





A Critical Analysis of the Evidence for the SARS-CoV-2 Origin Hypotheses

James C. Alwine,^{a,b} Arturo Casadevall,^c Editor in Chief, *mBio*,  Lynn W. Enquist,^{b,d}
 Felicia D. Goodrum,^{b,e} Editor in Chief, *Journal of Virology*, Michael J. Imperiale,^f Editor in Chief, *mSphere*

^aDepartment of Cancer Biology, University of Pennsylvania, Philadelphia, Pennsylvania, USA

^bDepartment of Immunobiology, University of Arizona, Tucson, Arizona, USA

^cBloomberg School of Public Health, Johns Hopkins University, Baltimore, Maryland, USA

^dDepartment of Molecular Biology, Princeton University, Princeton, New Jersey, USA

^eBIO5 Institute, University of Arizona, Tucson, Arizona, USA

^fDepartment of Microbiology and Immunology, University of Michigan, Ann Arbor, Michigan, USA

ABSTRACT When humans experience a new, devastating viral infection such as severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), significant challenges arise. How should individuals as well as societies respond to the situation? One of the primary questions concerns the origin of the SARS-CoV-2 virus that infected and was transmitted efficiently among humans, resulting in a pandemic. At first glance, the question appears straightforward to answer. However, the origin of SARS-CoV-2 has been the topic of substantial debate primarily because we do not have access to some relevant data. At least two major hypotheses have been suggested: a natural origin through zoonosis followed by sustained human-to-human spread or the introduction of a natural virus into humans from a laboratory source. Here, we summarize the scientific evidence that informs this debate to provide our fellow scientists and the public with the tools to join the discussion in a constructive and informed manner. Our goal is to dissect the evidence to make it more accessible to those interested in this important problem. The engagement of a broad representation of scientists is critical to ensure that the public and policy-makers can draw on relevant expertise in navigating this controversy.

KEYWORDS COVID-19, SARS-CoV-2, zoonosis, coronavirus, lab leak, origin

The Merriam-Webster online dictionary defines a hypothesis as a tentative assumption made to draw out and test its logical or empirical consequences. Hypotheses are part of the scientific method and are formulated to be tested by well-designed, controlled experiments that will support or refute them. The most powerful experiment is one that has the potential to eliminate a hypothesis or to reveal a deep flaw in it. Many experiments lack this potential and at best can only provide data that are consistent with a given hypothesis. While coincidences, correlations, or temporal sequences of events can be useful for formulating a hypothesis, they leave room for alternative explanations and are not conclusive. In the case of the origin of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the primary candidate hypotheses are not readily falsifiable with the currently available information; thus, we must acknowledge that a definitive determination of the origin is unlikely.

Two major hypotheses regarding the origin of SARS-CoV-2 have been debated: a direct zoonotic origin and the introduction of the virus into humans from a laboratory source (1–5). The zoonosis hypothesis states that the virus originated naturally from an animal-to-human spillover event followed by sustained human-to-human transmission. This scenario has happened repeatedly over time, causing epidemics and pandemics, including influenza, SARS, Middle East respiratory syndrome (MERS), Ebola, HIV, as well

Copyright © 2023 Alwine et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Felicia D. Goodrum, fgoodrum@arizona.edu.

The authors declare no conflict of interest.

Ed. Note: This editorial is being published by the following ASM journals: *Journal of Virology*, *mBio*, and *mSphere*.

The views expressed in this article do not necessarily reflect the views of the journal or of ASM.

Published 28 March 2023

as bacterial infections like bubonic plague. Viral emergence events through zoonoses often occur when host species alter their geographical range in response to environmental changes, habitat destruction, or urbanization. Hot spots for emergence in the human population exist in regions with high human population densities and frequent contact between humans and wildlife or agricultural species that harbor zoonotic pathogens. Only a small proportion of viruses that spill over into humans result in disease, and even fewer result in an outbreak with sustained human-to-human transmission. Most initial transmission events evade detection because they do not cause disease, do not come to medical attention, or are misdiagnosed. The lab leak hypothesis postulates that researchers constructed or simply cultured SARS-CoV-2 during research on bat-origin coronaviruses (CoVs) and that an accidental exposure to this lab-derived virus started the pandemic.

