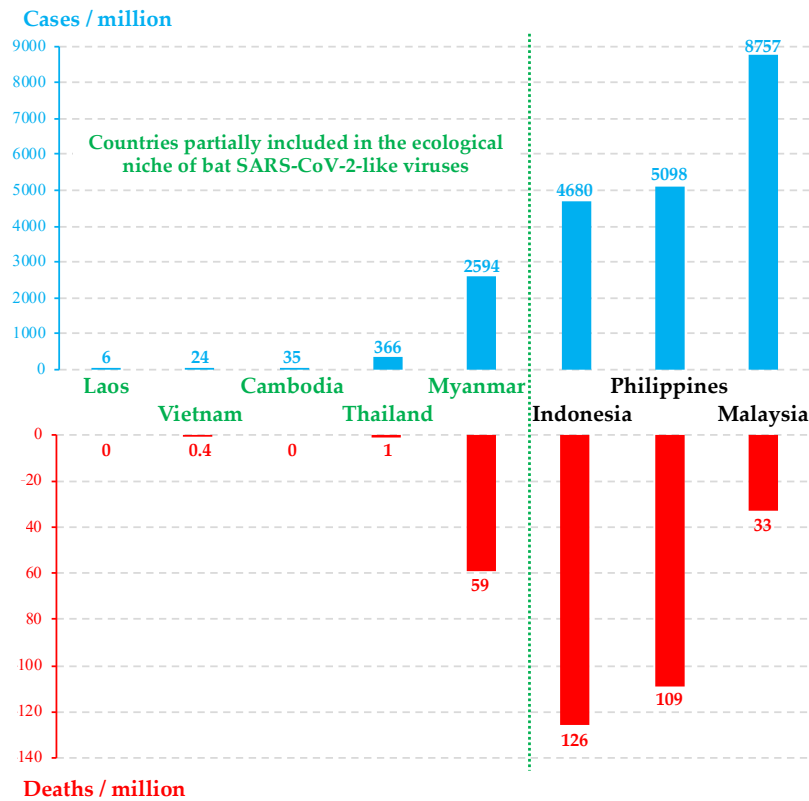


Figure 5. Number of Covid-19 patients per million inhabitants (in blue) and deaths per million inhabitants (in red) for the different countries of Southeast Asia (data extracted from the Worldometers website on February 23, 2021 [2])

The data available on human cases and deaths caused by the Covid-19 pandemic [2] indirectly support the hypothesis that the cradle of diversification of bat viruses related to SARS-CoV-2 is mainland Southeast Asia, and in particular the areas highlighted in green in Figures 4a and 4b. Indeed, human populations in Laos, Vietnam, Cambodia, and

Thailand appear to be much less affected by the Covid-19 pandemic than other countries of the region, such as Myanmar, Malaysia, the Philippines and Indonesia (Figure 5). This suggests that the human populations of these four countries may be benefiting from a level of herd immunity to SARS-CoV-2-like viruses because they have been regularly contaminated by bats and infected secondary hosts such as pangolins.

3.4. Pangolins contaminated by bats in Southeast Asia

Apart from bats, the Sunda pangolin (*Manis javanica*) and Chinese pangolin (*Manis pentadactyla*) are the only wild animals in which SARS-CoV-2-like viruses have been found so far. However, these discoveries were made in a rather special context, that of pangolin trafficking. Several sick pangolins were seized by Chinese customs in Yunnan province in 2017 (unpublished data), in Guangxi province in 2017-2018 [26] and in Guangdong province in 2019 [27]. Even if the viruses sequenced in pangolins are not that close to SARS-CoV-2 (one was 85% identical and the other 90%), they indicate that at least two sarbecoviruses could have been imported into China well before the emergence of Covid-19 epidemic. Indeed, it has been shown that Sunda pangolins collected from different Southeast Asian regions have contaminated each other while in captivity on Chinese territory [3]. The question remained on how the Sunda pangolins became infected initially. Could it have been in their natural Southeast Asian environment, before being captured? The discovery of two new viruses close to SARS-CoV-2 in bats from Cambodia and Thailand [7,8] supports this hypothesis, as *Rhinolophus* bats and pangolins can meet, at least occasionally, in forests of Southeast Asia, possibly in caves, tree hollows or burrows. Further substantiating this hypothesis, the geographic distribution of *Manis javanica* [28] overlaps the ecological niche here predicted for bat SARS-CoV-2 viruses (Figures 4a and 4b), and SARS-CoV-2 neutralizing antibodies have been recently detected in a Sunda pangolin collected in eastern Thailand [8]. Collectively, these data strengthen the hypothesis that pangolin trafficking is responsible for multiple exports of SARS-CoV-2-like viruses to China [3].

