

Commentary

Multi-omics in COVID-19: Seeing the unseen but overlooked in the clinic

Tian Lu,^{1,2,3} Yingrui Wang,^{1,2,3} and Tiannan Guo^{1,2,3,*}¹Westlake Laboratory of Life Sciences and Biomedicine, Key Laboratory of Structural Biology of Zhejiang Province, School of Life Sciences, Westlake University, Hangzhou, Zhejiang Province, China²Institute of Basic Medical Sciences, Westlake Institute for Advanced Study, Hangzhou, Zhejiang Province, China³Center for Infectious Disease Research, Westlake University, 18 Shilongshan Road, Hangzhou 310024, Zhejiang, China*Correspondence: guotiannan@westlake.edu.cn<https://doi.org/10.1016/j.xcrm.2022.100580>

COVID-19 is an ongoing pandemic of global concern and is unlikely to disappear. This commentary discusses how multi-omics technologies have helped uncover the molecular processes and dynamics underlying COVID-19 initiation, progression, and transmission, and how lack of standardization has limited their application in clinical settings.

As of February 12, 2022, more than 200 countries and territories have reported over 405 million confirmed cases of COVID-19, and 5.8 million recorded deaths. While vaccinations have not stopped the spread of SARS-CoV-2, they have reduced the risk of serious disease and death in adults and children. Investigations deploying multi-omics technologies reveal the underlying molecular structure of the pathogen and molecular host responses to the virus and vaccines (Figure 1 and 2).

The pathogen

Our understanding of SARS-CoV-2, as well as host responses to the virus, have been greatly deepened by multi-omics technologies including, but not limited to, next-generation sequencing (NGS) and proteomic and metabolomic approaches (Figure 2). Indeed, the virus causing COVID-19 was first identified by metagenomic RNA sequencing. Based on the RNA sequence of the virus genome, PCR-based assays were developed and rapidly applied in the clinic to replace initial temperature tests to diagnose COVID-19.

Multiple mutations in the SARS-CoV-2 sequence lead to the change of pathogenicity, infectivity, transmissibility, and/or antigenicity.¹ E484K, for example, was identified as an escape mutation that could reduce antibody-mediated neutralization.¹ The discovery of the Delta and Omicron variants relied on the prompt application of NGS techniques, which also provided

crucial information to track the spread of these variants across different geographic areas. A study from Zimbabwe of the genomic epidemiology of the SARS-CoV-2 variants found that 60% of cases were imported, indicating that human movement is a key factor in its transmission, further supporting the importance of quarantine and restriction of human movements.² Rational development of COVID-19 vaccines also relies on mass-spectrometry (MS)-based characterization of the spike (S) glycoprotein of the virus, whose conformational dynamics is key for vaccine design. The viral S protein interacts with ACE2 on host cells, and antibodies that interfere with this by targeting the S protein can potentially neutralize the virus. Watanabe et al. expressed recombinant spike trimers and determined the glycan compositions for 22 N-linked glycan sites by MS (e.g., principally oligomannose-type on N234 and N709), which not only deepens our understanding of the S protein as a vaccine target, it also provides a benchmark to assess the quality of immunogens in the development of vaccines and therapeutic antibodies.³

Host responses to the pathogen

Despite the progress we have made in understanding this virus, we still struggle to predict which patients will develop clinically severe COVID-19 illness. Should there be a precise and practical means to stratify patients according to disease severity and to identify the majority of patients who will survive SARS-CoV-2

infection, many concerns of the global pandemic could be relieved. Further, are there differences in immune responses to the emerging variants of concern? How do different vaccines and multiple vaccine doses impact anti-SARS-CoV-2 immunity? Addressing these and many other crucial questions requires a comprehensive understanding of host responses to this pathogen and to vaccines. Most multi-omics studies are focused on elucidating host responses to the pathogen in various organs and clinical specimens that are otherwise hidden using conventional approaches.