It is worthwhile to consider what type of evidence would definitively rule in or rule out either the zoonotic or lab leak hypothesis. Scientific evidence alone is likely to be insufficient to provide a definitive answer (1). In fact, the types of evidence needed for ruling in and out the two hypotheses are quite different. Establishing the lab leak hypothesis would require evidence that the Wuhan Institute of Virology (WIV) was working on a CoV very closely related to the original Wuhan strain, and such evidence would have to come from laboratory records. Had the WIV been working on such a virus, evidence of a laboratory accident and/or that some of the initial cases had come from individuals at the WIV would strongly support the possibility of a lab leak. While the Chinese government has denied that such work was being done by the WIV, transparency is lacking (2). Unless such evidence is forthcoming, the laboratory leak hypothesis cannot be confirmed. Conclusively establishing the zoonotic hypothesis requires finding evidence of the original animal-to-human infection event(s). This is difficult because these events likely occurred unnoticed, and there may be no record of their eventuality. However, the recovery of a CoV that is closely related to SARS-CoV-2 from bats and/or intermediate hosts would support the hypothesis that the coronavirus disease 2019 (COVID-19) pandemic was initiated by a zoonotic event. As discussed below, however, the needed exploration of multiple species for CoV prevalence and diversity will take time to achieve.

In addition to requiring very different sources of data, we must realize that the levels of certainty required to definitively prove either the lab leak or the zoonotic hypothesis differ significantly. Specifically, it should be possible to rule in or out the lab leak theory with information from the WIV, if such information exists. In contrast, unequivocally ruling in or out the zoonotic explanation is not possible since the evidence will always depend on certain probabilities, and certainty is impossible without knowledge of the initial events.

In this light, the only way to assess the viability of the hypotheses is to examine the best available evidence. It is important that scientists, the public, and public figures follow the facts and limit speculation that can become fodder for misinformation and conspiracy theories. In most instances, this is how science proceeds, following and building upon the best available evidence but always maintaining an open mind since all scientific knowledge is provisional, and the next experiment might change our opinions dramatically. In applying the scientific method, we ask a question, form a hypothesis, make a prediction, test the prediction, and then use the results to make a new hypothesis. The iterative nature of scientific investigation requires constantly revisiting previous conclusions as new data are collected. Scientific findings based on this philosophy are our best tool for directing our resources for effective responses to the many challenges that face our society. The scientific method has been broadly applied by scientists the world over to understand the origin of SARS-CoV-2.

To explore the origin question, we pose several relevant hypotheses below and critically consider the scientific data to support or refute them. Other evidence may emerge from ongoing investigations carried out by the scientific community, intelligence community, and government agencies that will hopefully shed further light on the SARS-CoV-2 origin question. It is important that as more evidence emerges, we revisit and reshape conclusions to reflect new data.

HYPOTHESIS 1: SARS-CoV-2 AROSE FROM A LABORATORY-ADAPTED CoV

Background. The genomic sequence of SARS-CoV-2 offers a means of testing the laboratory-origin hypothesis. Namely, if the virus had a laboratory origin, we would expect the viral genome to carry signatures resulting from the propagation of the biological isolate in the lab. This expectation is based on the reasoning that if the isolate were present in the lab but not propagated, the potential for laboratory-acquired infection would be extremely low due to the limited handling of the sample and likely low viral titers.

Evidence A: loss of the furin cleavage site in the spike protein during cell culture adaptation of initial isolates of SARS-CoV-2. Over the past 3 years, hundreds of research groups around the world have cultured human isolates of SARS-CoV-2 in cell lines. This collective experience has revealed a highly reproducible consequence of viral propagation in culture: the deletion of the portion of the spike gene encoding the furin cleavage site (3). If SARS-CoV-2 had a laboratory origin, it would have been amplified in a laboratory through a process of serial passage typically needed to recover high-titer stocks from environmental samples. In this process, the deletion of the furin cleavage site is expected, offering a signature of laboratory handling. However, early isolates of SARS-CoV-2 show the furin cleavage site to be intact, arguing against introduction into humans after laboratory cell culture.

Evidence B: initial SARS-CoV-2 isolates replicate poorly in traditional laboratory models. Traditional laboratory experiments would likely have included the adaptation of a virus in common laboratory animal models such as mice to study it with ease. Thus, a laboratory-derived virus released into the population would reasonably be expected to carry these adaptive markers in their genomes. However, early isolates of SARS-CoV-2 did not carry mutations that confer adaptation to common animal models, such as mutations in the receptor binding domain of spike that improve viral replication in mice (4, 5).