Supplementary Materials: The following are available online at <http://www.mdpi.com/>, Table S1. Cytochrome c oxidase subunit 1 (CO1) haplotypes of *Rhinolophus* species analysed in this study, Table S2: Geographic records used to infer the ecological niche of bat SARS-CoV-like viruses.

Author Contributions: Conceptualization, A.H.; methodology, A.H. and M.C.; formal analysis, A.H., M.C., and V.T.T.; resources, A.H., G.C., and V.T.T.; writing—original draft preparation, A.H.; writing—review and editing, A.H., G.C., M.C., and V.T.T.; funding acquisition, A.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the AAP RA-COVID-19, grant number XXXXX (the number will be provided soon), the CNRS, the MNHN, the INRA and the CEA (Genoscope).

Institutional Review Board Statement: Ethical review and approval were not available for this study because the field missions were carried out between 2004 and 2011, i.e., before the creation of the ethical committee at the *Muséum national d'Histoire naturelle*.

Informed Consent Statement: Not applicable

Data Availability Statement: DNA sequences generated for this study were deposited in the GenBank database (accession numbers XXXXXX-XXXXXX; They will be provided during review).

Acknowledgments: In Cambodia, we would like to acknowledge T.D. Vong, V. Hul and S. In for their support during the field surveys, and C. Garnerio-Morena, G. Diffloth, B. Kilian, A. Beschaouch and N. Hayashi-Denis for their assistance. The field mission was organized by UNESCO under the permissions of the National Authority for Preah Vihear, the President, His Excellency Mr. Chuch Phoeurn and the Director of the Preah Vihear temple, Prof. Hang Soth. In Laos, we would like to thank N. Beevang, J. Fuchs, E. Ifticene, and S. Xaythany for their support during the field surveys, and C. Bear Khue, F. Bouvier, C. Hatten, R. Humphrey, A.M. Ohler, P. Rousseau, Mr. Tipavanh, M. Thoummabouth, K. Valakone, Y. Varelides, and A.G. Verdier for their assistance. The fieldwork was supported and organized by the UNESCO and « La Maison du Patrimoine » (World Heritage House) of Luang Prabang, and the Phongsaly Forest Conservation and Rural Development Project,