COVID-19 host response studies were initially limited to the enumeration of clinical symptoms such as fever and cough and further extended to chest CT and circulating protein biomarkers such as CRP and SAA1, which had been used for other infectious diseases and empirically borrowed to monitor clinical progression of COVID-19. Multi-omics studies have greatly expanded our views on circulating molecular changes, but these are practically undetectable with conventional analytical methodologies (Figure 2). For instance, proteomic and metabolomic characterization of COVID-19 sera were effectively applied to measure about 2000 circulating molecules and identified 93 proteins and 204 metabolites that were significantly and specifically dysregulated in severe cases.⁴ This systematic investigation highlighted the critical roles of platelet degranulation, macrophage function, and



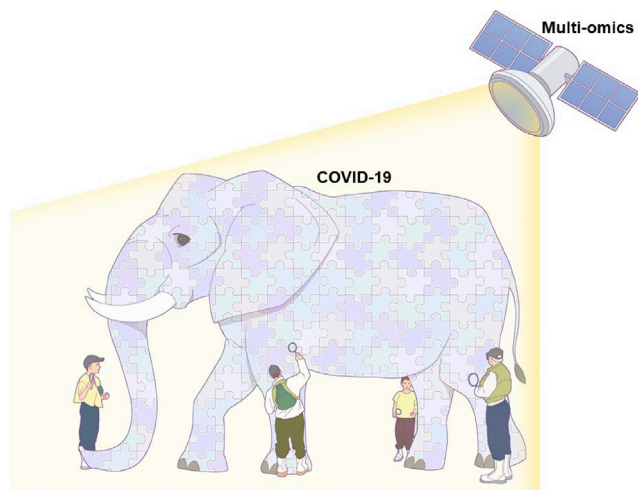


Figure 1. Multi-omics enables systematic characterization of molecular modulations in COVID-19

The elephant refers to a patient with COVID-19, while the magnifiers are conventional technologies for examining limited number of molecules. Multi-omics technologies, as indicated as the satellite, comprehensively profile molecular modulations, most of which could not be seen by the magnifiers.

the complement system. Most of the key molecular changes have been confirmed in other multi-omics studies of COVID-19 from multiple countries, further validating the technical reliability of the proteomic and metabolomic techniques and the potential value of these findings in clinical decision-making.

Remarkably, proteins and metabolites can be readily measured by mass spectrometry also in urine, a non-invasive clinical biospecimen. Close to 4000 proteins were detected in urine, while only about 1500 proteins were detected in serum using the same MS method due to the presence of high-abundance proteins in serum. Notably, 80% of the thus measured serum proteins were detectable in urine, while only 31% of urinary proteins were detected in serum.⁵

But why do we need to measure these molecules? How can these findings contribute to the management of COVID-19? Multiple omics studies have shown that in properly designed omics experiments, machine learning can be used to identify biomarkers that can be employed to classify and monitor disease progression; for example, a random forest model based on 22 proteins and 7 metabolites expressed in sera could reliably indicate the severity of COVID-19, and to a certain extent, predict the disease prognosis.⁴ Urinary protein-derived models performed as well as those from sera and

could be used to monitor COVID-19 disease progression.⁵

In addition to circulating biomarkers of host responses, local responses in multiple solid organs are also crucial. Histopathological examination of (sectioned) tissues provides the most relevant information at macroscopic and microscopic resolution and is widely accepted as the “gold standard” for disease diagnosis. However, histopathology of COVID-19 tissue specimens are not distinctive but rather highly similar to those described in SARS-CoV and Middle East respiratory syndrome coronavirus (MERS-CoV) patients.⁶ MS-based proteomics now allows characterization of over 10,000 proteins across multi-organ autopsy samples and the identification of multiple dysregulated proteins involved in coagulation, angiogenesis, fibrosis, and fatty acid metabolism.⁷ Without in-depth proteomics techniques, it will take much more effort and time, following the hypothesis-driven research paradigm, to uncover the potential link between upregulation of Cathepsin L in the lung and COVID-19-related mortality or the association between decreased INSL3 protein and impaired Leydig cells in the testis.⁷ Proteomics allows us to “see” molecular changes underlying the morphological observations. Tissue specimens, in particular lung tissues, are potentially infectious and must be inactivated and sterilized, usually by formalin fix-

ation, before analytical analysis, which makes transcriptomic analysis technically challenging due to the instability of mRNAs. In contrast, proteins could be effectively retrieved from formalin-fixed tissues for MS-analysis.⁸