HYPOTHESIS 2: SARS-CoV-2 IS A LABORATORY-CONSTRUCTED VIRUS

Background. CoV can be modified or generated from sequence information using molecular biology techniques (6). The existence of this technology has led some to suggest that SARS-CoV-2 not only may have been propagated in a lab but also may have been generated *de novo* in a lab. In this context, the furin cleavage site has been highlighted as a feature that may have been designed to increase virulence.

Evidence A: lack of any evidence of deliberate genetic engineering. Genetic engineering involves the use of recombinant DNA techniques to modify the virus. Such techniques often leave telltale evidence in the forms of novel restriction sites, differences in DNA base content, differences in gene organization from natural isolates, selection markers, and foreign nonviral sequences. While various papers have suggested that SARS-CoV-2 carries hallmarks of recombinant DNA technologies, this theory has been thoroughly refuted (7), including by the U.S. intelligence community.

Evidence B: the pathogenesis of SARS-CoV-2 may be dependent on the cleavage site loop length and not only the presence of the furin cleavage site. The furin cleavage site is an important determinant of virulence (4). A furin cleavage site is absent from the closest known relatives of SARS-CoV-2; however, these closest relatives are, in fact, distantly related. Furin cleavage sites are commonplace in other CoV spike proteins, including those of endemic human betacoronaviruses, and were first identified in the 1980s within a mouse coronavirus (8). Prior to the COVID-19 pandemic, it was commonly thought that the presence of the furin cleavage site was directly linked to the pathogenesis of CoV. However, a recent examination of SARS-CoV-2 has revealed a somewhat complex relationship between the presence of a furin cleavage site and viral pathogenicity. In fact, the context of the cleavage site defines the extent to which this site drives virulence. Specifically, the amino acids N terminal to the cleavage site, which comprise the cleavage loop, are important: both the loop's length and the presence of a glycosylation site within it define the contribution of the furin cleavage site to virulence (9). Since these nuances were not understood prior to 2020, deliberate engineering of the SARS-CoV-2 cleavage site to promote pathogenesis without this critical information is improbable.

Evidence C: loss of the furin cleavage site through cell culture propagation of the virus. As mentioned above in the evidence for hypothesis 1, the propagation of the virus in standard cell culture results in a deletion of the furin cleavage site. Constructing a virus in a laboratory would necessarily require that it be grown in a lab, where the site would likely be deleted, yet the subsequently lab-grown virus population must retain the furin cleavage site. This contradictory situation, as well as the fact that SARS-CoV-2 isolates retain a furin cleavage site, suggests that the virus had not been cultured in a lab at the time of emergence.

HYPOTHESIS 3: SARS-CoV-2 IS A BAT ZONOSIS INTRODUCED INTO THE HUMAN POPULATION

Background. It is widely accepted that bat species contain a large diversity of coronaviruses. Since the SARS-CoV outbreak in 2003, we furthermore know that bats act as a reservoir for CoVs that pose the potential for spillover. For this reason, bats have been subject to large-scale surveillance since 2003. In addition to CoVs, many human viruses have their origins in nonhuman hosts. Influenza A viruses are a pertinent example owing to the recurrent nature of influenza pandemics. Our most recent influenza pandemic, in 2009, was caused by a virus circulating in pigs. While pigs were known to harbor influenza viruses with zoonotic potential, and the 2009 strain showed similarities to swine strains, a genetically similar precursor virus was not identified in swine until 2016 (10).

Evidence A: a bat origin is highly plausible given the precedent of SARS-CoV and the diversity of CoVs in global bat populations. The source of the SARS-CoV outbreak of 2002 to 2004 is particularly relevant to consider given the similarities between this virus and SARS-CoV-2. In the years following the SARS-CoV outbreak, extensive molecular epidemiology and wildlife surveillance efforts determined its proximal source to be civet cats sold within live-animal markets in Guangdong, China (11). The civet cats, in turn, were likely to have been infected by bats (12). Although an intermediate animal host for SARS-CoV-2 has not been identified, a similar path of emergence from bats to wild animals to humans, facilitated by live-animal markets, is feasible. Viruses related to SARS-CoV-2 have been documented in multiple locations in Southeast Asia, primarily in bats (13–18). These data show that bats harbor a diverse gene pool of CoVs and are likely to have sustained circulation of the precursor lineage.

