

## Back to Science in Searching for SARS-CoV-2 Origins

William J Liu<sup>1,\*</sup>; Wenwen Lei<sup>1</sup>; Xiaozhou He<sup>1</sup>; Peipei Liu<sup>1</sup>; Qihui Wang<sup>2</sup>; Zhiqiang Wu<sup>3</sup>; Yun Tan<sup>4</sup>; Shuhui Song<sup>5</sup>; Gary Wong<sup>6</sup>; Jian Lu<sup>7</sup>; Jingkun Jiang<sup>8</sup>; Qiang Wei<sup>9</sup>; Mingkun Li<sup>5</sup>; Juncai Ma<sup>10</sup>; Xiaozhong Peng<sup>11</sup>; Yixue Li<sup>12</sup>; Baoxu Huang<sup>13</sup>; Yigang Tong<sup>14</sup>; Jun Han<sup>1</sup>; Guizhen Wu<sup>1,\*</sup>

In recent decades, emerging and re-emerging human-infecting pathogens have been represented as huge threats to public health and have become a global concern (1). After outbreaks of two coronaviruses (CoVs), severe acute respiratory syndrome coronavirus (SARS-CoV) and Middle East respiratory syndrome coronavirus (MERS-CoV), severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) became the first-known pandemic hastening CoV with tremendous wrecking to the world (2). The origin tracing of these emerging pathogens is of great significance in infectious disease prevention and control (3–4). The origin of SARS-CoV-2 remains elusive after the more than 3-year pandemic, though scientists around the world are making great efforts. From the experience of studying many other infectious pathogens, origin tracing is systematic and time-consuming work. The supposed origins of many infectious pathogens are still in debate, including SARS-CoV and human immunodeficiency virus, etc (5).

The establishment of a defined origin of an emerging human-infecting pathogen requires rigorous logical relationships, sufficient causal relationships, and a solid chain of evidence. Generally speaking, to understand the concept of “origin,” we believe that it can be divided into two perspectives: 1) From a macro perspective, for the global origin of coronavirus disease 2019 (COVID-19), zoonotic origin is still the most possible way. A natural host would exist, which may or may not require intermediate animals. However, based on the current global evidence available, there is no definitive conclusion for both natural and intermediate animals for SARS-CoV-2 yet. 2) In a narrower sense, for the origin (or more accurately, the “introduction” or the route of infection of the index case) in a particular geographic range (such as a country, city, school or market) or in a particular population, we believe that this origin may have multiple possibilities. For example, the initial case in a geographical area was infected by a person entering this area, by contact with the items entering the area as a vector of the pathogen,

or by an animal carrying the pathogen entering the area. However, it should be noted that even if the origin within a certain area (even as the early reported outbreak site) is found, it cannot be equated with the origin of the global pandemic, that is, “the location of discovery” does not equal to “the location of origin”.

To draw the scientific conclusion that a certain item is the origin of an emerging pathogen, scientific logic and evidence for research are also needed. First of all, to prove that the infection origin of item B, i.e., index case is item A (A and B refer to individuals rather than groups), which can be human, animal, or environmental medium items, the following representatives in the evidence chain are required: 1) Prior to the onset of disease, B has a history of exposure to A and the infection of A is prior to B (epidemiological evidence); 2) the pathogen is detected and/or isolated from A, in contact by B (laboratory evidence); and 3) the genome of the pathogen sequenced from the samples of A and B showed high homology or progeny relationship (molecular evidence). Considering the complexity and rigor of causality, these are necessary but not sufficient conditions to prove that A, as the origin, transmitted the pathogen to B.

On the issue of SARS-CoV-2 origin, all the studies and the related conclusions should be based on science. Chinese scientists have always maintained an open, transparent, and responsible attitude and practice and worked together with scientists from all over the world on origin tracing and contributed a lot of scientific references, with WHO-China Joint Report as one of the achievements (6).