a Lao-European cooperation. Survey permission and authorization for tissue samples collecting were granted by the Ministry of Agriculture and Forestry, Department of Livestock and Fisheries. In Vietnam, we would like to acknowledge P.D. Tien (IEBR), D.Q. Thang and N.X. Nghia (Ngoc Linh NR) and N.T. Son (Vu Quang NP) for their support during the field surveys. We are grateful to L.X. Canh, T.H. Thai, N.V. Sinh and other colleagues of the IEBR, Hanoi for administrative assistance. The field research was done under the permissions of the People's Committees of numerous provinces and the Vietnamese Ministry of Agriculture and Rural Development (Vietnam Administration of Forestry). The fieldwork was also supported by the "Société des Amis du Muséum et du Jardin des Plantes" and the National Research, Development and Innovation Fund of Hungary NKFIH KH130360. We are grateful to C. Bonillo, C. Ferreira, J. Lambourdière, and J. Utge (UMS 2700, MNHN) for their technical assistance. We would like to thank Huw Jones and Anne Ropiquet for helpful comments on the first version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Zhou, P.; Yang, X.L.; Wang, X.G.; Hu, B.; Zhang, L.; Zhang, W.; Si, H.R.; Zhu, Y.; Li, B.; Huang, C.L.; et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*, **2020**, *579*, 270-273. doi:10.1038/s41586-020-2012-7.
2. COVID-19 Coronavirus pandemic. <https://www.worldometers.info/coronavirus/>. Downloaded on 23 February 2021.
3. Hassanin, A.; Grandcolas, P.; Veron, G. Covid-19: natural or anthropic origin? *Mammalia* **2021**, *85*, 1-7. doi:10.1515/mammalia-2020-0044
4. Walker, P.J.; Siddell, S.G.; Lefkowitz, E.J.; Mushegian, A.R.; Adriaenssens, E.M.; Dempsey D.M.; Dutilh, B.E.; Harrach, B.; Harrison, R.L.; Hendrickson, R.C.; et al. Changes to virus taxonomy and the statutes ratified by the international committee on taxonomy of viruses. *Arch Virol.*, **2020**, *165*, 2737-2748. doi:10.1007/s00705-020-04752-x
5. Li, W.; Shi, Z.; Yu, M.; Ren, W.; Smith, C.; Epstein, J.H.; Wang, H.; Crameri, G.; Hu, Z.; Zhang, H.; et al. Bats are natural reservoirs of SARS-like coronaviruses. *Science*, **2005**, *310*, 676-679. doi:10.1126/science.1118391
6. Zhou, H.; Chen, X.; Hu, T.; Li, J.; Song, H.; Liu, Y.; Wang, P.; Liu, D.; Yang, J.; Holmes, E.C.; et al. A novel bat coronavirus closely related to SARS-CoV-2 contains natural insertions at the S1/S2 cleavage site of the spike protein. *Curr. Biol.*, **2020**, *30*, 3896. doi:10.1016/j.cub.2020.09.030.
7. Hul, V.; Delaune, D.; Karlsson E.A.; Hassanin A.; Tey P.O.; Baidaliuk A.; Gámbaro F.; Tu V.T.; Keatts L.; Mazet J.; et al. A novel SARS-CoV-2 related coronavirus in bats from Cambodia. doi:10.1101/2021.01.26.428212
8. Wacharapluesadee, S.; Tan, C.W.; Maneerorn, P.; Duengkhae, P.; Zhu, F.; Joyjinda, Y.; Kaewpom, T.; Chia, W.N.; Ampoot, W.; Lim, B. L.; et al. Evidence for SARS-CoV-2 related coronaviruses circulating in bats and pangolins in Southeast Asia. *Nat. Commun.*, **2021**, *12*, 972. doi:10.1038/s41467-021-21240-1
9. IUCN 2020. The IUCN Red List of Threatened Species. Version 2020-3. <https://www.iucnredlist.org>. Downloaded on 23 February 2021.
10. Hassanin, A.; Delsuc, F.; Ropiquet, A.; Hammer, C.; Jansen van Vuuren, B.; Matthee, C.; Ruiz-Garcia, M.; Catzeflis, F.; Areskoug, V.; Nguyen, T.T.; Couloux, A. Pattern and timing of diversification of Cetartiodactyla (Mammalia, Laurasiatheria), as revealed by a comprehensive analysis of mitochondrial genomes. *C.R. biol.*, **2012**, *335*, 32-50. doi:10.1016/j.crvi.2011.11.002
11. Tu, V.T.; Hassanin, A.; Görföl, T.; Arail, S.; Fukui, D.; Thanh, H.T.; Son, N.T.; Furey, N.M.; Csorba, G. Integrative taxonomy of the *Rhinolophus macrotis* complex (Chiroptera, Rhinolophidae) in Vietnam and nearby regions. *J. Zool. Syst. Evol. Res.*, **2017**, *55*, 177-198. doi:10.1111/jzs.12169
12. Larsson, A. AliView: a fast and lightweight alignment viewer and editor for large data sets. *Bioinformatics*, **2014**, *30*, 3276-3278. doi:10.1093/bioinformatics/btu531
13. Leigh, J.W.; Bryant, D. PopART: Full-feature software for haplotype network construction. *Methods Ecol. Evol.*, **2015**, *6*, 1110-1116. doi: 10.1111/2041-210X.12410
14. van Proosdij, A.S.J.; Sosef, M.S.M.; Wieringa, J.J.; Raes N. Minimum required number of specimen records to develop accurate species distribution models. *Ecography*, **2016**, *39*, 542-552. doi:10.5061/dryad.8sb8v.
15. Hijmans, R.J.; Cameron, S.E.; Parra, J.L.; Jones, P.G.; Jarvis, A. Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.*, **2005**, *25*, 1965-1978. doi:10.1002/joc.1276
16. Kuhn, M. Building predictive models in R using the caret package. *J. Stat. Softw.*, **2008**, *28*, 1-26. doi:10.18637/jss.v028.i05
17. Dormann, C.F.; Elith, J.; Bacher, S.; Buchmann, C.; Carl, G.; Carré, G.; Marquéz, J.R.G.; Gruber, B.; Lafourcade, B.; Leitao, P.J.; Münkemüller, T. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, **2013**, *36*, 27-46. doi:10.1111/j.1600-0587.2012.07348.x
18. Warren, D.L.; Matzke, N.J.; Cardillo, M.; Baumgartner, J.B.; Beaumont, L.J.; Turelli, M.; Glor, R.E.; Huron, N.A.; Simões, M.; Iglesias, T.L.; et al. ENMTools 1.0: an R package for comparative ecological biogeography. *Ecography*, **2021**, doi:10.1111/ecog.05485
19. Wisz, M.S.; Hijmans, R.J.; Li, J.; Peterson, A.T.; Graham, C.H.; Guisan, A. Effects of sample size on the performance of species distribution models. *Divers. Distrib.*, **2008**, *14*, 763-773. doi:10.1111/j.1472-4642.2008.00482.x