Evidence B: bat CoVs identified to date are not closely related to SARS-CoV-2. Phylogenetic analyses reveal a significant evolutionary gap between SARS-CoV-2 and the most closely related viruses sampled to date from bats. Bat CoVs were under study at the Wuhan Institute of Virology (WIV) in the period preceding the pandemic as part of an ongoing important surveillance effort. The most closely related virus known to be in the WIV collection, RaTG13, differs from early SARS-CoV-2 isolates by >1,000 nucleotides, a genetic distance of ~4% (19). Given this dissimilarity, RaTG13 is clearly not a proximal ancestor of SARS-CoV-2. However, owing to a lack of transparency, independent verification of the viruses present in the laboratories of the WIV is lacking. Further surveillance in bats since 2020 has identified additional CoVs that are related to SARS-CoV-2 (13–18, 20). Because surveillance coverage is far from complete, however, large gaps remain in our knowledge that preclude the identification of the viral lineage that seeded the pandemic or the intermediate host: the host niche harboring the direct precursors of SARS-CoV-2 has not yet been adequately sampled. It is important to recognize that interpreting this failure to identify a feral source of SARS-CoV-2 as evidence supporting a laboratory origin constitutes the logical error of using the absence of evidence as evidence for an absence. We note that, as for the 2009 influenza pandemic, it took many years to identify the feral sources of SARS-CoV and HIV after their initial isolation from humans (19, 21).

Evidence C: high rates of recombination of CoVs occur in nature. Recombination within the *Sarbecovirus* subgenus to which SARS-CoV-2 belongs is exceedingly common (18, 22, 23). Recombination naturally produces chimeric genomes, making lineage tracing to ancestors more complex, but also increases the natural diversity of CoVs. Thus, signatures of recombination in the SARS-CoV-2 genome do not point specifically to either a zoonotic or a laboratory origin.

Evidence D: numerous animal reservoirs for CoVs and SARS-CoV-2 are known to exist in nature. During the COVID-19 pandemic, many animal hosts at the human-animal interface were identified to be susceptible to the virus, such as cats and deer (24, 25). However, an analysis of animals at the Wuhan market at the time of the outbreak did not identify a single species as the intermediate host (26). SARS-CoV-2 sequences were detected in environmental samples obtained from cages and stalls where live animals susceptible to SARS-CoV infection had been kept, but those animals were no longer present (26). The broad tropism of CoVs in nature has led to speculation of pangolins and snakes as the intermediate hosts. The lack of surveillance data prior to the outbreak limits the ability to identify a possible intermediate host.

Evidence E: serological evidence for human exposure to bat viruses. Antibodies to bat coronavirus have been detected in humans living near caves in China, providing immunological evidence for the transmission of CoVs from bats to humans directly or through intermediate hosts (27). Recently, a novel rhabdovirus was isolated from bats, and virus-specific antibodies were detected in people who lived near the collection site, providing more evidence that bats can be a source of viruses that infect humans (27).

HYPOTHESIS 4: THE SARS-CoV-2 ORIGIN IS SUGGESTED BY EARLY CASES IN THE COVID-19 OUTBREAK

Background. The COVID-19 pandemic began in central Wuhan, a densely populated megacity of over 11 million people. At least four markets in Wuhan traded live wild animals in the months preceding the pandemic (2), an activity that brings large numbers of people into proximity with a diversity of wild animals.

Evidence A: known cases cluster at or near the Huanan Seafood Wholesale Market. Among the first 174 documented cases of SARS-CoV-2 infection in December 2019, approximately one-half had a recent association with a market, and approximately one-quarter had an association with the Huanan Seafood Wholesale Market in particular (2). None were scientists working on bat CoVs. Importantly, this set of early cases is almost certainly incomplete since many SARS-CoV-2 infections are asymptomatic, and mild cases are difficult to distinguish clinically from other respiratory infections. However, both in terms of a direct association and in considering geographical proximity, the early stages of the outbreak showed a clear link to the Huanan market (26).

Evidence B: emergence of two lineages of SARS-CoV-2. Early cases in Wuhan involved two lineages, distinguishable by the viral genome sequence, suggesting that there were multiple events of spillover into humans. This situation is more likely to have resulted from zoonosis than a lab accident as lab accidents are relatively rare, and both lineages of SARS-CoV-2 were found in sequences from the Hunan market (28). Additionally, the extreme transmission heterogeneity of SARS-CoV-2 is important to consider. Epidemiological analysis of the pandemic indicates that 80% of infections do not lead to onward transmission (29). This heterogeneity suggests that many spillover events, rather than a single introduction through an infected lab worker, were likely needed to initiate sustained spread.