During the early stage of COVID-19 outbreak in late December 2019, a certain percentage of cases were found to be linked to the Huanan Seafood Market (HSM) in Wuhan, China. Thus, the surveillance of SARS-CoV-2 in the environment and animal samples of the HSM is of great importance for investigating how the viruses were introduced into the market. The market was closed on the morning of January 1, 2020,

and at the same time, the Chinese Center for Disease Control and Prevention (China CDC) and local CDCs dispatched experts to the HSM to collect environmental samples. From January to March 2020, environmental samples from different locations within and around this market and animal samples including animal bodies, stray animals, and their feces were collected. The surveillance results reflect the profile of SARS-CoV-2 contamination by early cases in the market during the early phase of the outbreak. These results were timely and continuously informed to the public through official media (<http://www.gov.cn/xinwen/gwylflkjz16/wzsl.htm>). After the surveillance work in the field, the data were further analyzed and during the WHO-convened Global Study of Origins of SARS-CoV-2: China Part from July 2020 to February 2021, the results were reported to WHO experts and comprehensively discussed by a joint expert team of WHO and China. The joint team also entered the site of HSM in January 2021 and they recognized the importance of the surveillance methods, data, and results, which were written into the Joint Report (6). Furthermore, the Joint Report also suggested a list of more investigations and analyses to be performed in the future, including DNA barcoding analysis of the samples within the market. Meanwhile, the data on the HSM together with the barcoding analyses were summarized into a scientific paper submitted to a peer-review journal and also released as a preprint for the public (7–8) and further in-depth analyses were also performed during this process. Meanwhile, the raw data related to the study were deposited to GISAID (9). The data were not removed or deleted since the deposit. In accordance with the usual practice of publication, and by agreement with the journal and GISAID, the data would be released simultaneously with the formal publication, but the link to access the data for journal review has always existed. At present, all the data have been released simultaneously in four international databases, i.e. 1) GISAID; 2) Sequence Read Archive (SRA), National Center for Biotechnology Information (NCBI); 3) National Genomics Data Center (NGDC), China National Center for Bioinformation (CNCB); 4) China National Microbiology Data Center (NMDC).

The study included the SARS-CoV-2 detection results of 1,380 samples collected from the environment and the animals within the market in early 2020. By SARS-CoV-2-specific RT-qPCR, 73 environmental samples tested positive for SARS-CoV-2 and three live viruses were successfully isolated. The

viruses from the market shared nucleotide identity of 99.980% to 100% with the SARS-CoV-2 isolates from early COVID-19 cases. No virus was detected in the animal swabs covering 18 species of animals in the market. The RNA-seq analysis of SARS-CoV-2 positive and negative environmental samples showed the abundance of different vertebrate species. Through the DNA barcoding analysis, the high abundance of *Homo sapiens* within the environment samples highly suggests that the SARS-CoV-2 in the environmental samples was derived from early cases in HSM. Also, the results indicated the existence of *Sus*, *Bos*, *Gallus*, *Anas*, and *Nyctereutes* and other animals through gene barcode within the market before the closure. However, these environmental samples could not prove the infection of the animals. Furthermore, even if the animals were infected, the possibility of human-to-animal transmission occurring could not be ruled out after human infection. Thus, the possible potential introduction of the virus through human or cold chain products into the market can still not be ruled out. All these data showed that the market has just acted as an early amplifier during the pandemic due to the high number of consumers every day, causing many initially identified infection clusters. The origin of the virus involving animal-to-human transmission cannot be determined based on the current data.

At the same time as this study, Chinese scientists have also published a series of papers from the perspectives of serological testing of potential SARS-CoV-2 infection using blood from blood donors in Wuhan in 2019 (10), animal surveillance (11–13), the cold chain (14), etc (3–4,15). These studies have provided more scientific references for origin tracing and also provided scientific perspectives for future research.

