nature research

Corresponding author(s): Peter Daszak, Zhengli Shi

Last updated by author(s): May 25, 2020

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist .

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☒	☐	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☐	☒	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No software was used for data collection.
Data analysis	Nucleotide sequences were aligned using MUSCLE 3.8.31. Bayesian phylogenetic analysis were performed in BEAST 1.8.4. We used TEMPEST 1.5. to assess the temporal structure within our datasets. Convergence of the MCMC runs was confirmed using Tracer 1.6. Maximum clade credibility (MCC) tree annotated with discrete traits were generated in TreeAnnotator 1.8.4. and visualized using the software SpreaD3 0.9.6. The Mean Phylogenetic Distance (MPD) and the Mean Nearest Taxon Distance (MNTD) statistics were calculated in R 3.5.1. Mantel tests were performed in ARLEQUIN 3.5. Median-joining network was reconstructed in Network 10.0.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All sequences data generated for this project have been deposited in GenBank, the NIH genetic sequence database (https://www.ncbi.nlm.nih.gov/genbank/). All accession numbers are available in the supplementary material Note 1.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☐ Behavioural & social sciences ☒ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	We studied the diversity of bat coronaviruses in China using sequences generated from our own field-collected samples (oral and rectal swabs) and sequences available in GenBank to infer their cross-species transmission history and spatial spread.
Research sample	Our final datasets include 781 bat coronavirus sequences generated for this study and 490 sequences from GenBank (https://www.ncbi.nlm.nih.gov/genbank/) (list of GenBank accession numbers available in Supplementary Material). These bat CoV sequences were obtained from field samples collected on adult bats in China. Samples collected were oral and rectal swabs and fecal pellets to minimize the impact on bat populations. These kinds of samples are suitable for the detection of coronaviruses in wild animals. These bats belong to 50 species: Cynopterus sphinx, Eonycteris spelaea, Megaerops sp., Rousettus leschenaultii, Rousettus sp., Aselliscus stoliczkanus, Hipposideros armiger, Hipposideros cineraceus, Hipposideros pomona, Hipposideros pratti, Rhinolophus affinis, Rhinolophus ferrumequinum, Rhinolophus hipposideros, Rhinolophus macrotis, Rhinolophus malayanus, Rhinolophus monoceros, Rhinolophus pearsonii, Rhinolophus pusillus, Rhinolophus rex, Rhinolophus sinicus, Rhinolophus stheno, Rhinolophus thomasi, Rhinolophus sp., Miniopterus schreibersii, Miniopterus fuscus, Miniopterus magnater, Miniopterus pusillus, Miniopterus sp., Eptesicus serotinus, Hypsugo sp., Ia io, Murina leucogaster, Myotis chinensis, Myotis daubentonii, Myotis davidii, Myotis fimbriatus, Myotis horsfieldii, Myotis myotis, Myotis pequinius, Myotis ricketti, Myotis siligorensis, Myotis sp., Nyctalus plancyi, Pipistrellus abramus, Pipistrellus pipistrellus, Pipistrellus sp., Scotophilus kuhlii, Tylonycteris pachypus, Tylonycteris robustula, Vespertilio sinensis. Additionally, we accessed sequences from GenBank and GISAID from humans (SARS-CoV, SARS-CoV-2) and pangolins (Manis javanica).
Sampling strategy	Samples were collected to cover a large proportion of the bat taxonomic diversity in China including 50 species: Cynopterus sphinx, Eonycteris spelaea, Megaerops sp., Rousettus leschenaultii, Rousettus sp., Aselliscus stoliczkanus, Hipposideros armiger, Hipposideros cineraceus, Hipposideros pomona, Hipposideros pratti, Rhinolophus affinis, Rhinolophus ferrumequinum, Rhinolophus hipposideros, Rhinolophus macrotis, Rhinolophus monoceros, Rhinolophus pearsonii, Rhinolophus pusillus, Rhinolophus rex, Rhinolophus sinicus, Rhinolophus stheno, Rhinolophus thomasi, Rhinolophus sp., Miniopterus schreibersii, Miniopterus fuscus, Miniopterus magnater, Miniopterus pusillus, Miniopterus sp., Eptesicus serotinus, Hypsugo sp., Ia io, Murina leucogaster, Myotis chinensis, Myotis daubentonii, Myotis davidii, Myotis fimbriatus, Myotis horsfieldii, Myotis myotis, Myotis pequinius, Myotis ricketti, Myotis siligorensis, Myotis sp., Nyctalus plancyi, Pipistrellus abramus, Pipistrellus pipistrellus, Pipistrellus sp., Scotophilus kuhlii, Tylonycteris pachypus, Tylonycteris robustula, Vespertilio sinensis. No sample size calculation was performed as this was not required for this study. Bats were captured using mist nets at their roost site or feeding areas. Each captured bat was stored into a cotton bag, all sampling was non-lethal and bats were released at the site of capture immediately after sample collection. Oral and fecal swabs were collected. These kinds of samples are suitable for the detection of coronaviruses in wild animals.
Data collection	Field samples were collected by experienced wildlife zoologists using standard procedures approved by Tufts University IACUC committee (proposal #G2017-32) and Wuhan Institute of Virology Chinese Academy of Sciences IACUC committee (proposal WIVA05201705). Bats were captured using mist nets at their roost site or feeding areas. Each captured bat was stored into a cotton bag, all sampling was non-lethal and bats were released at the site of capture immediately after sample collection. Oral and fecal swabs were collected. RNA was extracted and tested for the presence of coronaviruses as described in the Methods.
Timing and spatial scale	Our samples were collected in 15 Chinese provinces (Anhui, Beijing, Guangdong, Guangxi, Guizhou, Hainan, Henan, Hubei, Hunan, Jiangxi, Macau, Shanxi, Sichuan, Yunnan, and Zhejiang) from December 2010 to June 2015 when bats were not hibernating. Samples were collected each month in different locations.
Data exclusions	Positive results detected in bat genera that were not known to harbor a specific CoV lineage previously were repeated a second time (PCR + sequencing) as a confirmation. Field species identifications were also confirmed and re-confirmed by cytochrome (cytb) DNA barcoding using DNA extracted from the feces or swabs. Only viral detection and barcoding results confirmed at least twice were included in this study. Viral detection in bat genera that were not known to harbor a specific CoV lineage previously and that were not reconfirmed at least twice were excluded from the study.
Reproducibility	Positive results detected in bat genera that were not known to harbor a specific CoV lineage previously were repeated a second time (PCR + sequencing) as a confirmation. Field species identifications were also confirmed and re-confirmed by cytochrome (cytb) DNA barcoding using DNA extracted from the feces or swabs. Only viral detection and barcoding results confirmed at least twice were included in this study. All sequences data generated for this project have been deposited in GenBank.
Randomization	Our random data subsets were created by randomly selecting a smaller number of sequence, without replacement, from each host or province in Excel. Experimental groups were not created for this study.
Blinding	No blinding was used. This was not necessary for our experiment design given that no observer-biased measurement were taken and only PCR lab experiments, which don't require blinding, were performed for this study.