Evidence C: lack of epidemiological information from the WIV. Because data are not available, it is unclear how comprehensively Wuhan-based scientists working on CoVs were tested. While it has been reported that three scientists from the WIV had a respiratory illness in the fall of 2019 (30), they were not tested for any specific pathogen, and therefore, the causative agent of their illness cannot be determined. Human respiratory illnesses are common in the general population and can be caused by several viruses such as influenza virus, parainfluenza virus, and respiratory syncytial virus (RSV) as well as some bacteria and fungi. For many respiratory illnesses, the infectious etiology is never known.

CONCLUSIONS DRAWN FROM EXISTING EVIDENCE

Scientific conclusions are based on likelihood given the scientific data, and conclusions can change as new data are obtained. Based on the scientific data collected in the last 3 years by virologists worldwide, hypotheses 1 and 2 are unlikely. Hypotheses 3 and 4 cannot be ruled out by existing evidence. Since hypotheses 1 and 2 support the lab leak

theory and hypotheses 3 and 4 are consistent with a zoonotic origin, the lab leak- and zoonotic-origin explanations are not equally probable, and the available evidence favors the latter. Further insight into CoVs in animals at the animal-human interface requires additional surveillance of circulating virus sequences from animals. There is ample precedent for the seeding of pandemics and more geographically limited outbreaks from nonhuman species. Common-cold CoVs, SARS-CoV, Ebola virus, HIV, influenza A virus, mpox virus, and others all have zoonotic origins (31–33). SARS-CoV-2 is the ninth documented coronavirus to enter the human population. The best existing scientific evidence supports a direct zoonotic origin. As new evidence continues to emerge from scientific studies or other investigations, our understanding of the origin of SARS-CoV-2 will continue to evolve. Nevertheless, it is possible that its origin may never be known with certainty.

CONCLUDING REMARKS

The COVID-19 pandemic brought tremendous suffering to humanity. It is reasonable to want to hold an individual(s) accountable, and it may be unsatisfying to attribute such calamities to natural causes. We understand the anger that people feel because of losing or being separated from loved ones, homeschooling their children, and having their lives upended. We have lived that experience as well. Science was our best ally during the pandemic, working to understand virus replication, spread, and disease to produce life-saving vaccines, antivirals, rapid tests, and treatments. Science is also a key ally in addressing origin hypotheses. Given the currently available evidence, the two plausible possibilities for the origin of SARS-CoV-2 are not equally likely. Continuing to frame them as such does us all a disservice, dilutes our efforts, and misdirects resources—only serving to weaken our ability to respond to future pandemics.

