Check for updates

The network effect: studying COVID-19 pathology with the Human Cell Atlas

Sarah Teichmann[iD][1,2 ✉] *and Aviv Regev*[iD][3,4 ✉]

The SARS coronavirus 2 (SARS-CoV-2) pandemic is a global challenge, which the scientific community has tackled by mounting extensive global collaborations, combining expertise in diverse areas and applying cutting-edge techniques, to dissect mechanisms underlying coronavirus disease-19 (COVID-19) pathology. The Human Cell Atlas consortium is emerging as an important contributor to furthering this progress.

Coronavirus disease-19 (COVID-19), caused by SARS coronavirus 2 (SARS-CoV-2), typically manifests as a period of low-grade respiratory tract infection, which progresses to a more severe disease in some patients, and this progression is correlated with certain epidemiological factors, including older age, male sex, habits such as smoking and underlying conditions such as diabetes. In addition to lung injury, SARS-CoV-2 infection can also cause failure in multiple other systems and organs, including the heart, kidney, vasculature and brain, and systemic inflammatory responses, some with delayed onset, for example, in children.

Most studies of COVID-19 to date have focused on understanding SARS-CoV-2 infection at the tissue and organ level. Nevertheless, viral infection and the response to it are events that initiate at the single-cell level, and many cell types, even those that do not express the molecular machinery needed for effective infection, may play key roles in viral pathology. Unlocking the identity of the specific cells involved in infection can clarify transmission patterns, pathogenesis and risk differences between individuals and will be key to developing effective therapeutic approaches. Single-cell approaches, including single-cell RNA-seq (scRNA-Seq) are uniquely poised to achieve this goal.

Excitingly, efforts along these lines have been underway across the world over the recent months, including those enabled by the Human Cell Atlas (HCA). The HCA — currently spanning >1,700 researchers across >70 countries — was launched in 2016 to create comprehensive reference maps of all human cells as a basis for both understanding human health and diagnosing, monitoring and treating disease. Initial work aimed at understanding SARS-CoV-2 pathology at the single-cell level — some of which used previously generated HCA data — focused on understanding which cell types expressed the gene encoding the SARS-CoV-2 receptor, *ACE2*. This revealed that type II alveolar (AT2) cells are likely to be most susceptible to infection in the lung[1], but also that cells in other tissues, including gut[2],

might be susceptible to infection. Inspired in part by these initial studies, the HCA community has in parallel embarked on combining our datasets and expertise to further the understanding of SARS-CoV-2 pathology across the entire human body. Together, we pooled data from millions of cells across the body derived from healthy individuals and coming from dozens of studies — including ongoing studies with unpublished datasets — to identify the tissues where cells expressing genes relevant for SARS-CoV-2 viral entry and activation reside. In this way we attempt to address key questions in SARS-CoV-2 transmission, epidemiology and pathogenesis, with several publications and preprints already being available.

In the context of routes of infection and transmission, we found that in addition to AT2 cells in the lung, corneal epithelial cells and two specific cell types within the nose also show high levels of expression of the *ACE2* gene, suggesting intranasal administration as a potential vaccination or treatment route[3]. The demonstration of *ACE2* expression in gastrointestinal enterocytes implicates the oral–faecal route of transmission, which could explain high rates of transmission of SARS-CoV-2 as compared with other coronaviruses. In addition, analysing multiple datasets of placental and uterine tissue revealed localization of *ACE2* at the uterine–placental interface, indicating a possible route for transmission from mother to baby (supporting previous anecdotal reports)[3].

To understand some of the epidemiological factors associated with COVID-19, especially the impact of age, sex or smoking, we pooled data from 22 studies (16 of them not yet published) to test for associations between cell-type specific expression levels of *ACE2* and each of two proteases also needed for viral entry, and these epidemiological factors[4]. Focusing on data from lung and airways, we found that *ACE2* and accessory protease expression in airway epithelial and alveolar AT2 cells increases with age and is higher in men, and that smoking increases *ACE2* expression in airway epithelial cells, while it may decrease it in AT2 cells. Overall,

[1]*Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton, Cambridge, UK.*

[2]*Theory of Condensed Matter Group, Cavendish Laboratory/ Department of Physics University of Cambridge, Cambridge, UK.*

[3]*Klarman Cell Observatory, Broad Institute of MIT and Harvard, Cambridge, MA, USA.*

[4]*Howard Hughes Medical Institute, Koch Institute of Integrative Cancer Research, Department of Biology, Massachusetts Institute of Technology, Cambridge, MA, USA.*

✉*e-mail: st9@sanger.ac.uk; aregev@broadinstitute.org*

these observations may explain some of the differences in disease susceptibility between individuals, encouraging follow-up studies. Notably, this analysis required large numbers of samples to draw robust statistical associations, necessitating coordinated effort in data sharing, annotation and analysis between many groups. Moreover, because the data were generated using a variety of sample sources and lab protocols, the network had to devise a statistical meta-analysis approach — the first for any single-cell effort — to consolidate the data without requiring re-analysis of the results for each study.