Based on the current scientific logic and the evidence available globally, the several hypotheses and the conclusions on the origins of SARS-CoV-2 in the previous WHO-China Joint Report are scientific and objective (6). That is, traceability is a scientific issue that requires scientific evidence and logical inference. The origin tracing of SARS-CoV-2 may still be a long way off. However, it is believed that the experience accumulated in the studies on COVID-19 will shed light on the prevention and control of other emerging and re-emerging infectious diseases in the future (16).

doi: 10.46234/ccdcw2023.055

\* Corresponding authors: William J. Liu, [liujun@ivdc.chinacdc.cn](mailto:liujun@ivdc.chinacdc.cn); Guizhen Wu, [wuzg@ivdc.chinacdc.cn](mailto:wuzg@ivdc.chinacdc.cn).

<sup>1</sup> NHC Key Laboratory of Biosafety, National Institute for Viral Disease Control and Prevention, Chinese Center for Disease Control and Prevention, Beijing, China; <sup>2</sup> CAS Key Laboratory of Pathogen Microbiology and Immunology, Institute of Microbiology, Chinese Academy of Sciences, Beijing, China; <sup>3</sup> NHC Key Laboratory of Systems Biology of Pathogens, Institute of Pathogen Biology, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, China; <sup>4</sup> Shanghai Institute of Hematology, State Key Laboratory of Medical Genomics, National Research Center for Translational Medicine, Ruijin Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, Shanghai, China; <sup>5</sup> Beijing Institute of Genomics, Chinese Academy of Sciences, and China National Center for Bioinformation, Beijing, China; <sup>6</sup> Laboratory of Molecular Virology and Immunology, Institut Pasteur of Shanghai, Chinese Academy of Sciences, Shanghai, China; <sup>7</sup> State Key Laboratory of Protein and Plant Gene Research, Center for Bioinformatics, School of Life Sciences, Peking University, Beijing, China; <sup>8</sup> School of Environment, Tsinghua University, Beijing, China; <sup>9</sup> National Pathogen Resource Center, Chinese Center for Disease Control and Prevention (China CDC), Beijing, China; <sup>10</sup> Microbial Resource and Big Data Center, Chinese National Microbiology Data Center (NMDC), Institute of Microbiology, Chinese Academy of Sciences, Beijing, China; <sup>11</sup> State Key Laboratory of Medical Molecular Biology, Department of Molecular Biology and Biochemistry, Institute of Basic Medical Sciences, Medical Primate Research Center, Neuroscience Center, Chinese Academy of Medical Sciences, School of Basic Medicine, Peking Union Medical College, Beijing, China; <sup>12</sup> School of Life Science and Technology, ShanghaiTech University, Shanghai, China; Bio-Med Big Data Center, Key Laboratory of Computational Biology, CAS-MPG Partner Institute for Computational Biology, Shanghai Institute of Nutrition and Health, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Shanghai, China; <sup>13</sup> China Animal Health and Epidemiology Center, Qingdao City, Shandong Province, China; <sup>14</sup> College of Life Science and Technology, Beijing University of Chemical Technology, Beijing, China.