REFERENCES

- Casadevall A, Weiss SR, Imperiale MJ. 2021. Can science help resolve the controversy on the origins of the SARS-CoV-2 pandemic? *mBio* 12:e01948-21. <https://doi.org/10.1128/mBio.01948-21>.
- WHO. 2021. WHO-convended global study of origins of SARS-CoV-2: China part. WHO, Geneva, Switzerland. <https://www.who.int/publications/i/item/who-convended-global-study-of-origins-of-sars-cov-2-china-part>. Accessed 5 March 2023.
- Hale BG. 2021. Avoiding culture shock with the SARS-CoV-2 spike protein. *Elife* 10:e69496. <https://doi.org/10.7554/eLife.69496>.
- Leist SR, Dinnon KH, III, Schafer A, Tse LV, Okuda K, Hou YJ, West A, Edwards CE, Sanders W, Fritch EJ, Gully KL, Scobey T, Brown AJ, Sheahan TP, Moorman NJ, Boucher RC, Gralinski LE, Montgomery SA, Baric RS. 2020. A mouse-adapted SARS-CoV-2 induces acute lung injury and mortality in standard laboratory mice. *Cell* 183:1070–1085.e12. <https://doi.org/10.1016/j.cell.2020.09.050>.
- Zhang W, Shi K, Geng Q, Ye G, Aihara H, Li F. 2022. Structural basis for mouse receptor recognition by SARS-CoV-2 omicron variant. *Proc Natl Acad Sci U S A* 119:e2206509119. <https://doi.org/10.1073/pnas.2206509119>.
- Yount B, Denison MR, Weiss SR, Baric RS. 2002. Systematic assembly of a full-length infectious cDNA of mouse hepatitis virus strain A59. *J Virol* 76:11065–11078. <https://doi.org/10.1128/jvi.76.21.11065-11078.2002>.
- Johns Hopkins Center for Health Security. 2020. In response: Yan et al. Preprint examinations of the origin of SARS-CoV-2. Johns Hopkins Center for Health Security, Baltimore, MD. https://www.centerforhealthsecurity.org/our-work/pubs_archive/pubs-pdfs/2020/200921-in-response-yan.pdf. Accessed 5 March 2023.
- Sturman LS, Ricard CS, Holmes KV. 1985. Proteolytic cleavage of the E2 glycoprotein of murine coronavirus: activation of cell-fusing activity of virions by trypsin and separation of two different 90K cleavage fragments. *J Virol* 56:904–911. <https://doi.org/10.1128/JVI.56.3.904-911.1985>.
- Vu MN, Lokugamage KG, Plante JA, Scharton D, Bailey AO, Sotcheff S, Swetnam DM, Johnson BA, Schindewolf C, Alvarado RE, Crocquet-Valdes PA, Debbink K, Weaver SC, Walker DH, Russell WK, Routh AL, Plante KS, Menachery VD. 2022. QTQTN motif upstream of the furin-cleavage site plays a key role in SARS-CoV-2 infection and pathogenesis. *Proc Natl Acad Sci U S A* 119:e2205690119. <https://doi.org/10.1073/pnas.2205690119>.
- Mena I, Nelson MI, Quezada-Monroy F, Dutta J, Cortes-Fernandez R, Lara-Puente JH, Castro-Peralta F, Cunha LF, Trovao NS, Lozano-Dubernard B, Rambaut A, van Bakel H, Garcia-Sastre A. 2016. Origins of the 2009 H1N1 influenza pandemic in swine in Mexico. *Elife* 5:e16777. <https://doi.org/10.7554/eLife.16777>.
- Song H-D, Tu C-C, Zhang G-W, Wang S-Y, Zheng K, Lei L-C, Chen Q-X, Gao Y-W, Zhou H-Q, Xiang H, Zheng H-J, Chern S-WW, Cheng F, Pan C-M, Xuan H, Chen S-J, Luo H-M, Zhou D-H, Liu Y-F, He J-F, Qin P-Z, Li L-H, Ren Y-Q, Liang W-J, Yu Y-D, Anderson L, Wang M, Xu R-H, Wu X-W, Zheng H-Y, Chen J-D, Liang G, Gao Y, Liao M, Fang L, Jiang L-Y, Li H, Chen F, Di B, He L-J, Lin J-Y, Tong S, Kong X, Du L, Hao P, Tang H, Bernini A, Yu X-J, Spiga O, Guo Z-M, et al. 2005. Cross-host evolution of severe acute respiratory syndrome coronavirus in palm civet and human. *Proc Natl Acad Sci U S A* 102:2430–2435. <https://doi.org/10.1073/pnas.0409608102>.
- Li W, Shi Z, Yu M, Ren W, Smith C, Epstein JH, Wang H, Cramer G, Hu Z, Zhang H, Zhang J, McEachern J, Field H, Daszak P, Eaton BT, Zhang S, Wang LF. 2005. Bats are natural reservoirs of SARS-like coronaviruses. *Science* 310:676–679. <https://doi.org/10.1126/science.1118391>.
- Zhou H, Chen X, Hu T, Li J, Song H, Liu Y, Wang P, Liu D, Yang J, Holmes EC, Hughes AC, Bi Y, Shi W. 2020. 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* 30:3896. <https://doi.org/10.1016/j.cub.2020.09.030>.
- Delaune D, Hul V, Karlsson EA, Hassanin A, Ou TP, Baidaliuk A, Gambaro F, Prot M, Tu VT, Chea S, Keatts L, Mazet J, Johnson CK, Buchy P, Dussart P, Goldstein T, Simon-Lorriere E, Duong V. 2021. A novel SARS-CoV-2 related coronavirus in bats from Cambodia. *Nat Commun* 12:6563. <https://doi.org/10.1038/s41467-021-26809-4>.
- Li L-L, Wang J-L, Ma X-H, Sun X-M, Li J-S, Yang X-F, Shi W-F, Duan Z-J. 2021. A novel SARS-CoV-2 related coronavirus with complex recombination isolated from bats in Yunnan Province, China. *Emerg Microbes Infect* 10:1683–1690. <https://doi.org/10.1080/22221751.2021.1964925>.
- Wacharapluesadee S, Tan CW, Maneerom P, Duengkae P, Zhu F, Joyjinda Y, Kaewpom T, Chia WN, Ampoot W, Lim BL, Worachotsueptrakun K, Chen VC-W, Sirichan N, Ruchisrisarod C, Rodpan A, Noradechanon K, Phaichana T, Jantarat N, Thongnumchaima B, Tu C, Cramer G, Stokes MM, Hemachudha T, Wang L-F. 2021. Evidence for SARS-CoV-2 related coronaviruses circulating in bats and pangolins in Southeast Asia. *Nat Commun* 12:972. <https://doi.org/10.1038/s41467-021-21768-2>.
- Nga NTT, Latinne A, Thuy HB, Long NV, Ngoc PTB, Anh NTL, Thai NV, Phuong TQ, Thai HV, Hai LK, Long PT, Phuong NT, Hung VV, Quang LTV, Lan NT, Hoa NT, Johnson CK, Mazet JAK, Robertson SI, Walzer C, Olson SH, Fine AE. 2022.