20. Raes, N.; ter Steege, H. A null-model for significance testing of presence-only species distribution models. *Ecography*, **2007**, *30*, 727-736. doi:10.1111/j.2007.0906-7590.05041.x
21. Thong, V.D.; Thanh, H.T.; Soisook, P.; Csorba, G. *Rhinolophus acuminatus*. The IUCN Red List of Threatened Species **2019**, e.T19520A21974227. doi:10.2305/IUCN.UK.2019-3.RLTS.T19520A21974227.en.
22. Ith, S.; Bumrungsri, S.; Furey, N.M.; Bates, P.J.; Wonglapsuwan, M.; Khan, F.; Thong, V.D.; Soisook, P.; Satasook, C.; Thomas, N.M. Taxonomic implications of geographical variation in *Rhinolophus affinis* (Chiroptera: Rhinolophidae) in mainland South-east Asia. *Zool. Stud.*, **2015**, *54*, e31. doi:10.1186/s40555-015-0109-8
23. Song, Y.; Achberger, C.; Linderholm, H.W. Rain-season trends in precipitation and their effect in different climate regions of China during 1961–2008. *Environ. Res. Lett.*, **2011**, *6*, 034025. doi:10.1088/1748-9326/6/3/034025
24. Hu, B.; Zeng, L.P.; Yang, X.L.; Ge, X.Y.; Zhang, W.; Li, B.; Xie, J.Z.; Shen, X.R.; Zhang, Y.Z.; Wang, N.; et al. Discovery of a rich gene pool of bat SARS-related coronaviruses provides new insights into the origin of SARS coronavirus. *PLoS Pathog.*, **2017**, *13*, e1006698. doi:10.1371/journal.ppat.1006698
25. Fan, Y.; Zhao, K.; Shi, Z. L.; Zhou, P. Bat Coronaviruses in China. *Viruses*, **2019**, *11*, 210. doi:10.3390/v11030210
26. Lam, T.T.; Jia, N.; Zhang, Y.W.; Shum, M.H.; Jiang, J.F.; Zhu, H.C.; Tong, Y.G.; Shi, Y.X.; Ni, X.B.; Liao, Y.S.; et al. Identifying SARS-CoV-2-related coronaviruses in Malayan pangolins. *Nature*, **2020**, *583*, 282-285. doi:10.1038/s41586-020-2169-0
27. Liu, P.; Chen, W.; Chen, J.P. Viral metagenomics revealed Sendai virus and coronavirus infection of Malayan pangolins (*Manis javanica*). *Viruses*, **2019**, *11*, 979. doi:10.3390/v11110979
28. Challender, D.; Willcox, D.H.A.; Panjang, E.; Lim, N.; Nash, H.; Heinrich, S.; Chong, J. 2019. *Manis javanica*. The IUCN Red List of Threatened Species 2019: e.T12763A123584856. <https://dx.doi.org/10.2305/IUCN.UK.2019-3.RLTS.T12763A123584856.en>. Downloaded on 23 February 2021.