Submitted: March 29, 2023; Accepted: March 30, 2023

## REFERENCES

- Liu WJ, Liu D. The triphibious warfare against viruses. *Sci China Life Sci* 2017;60(12):1295 – 8. <http://dx.doi.org/10.1007/s11427-017-9252-y>.
- Tan WJ, Zhao X, Ma XJ, Wang WL, Niu PH, Xu WB, et al. A novel coronavirus genome identified in a cluster of pneumonia cases — Wuhan, China 2019–2020. *China CDC Wkly* 2020;2(4):61 – 2. <http://dx.doi.org/10.46234/ccdcw2020.017>.
- Tong YG, Liu WL, Liu PP, Liu WJ, Wang QH, Gao GF. The origins of viruses: discovery takes time, international resources, and cooperation. *Lancet* 2021;398(10309):1401 – 2. [http://dx.doi.org/10.1016/S0140-6736\(21\)02180-2](http://dx.doi.org/10.1016/S0140-6736(21)02180-2).
- Wang QH, Chen H, Shi Y, Hughes AC, Liu WJ, Jiang JK, et al. Tracing the origins of SARS-CoV-2: lessons learned from the past. *Cell Res* 2021;31(11):1139 – 41. <http://dx.doi.org/10.1038/s41422-021-00575-w>.
- Liu WL, Liu PP, Liu WJ, Wang QH, Tong YG, Gao GF. Origins of HIV, HCoV-HKU1, SFTSV, and MERS-CoV and beyond. *China CDC Wkly* 2022;4(37):823 – 7. <http://dx.doi.org/10.46234/ccdcw2022.171>.
- WHO-convened global study of origins of SARS-CoV-2: China Part. 2021. <https://www.who.int/publications/i/item/who-convened-global-study-of-origins-of-sars-cov-2-china-part>. [2023-03-29].
- Gao GF, Liu WJ, Liu PP, Lei WW, Jia ZY, He XZ, et al. Surveillance of SARS-CoV-2 in the environment and animal samples of the Huanan Seafood Market. *Res Square* 2022. 022. <http://dx.doi.org/10.21203/rs.3.rs-1370392/v1>.
- Liu WJ, Liu PP, Lei WW, Jia ZY, He XZ, Shi WF, et al. Surveillance of SARS-CoV-2 at the Huanan Seafood Market. *CSTR:32003.36.ChinaXiv.202303.10351.V1*. 2023. <http://dx.doi.org/10.12074/202303.10351V1>.
- Khare S, Gurry C, Freitas L, Schultz MB, Bach G, Diallo A, et al. GISAID's role in pandemic response. *China CDC Wkly* 2023;3(49):1049 – 51. <http://dx.doi.org/10.46234/ccdcw2021.255>.
- Chang L, Zhao L, Xiao Y, Xu TT, Chen L, Cai Y, et al. Serosurvey for SARS-CoV-2 among blood donors in Wuhan, China from September to December 2019. *Protein Cell* 2022;14(1):28 – 36. <http://dx.doi.org/10.1093/procel/pwac013>.
- Wang W, Tian JH, Chen X, Hu RX, Lin XD, Pei YY, et al. Coronaviruses in wild animals sampled in and around Wuhan at the beginning of COVID-19 emergence. *Virus Evol* 2022;8(1):veac046. <http://dx.doi.org/10.1093/ve/veac046>.
- Xiao X, Newman C, Buesching CD, Macdonald DW, Zhou ZM. Animal sales from Wuhan wet markets immediately prior to the COVID-19 pandemic. *Sci Rep* 2021;11(1):11898. <http://dx.doi.org/10.1038/s41598-021-91470-2>.
- Li Y, Zhuang QY, Jiang LJ, Jiang WM, Peng C, Jiang N, et al. Traceable surveillance and genetic diversity analysis of coronaviruses in poultry from China in 2019. *Virus Res* 2021;306:198566. <http://dx.doi.org/10.1016/j.virusres.2021.198566>.
- Li YY, Liu HX, Xia W, Wong GWK, Xu SQ. Cold chain logistics: a possible mode of SARS-CoV-2 transmission? *BMJ* 2021;375:e066129. <http://dx.doi.org/10.1136/bmj-2021-066129>.
- Wu ZQ, Jin Q, Wu GZ, Lu J, Li MK, Guo DY, et al. SARS-CoV-2's origin should be investigated worldwide for pandemic prevention. *Lancet* 2021;398(10308):1299 – 303. [http://dx.doi.org/10.1016/S0140-6736\(21\)02020-1](http://dx.doi.org/10.1016/S0140-6736(21)02020-1).
- Liu WJ, Wu GZ. Convincing the confidence to conquer COVID-19: from epidemiological intervention to laboratory investigation. *Biosaf Health* 2020;2(4):185 – 6. <http://dx.doi.org/10.1016/j.bshealth.2020.11.005>.