- Evidence of SARS-CoV-2 related coronaviruses circulating in Sunda pangolins (*Manis javanica*) confiscated from the illegal wildlife trade in Viet Nam. *Front Public Health* 10:826116. <https://doi.org/10.3389/fpubh.2022.826116>.
18. Zhou H, Ji J, Chen X, Bi Y, Li J, Wang Q, Hu T, Song H, Zhao R, Chen Y, Cui M, Zhang Y, Hughes AC, Holmes EC, Shi W. 2021. Identification of novel bat coronaviruses sheds light on the evolutionary origins of SARS-CoV-2 and related viruses. *Cell* 184:4380–4391.e14. <https://doi.org/10.1016/j.cell.2021.06.008>.
 19. Zhou P, Yang X-L, Wang X-G, Hu B, Zhang L, Zhang W, Si H-R, Zhu Y, Li B, Huang C-L, Chen H-D, Chen J, Luo Y, Guo H, Jiang R-D, Liu M-Q, Chen Y, Shen X-R, Wang X, Zheng X-S, Zhao K, Chen Q-J, Deng F, Liu L-L, Yan B, Zhan F-X, Wang Y-Y, Xiao G-F, Shi Z-L. 2020. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* 579:270–273. <https://doi.org/10.1038/s41586-020-2012-7>.
 20. Wahba L, Jain N, Fire AZ, Shoura MJ, Artilles KL, McCoy MJ, Jeong D-E. 2020. An extensive meta-metagenomic search identifies SARS-CoV-2-homologous sequences in pangolin lung viromes. *mSphere* 5:e00160-20. <https://doi.org/10.1128/mSphere.00160-20>.
 21. Sharp PM, Hahn BH. 2011. Origins of HIV and the AIDS pandemic. *Cold Spring Harb Perspect Med* 1:a006841. <https://doi.org/10.1101/cshperspect.a006841>.
 22. Lytras S, Hughes J, Martin D, Swanepoel P, de Klerk A, Lourens R, Kosakovsky Pond SL, Xia W, Jiang X, Robertson DL. 2022. Exploring the natural origins of SARS-CoV-2 in the light of recombination. *Genome Biol Evol* 14:evac018. <https://doi.org/10.1093/gbe/evac018>.
 23. Boni MF, Lemey P, Jiang X, Lam TT-Y, Perry BW, Castoe TA, Rambaut A, Robertson DL. 2020. Evolutionary origins of the SARS-CoV-2 sarbecovirus lineage responsible for the COVID-19 pandemic. *Nat Microbiol* 5:1408–1417. <https://doi.org/10.1038/s41564-020-0771-4>.
 24. EFSA Panel on Animal Health and Welfare (AHAW), Nielsen SS, Alvarez J, Bicoût DJ, Calistri P, Canali E, Drewe JA, Garin-Bastuji B, Gonzales Rojas JL, Gortazar C, Herskin M, Michel V, Miranda Chueca MA, Padalino B, Pasquali P, Roberts HC, Spoolder H, Velarde A, Viltrop A, Winckler C, Adlhoch C, Aznar I, Baldinelli F, Boklund A, Broglia A, Gerhards N, Mur L, Nannapaneni P, Stahl K. 2023. SARS-CoV-2 in animals: susceptibility of animal species, risk for animal and public health, monitoring, prevention and control. *EFSA J* 21:e07822. <https://doi.org/10.2903/j.efsa.2023.7822>.
 25. Marques AD, Sherrill-Mix S, Everett JK, Adhikari H, Reddy S, Ellis JC, Zelif H, Greening SS, Cannuscio CC, Strelau KM, Collman RG, Kelly BJ, Rodino KG, Bushman FD, Gagne RB, Anis E. 2022. Multiple introductions of SARS-CoV-2 Alpha and Delta variants into white-tailed deer in Pennsylvania. *mBio* 13:e02101-22. <https://doi.org/10.1128/mbio.02101-22>.