Omics technologies also enable the identification of target cells of the virus, which is hardly achievable with conventional techniques. Single-cell RNA sequencing (scRNA-seq) analysis of nasopharyngeal swabs showed ciliated cell losses with secretory, deuterosomal expansion, and increase of macrophages during COVID-19.⁹ No alternative technology could directly provide such insights in clinical specimens. scRNA-seq also deconvolutes mixed cells. For example, lymphopenia, the substantial decrease of lymphocytes circulating in the blood, is a common symptom in patients with COVID-19. Various lymphocytes play diverse roles in anti-virus immunity. scRNA-seq analysis of the lymphocytes uncovered three patterns of lymphocyte responses associated with different clinical outcomes.¹⁰

Host responses to vaccines

Multi-omics studies could also provide potentially invaluable guidance for the use of vaccines, but unfortunately, such data are still inadequate. Over 10 billion vaccine doses have been administered globally, including mRNA vaccines, inactivated viral particles, and adenoviral-based vaccines, among others. While all have provided some protection against SARS-CoV-2, the mRNA vaccines elicit a higher titer of neutralizing antibodies than the others.¹¹ However, this protection is mediated not only by the neutralizing antibodies, but also other immune effector mechanisms including T cells and innate immune cells. For example, mRNA vaccines showed 80% efficacy against symptomatic infection even after only one dose in the absence of detectable neutralizing antibodies.¹¹ Moreover, up to 6% of recipients are seronegative after the second injection.¹¹ According to the US Centers for Disease Control and Prevention (www.cdc.gov), as of February 12, 2022, more than 0.004% of recipients experienced severe side effects, including anaphylaxis, thrombosis with thrombocytopenia syndrome,

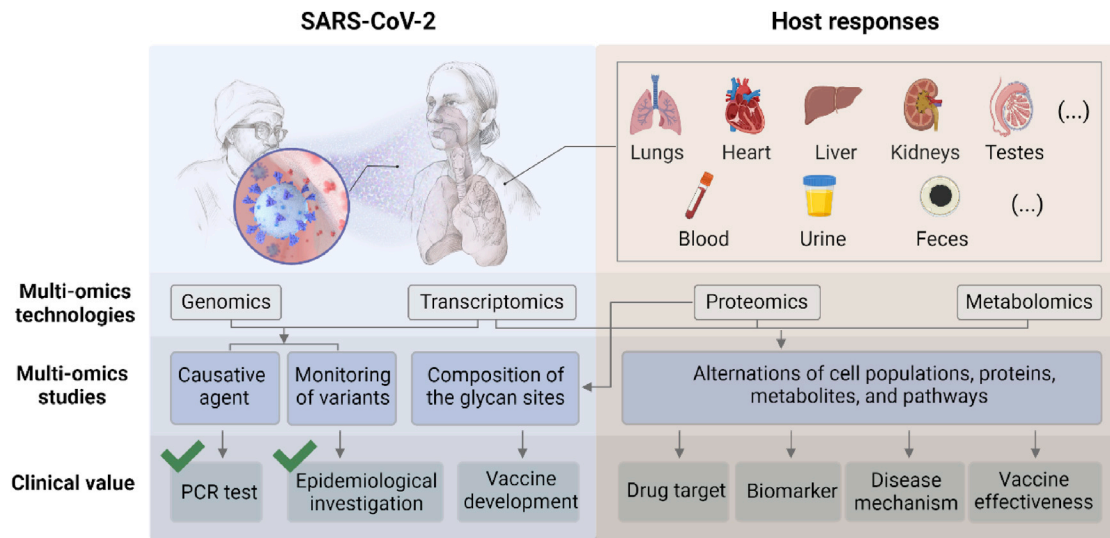


Figure 2. Multi-omics studies on SARS-CoV-2 and host response

Adapted from “Potential Transmission Mechanisms of SARS-CoV-2 Variants of Concern,” by BioRender.com (2022). Retrieved from <https://app.biorender.com/biorender-templates>. Studies of the viral genome and proteome led to rapid establishment of PCR tests for COVID-19 diagnosis. The tracking of the genomes enabled epidemiological investigation of the virus evolution and spreading. Multiple vaccines have been designed based on the sequence and surface proteins. Most multi-omics analyses are focused on the host responses toward infection or vaccination. In this regard, proteomics and metabolomics are the major workhorses.

