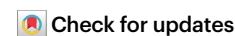


Coronavirus research: knowledge gaps and research priorities

Stanley Perlman & Malik Peiris



Decades of coronavirus research and intense studies of SARS-CoV-2 since the beginning of the COVID-19 pandemic have led to an unprecedented level of knowledge of coronavirus biology and pathogenesis, yet many outstanding questions remain. Here, we discuss knowledge gaps and research priorities in the field.

Introduction

The COVID-19 pandemic showed that, based on previous research efforts, we understood many aspects of coronavirus biology and pathogenesis, but also that there was much we did not know. In 2019, the worldwide number of coronavirus investigators was small, having increased after the severe acute respiratory syndrome coronavirus (SARS-CoV) outbreak in 2003 but decreasing thereafter. The influx of scientists with diverse expertise into the field after the pandemic onset contributed to an increased understanding of coronavirus replication, epidemiology, SARS-CoV-2 pathogenesis and immune responses in humans, to the development and characterization of experimentally infected animal models for COVID-19, and to SARS-CoV-2 vaccine and antiviral drug development. Here, as investigators who have studied coronaviruses for decades, we outline some of the outstanding research questions that we think need to be addressed.

SARS-CoV-2 emergence

Where did SARS-CoV-2 originate and how did it evolve to infect humans? The emergence of SARS-CoV-2 continues to be an area of controversy and has been, and is being, investigated by many national and international organizations, including the WHO (World Health Organization). It is almost certain that the virus originated in bats and crossed species to humans either directly or indirectly via intermediary hosts. There remains debate on whether the virus first infected humans from a zoonotic source or from a research laboratory, but, no matter what the answer to this question is, it is clear to us that in order to be prepared for the next pandemic, we need to further delineate the panoply of coronaviruses present in bats and possible intermediary hosts¹. We need to better understand coronavirus circulation in hotspots, such as parts of China and Southeast Asia, where humans, wildlife gathered for food or medicinal purposes and bats are in close proximity. These investigations should include surveillance (virological and serological) of humans in close contact with bats and the game animal trade, with or without respiratory disease, for evidence of coronavirus infection. A related question, discussed below, is why coronaviruses are especially good at jumping species, to humans and other animals.

Zoonotic risk

Once coronaviruses in animal reservoirs are identified, can they be better risk assessed for threats for human spillover? Surveillance of bat reservoirs of sarbecoviruses (Sarbecovirus is the subgenus to which SARS-CoV-2 belongs) had previously found evidence of viruses with a capacity for infecting human cells using the angiotensin converting enzyme 2 (ACE2) receptor (reminiscent of SARS-CoV)². Serological evidence of viral spillover to humans was demonstrated before the emergence of SARS-CoV-2 (ref.³). Arguably, these signals together should have been triggers for action to develop countermeasures with greater urgency. The availability of human organoid cultures and ex vivo cultures of human respiratory tissue may enable the use of physiologically relevant systems for a more systematic risk assessment of animal coronaviruses in the future, analogous to ongoing risk assessments being carried out for animal influenza viruses⁴.

SARS-CoV-2 transmissibility

What explains the high transmissibility of SARS-CoV-2 compared with SARS-CoV or Middle East respiratory syndrome coronavirus (MERS-CoV)? A critical factor leading to the COVID-19 pandemic was the ability of SARS-CoV-2 to grow to high levels in the upper respiratory tract and therefore to readily transmit to other humans. Titres of SARS-CoV and MERS-CoV in the upper respiratory tract peak at later times after infection⁵, consistent with the ability to interrupt transmission with relevant public health infection-prevention methods. A second, related question is why SARS-CoV and a common cold coronavirus, HCoV-NL63, which both use the same receptor as SARS-CoV-2 (ACE2)⁶, have such different patterns of infection within the human respiratory tract. HCoV-NL63 rarely infects the lower respiratory tract, whereas SARS-CoV preferentially causes pneumonia. These different patterns of infection most likely relate to differences in cell entry, including differences in co-receptor usage, host protease usage or fusogenicity of the spike protein, but there are other possibilities. Understanding these differences will provide information on which coronavirus might be expected to be transmissible and to identify additional targets for therapeutic interventions. Further elucidation of the factors that contribute to virus spread will require additional experimental animal models of coronavirus transmission.

The SARS-CoV-2 outbreak also highlighted the lack of evidence-based data on the transmission of coronaviruses, or indeed respiratory viruses in general, and on which non-pharmaceutical countermeasures (for example, social distancing and masks (surgical versus N99/FFP3 masks)) are effective or not. The SARS-CoV-2 outbreak demonstrated that the only effective control options available in the first months of the pandemic were non-pharmaceutical, but our understanding of the efficacy of specific measures is limited.