 26. Worobey M, Levy JI, Malpica Serrano L, Crits-Christoph A, Pekar JE, Goldstein SA, Rasmussen AL, Kraemer MUG, Newman C, Koopmans MPG, Suchard MA, Wertheim JO, Lemey P, Robertson DL, Garry RF, Holmes EC, Rambaut A, Andersen KG. 2022. The Huanan Seafood Wholesale Market in Wuhan was the early epicenter of the COVID-19 pandemic. *Science* 377:951–959. <https://doi.org/10.1126/science.abp8715>.
 27. Wang N, Li S-Y, Yang X-L, Huang H-M, Zhang Y-J, Guo H, Luo C-M, Miller M, Zhu G, Chmura AA, Hagan E, Zhou J-H, Zhang Y-Z, Wang L-F, Daszak P, Shi Z-L. 2018. Serological evidence of bat SARS-related coronavirus infection in humans, China. *Viral Sin* 33:104–107. <https://doi.org/10.1007/s12250-018-0012-7>.
 28. Pekar JE, Magee A, Parker E, Moshiri N, Izhikevich K, Havens JL, Gangavarapu K, Malpica Serrano LM, Crits-Christoph A, Matteson NL, Zeller M, Levy JI, Wang JC, Hughes S, Lee J, Park H, Park M-S, Ching Zi Yan K, Lin RTP, Mat Isa MN, Noor YM, Vasylyeva TI, Garry RF, Holmes EC, Rambaut A, Suchard MA, Andersen KG, Worobey M, Wertheim JO. 2022. The molecular epidemiology of multiple zoonotic origins of SARS-CoV-2. *Science* 377:960–966. <https://doi.org/10.1126/science.abp8337>.
 29. Adam DC, Wu P, Wong JY, Lau EHY, Tsang TK, Cauchemez S, Leung GM, Cowling BJ. 2020. Clustering and superspreading potential of SARS-CoV-2 infections in Hong Kong. *Nat Med* 26:1714–1719. <https://doi.org/10.1038/s41591-020-1092-0>.
 30. Bertrand N, Atwood K, Williams KB, Cohen Z. 2021. New information on Wuhan researchers' illness further debate on pandemic origins, *on CNN*. <https://www.cnn.com/2021/05/23/politics/us-intelligence-report-wuhan-lab-researchers-coronavirus-origin/index.html>. Accessed 24 May 2021.
 31. Plowright RK, Parrish CR, McCallum H, Hudson PJ, Ko AI, Graham AL, Lloyd-Smith JO. 2017. Pathways to zoonotic spillover. *Nat Rev Microbiol* 15:502–510. <https://doi.org/10.1038/nrmicro.2017.45>.
 32. Albery GF, Becker DJ, Brierley L, Brook CE, Christofferson RC, Cohen LE, Dallas TA, Eskew EA, Fagre A, Farrell MJ, Glennon E, Guth S, Joseph MB, Mollentze N, Neely BA, Poisot T, Rasmussen AL, Ryan SJ, Seifert S, Sjodin AR, Sorrell EM, Carlson CJ. 2021. The science of the host-virus network. *Nat Microbiol* 6:1483–1492. <https://doi.org/10.1038/s41564-021-00999-5>.
 33. Taubenberger JK, Kash JC. 2010. Influenza virus evolution, host adaptation, and pandemic formation. *Cell Host Microbe* 7:440–451. <https://doi.org/10.1016/j.chom.2010.05.009>.