Guillain-Barré Syndrome, and even death. Circulating neutralizing antibodies decline over time. Breakthrough infections occur in vaccinated individuals, especially from emerging and immune escape variants such as Omicron.¹² Customized vaccination against specific variants has not been realized to date, although its importance in preventing COVID-19 is undeniable. A controversial but pragmatic policy to address the uncertainties in protective immunity is to make booster vaccination mandatory.

All these are limitations of the current vaccination strategies that are being implemented for billions of individuals in many countries. Timely monitoring of serological host responses is theoretically informative for epidemiological tracking in a population and could potentially offer useful information to guide vaccine dosage and dose spacing. However, even circulating neutralizing antibody tests have not been widely adopted to guide clinical decision making, probably because of the limited practicality of mass blood sampling and the absence of universal standards. Monitoring urinary proteins might be a more practical means to evaluate immunity against the virus; however, no published data are available yet.

Why are proteomics and metabolomics largely overlooked in the clinic?

While close to 10,000 COVID-19-related omics papers have been published (PubMed search as of February 12, 2022), no findings from omics studies except for PCR and genomic analysis of the virus have been successfully translated for clinical management of COVID-19. One possible reason is that multi-omics technologies, mainly proteomics and metabolomics, are not mature, reproducible, and sufficiently robust to provide clinically valid results. However, recent proteomics data collated from many laboratories globally show dysregulated molecules to be surprisingly consistent. In contrast, false negatives are notoriously frequent in PCR tests of SARS-CoV-2. A recent study of 95,919 patients reported a false-negative rate (FNR) and sensitivity of 9.3% (95% CI 1.5%–17.0%) and 90.7% (95% CI 82.6%–98.9%), respectively.¹³ Some may argue that proteins are too vulnerable to be robust biomarkers in real-world applications. However, data have shown that proteins as measured by mass spectrometry are more stable than transcripts measured by NGS.¹⁴

Is it because the measurement of proteins and metabolites is too technically and analytically challenging, and expensive? Indeed, the acquisition of proteomics and metabolomics data involves multiple steps and requires special expertise and expensive instruments, such as high-resolution mass spectrometers for proteomics. However, these requirements should not be insurmountable given that much more sophisticated and expensive techniques such as positron emission tomography (PET) are already widely used clinically. Moreover, the cost of measuring a protein by mass spectrometry has dropped from about \$3 in 2006 to less than \$0.1 in 2020.¹⁵

The major hurdle delaying clinical applications of proteomics and metabolomics is probably the lack of standardization. Few research laboratories in the rapidly developing field of proteomics use consensus standard operating protocols (SOPs), while the situation in metabolomics is even more confusing because even raw data files generated in different laboratories are rarely shared due to commercial considerations. Although commercialization of metabolomics preceded that of proteomics, most metabolomic platforms developed closed proprietary resources (e.g., compound libraries for molecular

measurement) and techniques and do not share their raw data even for scientific communications. It is unlikely that this daunting situation can be easily reversed to achieve the necessary transparency and standardization. In contrast, the proteomics community was founded on an ethos of scientific transparency and open sharing. We are optimistic that several international consensus SOPs for clinical proteomics will be developed in the near future through concerted efforts of proteomics scientists, clinical practitioners, and other stakeholders.

Other challenges

The ability to interrogate biomedical specimens on diverse omic platforms poses new challenges of integrating multi-omics data. Genomic data are mostly qualitative, while the expression data of transcripts, proteins, and metabolites are quantitative. Therefore, it is inherently not straightforward to integrate them by algorithms. Furthermore, unlike the genome which is basically identical among different tissue specimens of an individual, proteomes and metabolomes are tissue-specific and context-dependent, reflecting the actual pathophysiological state and complexity of biology and diseases, thus posing great challenges to identify the critical molecular changes over space and time.