Coronavirus genome complexity

Why do coronavirus genomes encode so many more proteins than other RNA viruses? Coronavirus genomes are bigger than those of any other RNA virus, apart from those of related members of the

Nidovirales order. The genomes are so large that they require genomic proofreading activity to avoid error catastrophe⁷. A large genome size may contribute to enhanced cross-species transmission, but, at present, this notion is speculative. In any case, an important question is to understand the function of the many non-structural proteins involved in virus replication. Development of a cell-free or entirely in vitro replication system would facilitate detailed probing of the role of individual proteins in replication and transmission. Efforts to develop such cell-free systems were initiated 40 years ago, but it is only in the past few years with the advent of cryo-electron microscopy and new biochemical approaches that progress has been made. These efforts are expected to complement studies in intact cells, which use high-resolution microscopy and related techniques to analyse macromolecular interactions and function at the subcellular level.

Related to the previous question, why do coronaviruses encode so many proteins with apparent immunoevasive function? Coronaviruses encode a variable number of accessory proteins, the genes of which are intermingled within the structural protein genes located at the 3' end of the genome. For example, SARS-CoV-2 encodes at least six such proteins, with several other putative open reading frames in the genome hypothesized to be expressed and have immunoevasive properties⁸. Confusingly, these genes are often deleted in viruses isolated from infected animals, without apparent loss of virulence. This was shown most clearly in the case of MERS-CoV, in which diverse deletions and insertions in accessory genes were detected in some isolates obtained in Africa from camels, the primary host of the virus⁹. These genetic changes may have unpredictable consequences for virus transmissibility or pathogenesis. Deletion of these genes occasionally leads to increased virulence¹⁰. The variable and sometimes unexpectedly high numbers of these proteins suggest that they have redundant and, perhaps, additional functions. Such redundancy could contribute to cross-species transmission. The genetic instability of MERS-CoV camels in Africa therefore needs to be monitored and evidence for human spillover needs to be continually assessed.

Predictive evolution

Can coronavirus evolution in infected human or other animal hosts be predicted? Coronaviruses readily mutate and recombine as they adapt to a new host. This is well illustrated by the COVID-19 pandemic, in which ancestral strains of SARS-CoV-2 initially mutated to better infect humans, and later evolved to evade the human immune response,

generating a series of variants of concern. Several studies have modelled SARS-CoV-2 evolution but so far it has not been possible to predict how the virus will evolve in the future. Such predictive modelling is recognized to be difficult, but would be very useful in the present pandemic as well as in future coronavirus outbreaks or pandemics for vaccine development, for anticipating clinical disease and pathogenesis, and for risk assessment of animal viruses with zoonotic potential.

Stanley Perlman¹✉ & Malik Peiris^{2,3}

¹Department of Microbiology and Immunology, University of Iowa, Iowa City, IA, USA. ²HKU-Pasteur Research Pole, The University of Hong Kong (HKU), Hong Kong Special Administrative Region, People's Republic of China. ³School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong (HKU), Hong Kong Special Administrative Region, People's Republic of China.

✉ e-mail: Stanley-Perlman@uiowa.edu

Published online: 15 February 2023

References

1. Keusch, G. T. et al. Pandemic origins and a One Health approach to preparedness and prevention: solutions based on SARS-CoV-2 and other RNA viruses. *Proc. Natl Acad. Sci. USA* **119**, e2202871119 (2022).
2. Menachery, V. D. et al. A SARS-like cluster of circulating bat coronaviruses shows potential for human emergence. *Nat. Med.* **21**, 1508–1513 (2015).
3. Li, H. et al. Human-animal interactions and bat coronavirus spillover potential among rural residents in Southern China. *Biosaf. Health* **1**, 84–90 (2019).
4. Cox, N. J., Tock, S. C. & Burke, S. A. Pandemic preparedness and the Influenza Risk Assessment Tool (IRAT). *Curr. Top. Microbiol. Immunol.* **385**, 119–136 (2014).
5. Cevik, M. et al. SARS-CoV-2, SARS-CoV, and MERS-CoV viral load dynamics, duration of viral shedding, and infectiousness: a systematic review and meta-analysis. *Lancet Microbe* **2**, e13–e22 (2021).
6. Hofmann, H. et al. Human coronavirus NL63 employs the severe acute respiratory syndrome coronavirus receptor for cellular entry. *Proc. Natl Acad. Sci. USA* **102**, 7988–7993 (2005).
7. Eckerle, L. D., Lu, X., Sperry, S. M., Choi, L. & Denison, M. R. High fidelity of murine hepatitis virus replication is decreased in nsp14 exoribonuclease mutants. *J. Virol.* **81**, 12135–12144 (2007).
8. Lowery, S. A., Sariol, A. & Perlman, S. Innate immune and inflammatory responses to SARS-CoV-2: implications for COVID-19. *Cell Host Microbe* **29**, 1052–1062 (2021).
9. Chu, D. K. W. et al. MERS coronaviruses from camels in Africa exhibit region-dependent genetic diversity. *Proc. Natl Acad. Sci. USA* **115**, 3144–3149 (2018).
10. Gutierrez-Alvarez, J. et al. Middle East respiratory syndrome coronavirus gene 5 modulates pathogenesis in mice. *J. Virol.* **95**, e01172-20 (2021).

Competing interests

The authors declare no competing interests.