In the context of pathogenesis, the network's broader surveys of 25 tissues, assembling both published and unpublished data, allowed us to identify cells that may be susceptible to viral entry across the body[3]. In addition to gut enterocytes and corneal epithelial cells implicated in transmission, cells in diverse organs including the heart (ventricular cardiomyocytes and heart pericytes and fibroblasts[5,6]), olfactory system (olfactory sustentacular cells), kidney (renal epithelial cells) and brain (oligodendrocytes) expressed *ACE2* and one of the accessory proteases, in concordance with the wide range of clinical symptoms reported for COVID-19. In many cases, these cells are rare, or express *ACE2* at relatively low levels, such that these patterns would likely not be resolved in bulk tissue profiling. Further analysis in the lung, nose[3] and gut suggested that these cells exhibit an accompanying gene expression programme involving genes modulating immunity. Finally, through single-cell analysis, the network revealed that the expression of genes relevant for SARS-CoV-2 infection is conserved in non-human primates and mice, suggesting that they could be suitable models for further study of the biology of the disease[4,7,8].

As the pandemic is sweeping through the world, we have now pivoted to further contribute our expertise through single-cell analysis of samples from patients with COVID-19 — in addition to samples from healthy subjects — including blood samples, biopsies or aspirates from the airways and post mortem tissue. The HCA community is also coupling these analyses with human in vitro systems, such as organoids, and with animal infection models as mentioned above, which can serve to deliver further insights into SARS-CoV-2 pathology at single-cell resolution, including screening potential drugs. These studies will help us to understand the role of immune cells in COVID-19 pathology and their response to potential treatment regimens.

Overall, collaborating at this new scale allows us to perform comprehensive analysis not only within one tissue, but across the entire body, yielding new and important insights into the aetiology of a disease that attacks many organ systems. In addition to rapidly accelerating our understanding of COVID-19 biology, by working together, we have gained valuable lessons in how to do science quickly, at a massive scale, in a time of urgent need. We observe that the pre-existing effort and network of HCA, which already brought together hundreds of researchers across many countries in a joint mission, allowed us to function efficiently by taking advantage of diverse modes of communication. Furthermore, the rapid progress has been supported by the engagement of trainees (students, postdocs and clinical fellows), who through this experience build relationships, trust and a style of open collaboration that offers to sprout many long-term future projects. On the whole, in this global effort, we have benefited from a culture of openness and from shared resources: as we pool many datasets together and combine the expertise and critical input of many analysts, we can draw conclusions that may not be possible to achieve within a single lab, institute, or even national initiative, and befitting a question focused on global health.

Finally, the global coordination that we have established spans beyond the HCA and includes not only collaborative analysis of collected data, but also developing and sharing protocols for specimen collection and processing. In this extended network, the HCA leverages its relationships with other communities, such as the LifeTime Initiative, the Seed Networks and COVID-19-specific projects funded by various philanthropic organizations as well as our colleagues in the human genetics and viral genomics communities. This extended network will additionally broaden the scope of the biological questions we, as a research community, can answer together to help combat the COVID-19 pandemic.

1. Zhao, Y. et al. Single-cell RNA expression profiling of ACE2, the receptor of SARS-CoV-2. Preprint at *bioRxiv* https://doi.org/10.1101/2020.01.26.919985 (2020).
2. Lee, J. J. et al. Relative abundance of SARS-CoV-2 entry genes in the enterocytes of the lower gastrointestinal tract. Preprint at *bioRxiv* https://doi.org/10.1101/2020.04.08.033001 (2020).
3. Sungnak, W. et al. SARS-CoV-2 entry factors are highly expressed in nasal epithelial cells together with innate immune genes. *Nat. Med.* **26**, 681–687 (2020).
4. Muus, C. et al. Integrated analyses of single-cell atlases reveal age, gender, and smoking status associations with cell type-specific expression of mediators of SARS-CoV-2 viral entry and highlights inflammatory programs in putative target cells. Preprint at *bioRxiv* https://doi.org/10.1101/2020.04.19.049254 (2020).
5. Litviňuková, M. et al. Cells and gene expression programs in the adult human heart. Preprint at *bioRxiv* https://doi.org/10.1101/2020.04.03.024075 (2020).
6. Tucker, N. R. et al. Myocyte specific upregulation of ACE2 in cardiovascular disease: implications for SARS-CoV-2 mediated myocarditis. Preprint at *medRxiv* https://doi.org/10.1101/2020.04.09.20059204 (2020).
7. Han, L. et al. Single-cell atlas of a non-human primate reveals new pathogenic mechanisms of COVID-19. Preprint at *bioRxiv* https://doi.org/10.1101/2020.04.10.022103 (2020).
8. Ziegler, C. et al. SARS-CoV-2 receptor ACE2 is an interferon-stimulated gene in human airway epithelial cells and is enriched in specific cell subsets across tissues. *Cell* https://doi.org/10.2139/ssrn.3555145 (2020).

Author contributions
The authors contributed equally to all aspects of the article.

RELATED LINKS
Currently published reports, non-peer reviewed preprints and resources on COVID-19 generated within the Human Cell Atlas network: https://www.humancellatlas.org/covid-19/
Human Cell Atlas: https://www.humancellatlas.org/
Human genetics community: https://www.covid19hg.org/
LifeTime Initiative: https://lifetime-fetflagship.eu/
Seed Networks: https://chanzuckerberg.com/science/programs-resources/humancellatlas/seednetworks/
Viral genomics community: https://nextstrain.org/ncov/global