Did the study involve field work? ☒ Yes ☐ No

Field work, collection and transport

Field conditions	Environmental conditions (rainfall, temperature) were not recorded during field work as this data was not necessary for our study.
Location	Several locations in 15 Chinese provinces (Anhui 31°50′ N 117°0′E, Beijing 39°55′N 116°23′E, Guangdong 23°24′N 113°30′E, Guangxi 23.6°N 108.3°E, Guizhou 26°50′N 106°50′E, Hainan 19°12′N 109°42′E, Henan 33°54′N 113°30′E, Hubei 31°12′N 112°18′E, Hunan 28°06′46″N 112°59′00″E, Jiangxi 27°18′N 116°00′E, Macau 22°12′N 113°33′E, Shanxi 37°42′N 112°24′E, Sichuan 30°08′N 102°56′E, Yunnan 25°03′N 101°52′E, and Zhejiang 29°12′N 120°30′E). Bats were captured using mist nets at their roost site or feeding areas. Details of the sampling sites for each sample are given in the GenBank accession data.
Access and import/export	No import/export permits were needed as all field samples were collected and then tested for coronaviruses in China.
Disturbance	Disturbance in bat colony was minimized and limited to a few hours during a single day at each sampling site. Samples were collected by experienced wildlife zoologists. Each captured bat was stored into a cotton bag, all sampling was non-lethal and bats were released at the site of capture immediately after sample collection. No bats were killed for this study.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☐	☒ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	No laboratory animals were used in this study
Wild animals	Adult bats were captured using mist nets at their roost site or feeding areas. These bats belong to 50 species: Cynopterus sphinx, Eonycteris spelaea, Megaerops sp.,Rousettus leschenaultii, Rousettus sp., Aselliscus stoliczkanus, Hipposideros armiger, Hipposideros cineraceus, Hipposideros pomona, Hipposideros pratti, Rhinolophus affinis, Rhinolophus ferrumequinum, Rhinolophus hipposideros, Rhinolophus macrotis, Rhinolophus monoceros, Rhinolophus pearsonii, Rhinolophus pusillus, Rhinolophus rex, Rhinolophus sinicus, Rhinolophus stheno, Rhinolophus thomasi, Rhinolophus sp., Miniopterus schreibersii, Miniopterus fuscus, Miniopterus magnater, Miniopterus pusillus, Miniopterus sp., Eptesicus serotinus, Hypsugo sp., Ia io, Murina leucogaster, Myotis chinensis, Myotis daubentonii, Myotis davidii, Myotis fimbriatus, Myotis horsfieldii, Myotis myotis, Myotis pequinius, Myotis ricketti, Myotis siligorensis, Myotis sp., Nyctalus plancyi, Pipistrellus abramus, Pipistrellus pipistrellus, Pipistrellus sp., Scotophilus kuhlii, Tylonycteris pachypus, Tylonycteris robustula, Vespertilio sinensis. Each captured bat was stored into a cotton bag, all sampling was non-lethal and bats were released at the site of capture immediately after sample collection. Oral and fecal swabs were collected. A wing punch was also collected for barcoding purpose. Bat handling methods were approved by Tufts University IACUC committee (proposal #G2017-32) and Wuhan Institute of Virology Chinese Academy of Sciences IACUC committee (proposal WIVA05201705). No bats were killed for this study.
Field-collected samples	This study did not involve laboratory experiments on field-caught animals and samples.
Ethics oversight	Bat handling methods were approved by Tufts University IACUC committee (proposal #G2017-32) and Wuhan Institute of Virology Chinese Academy of Sciences IACUC committee (proposal WIVA05201705).

Note that full information on the approval of the study protocol must also be provided in the manuscript.