Besides these technical limitations, collection of potentially infectious samples during the pandemic poses additional challenges. The most common samples used in COVID-19 studies are blood or bronchoalveolar lavage fluid samples, while data from non-invasive samples like urine and feces have gradually proved to be useful in reflecting host responses. However, there remains a lack of standard operating procedures (SOPs) for hospitals to safely collect and competently pretreat these samples for multi-omics research, which can be a source of bias within a cohort and across different cohorts in separate studies. Biosafety practices restrict access to COVID-19 samples to minimize the risk of accidental virus transmission but also limit comprehensive investigation of these precious specimens to understand and

control COVID-19. Had proper and technically feasible SOPs been established for processing these potentially infectious specimens, our understanding of COVID-19 would have progressed much faster. COVID-19 is unlikely to disappear, nor will it be the cause of the last pandemic. Therefore, there is no place for complacency. The global community must act now to be better prepared for the future.

ACKNOWLEDGMENTS

This work is supported by grants from National Key R&D Program of China (2021YFA1301602, 2020YFE0202200), the National Natural Science Foundation of China (81972492, 21904107), Zhejiang Provincial Natural Science Foundation for Distinguished Young Scholars (LR19C050001), Hangzhou Agriculture and Society Advancement Program (20190101A04), and Westlake Education Foundation. We thank Dr. O.L. Kon for insightful revisions and Mr. Q. Xiong for assisting in figure preparation.

DECLARATION OF INTERESTS

T.G. is a shareholder of Westlake Omics, Inc., and a member of the advisory board of the journal. The remaining authors declare no competing interests.

REFERENCES

- Harvey, W.T., Carabelli, A.M., Jackson, B., Gupta, R.K., Thomson, E.C., Harrison, E.M., Ludden, C., Reeve, R., Rambaut, A., Peacock, S.J., and Robertson, D.L.; COVID-19 Genomics UK (COG-UK) Consortium (2021). SARS-CoV-2 variants, spike mutations and immune escape. *Nat. Rev. Microbiol.* **19**, 409–424.
- Mashe, T., Takawira, F.T., de Oliveira Martins, L., Gudza-Mugabe, M., Chirenda, J., Munyanyi, M., Chaibva, B.V., Tarupiwa, A., Gumbo, H., Juru, A., et al.; COVID-19 Genomics UK Consortium; SARS-CoV-2 Research Group (2021). Genomic epidemiology and the role of international and regional travel in the SARS-CoV-2 epidemic in Zimbabwe: a retrospective study of routinely collected surveillance data. *Lancet Glob. Health* **9**, e1658–e1666.
- Watanabe, Y., Allen, J.D., Wrapp, D., McLellan, J.S., and Crispin, M. (2020). Site-specific glycan analysis of the SARS-CoV-2 spike. *Science* **369**, 330–333.
- Shen, B., Yi, X., Sun, Y., Bi, X., Du, J., Zhang, C., Quan, S., Zhang, F., Sun, R., Qian, L., et al. (2020). Proteomic and Metabolomic Characterization of COVID-19 Patient Sera. *Cell* **182**, 59–72.e15.

- Bi, X., Liu, W., Ding, X., Liang, S., Zheng, Y., Zhu, X., Quan, S., Yi, X., Xiang, N., Du, J., et al. (2022). Proteomic and metabolomic profiling of urine uncovers immune responses in patients with COVID-19. *Cell Rep.* **38**, 110271.
- Deshmukh, V., Motwani, R., Kumar, A., Kumari, C., and Raza, K. (2021). Histopathological observations in COVID-19: a systematic review. *J. Clin. Pathol.* **74**, 76–83.
- Nie, X., Qian, L., Sun, R., Huang, B., Dong, X., Xiao, Q., Zhang, Q., Lu, T., Yue, L., Chen, S., et al. (2021). Multi-organ proteomic landscape of COVID-19 autopsies. *Cell* **184**, 775–791.e14.
- Zhu, Y., Weiss, T., Zhang, Q., Sun, R., Wang, B., Yi, X., Wu, Z., Gao, H., Cai, X., Ruan, G., et al. (2019). High-throughput proteomic analysis of FFPE tissue samples facilitates tumor stratification. *Mol. Oncol.* **13**, 2305–2328.
- Ziegler, C.G.K., Miao, V.N., Owings, A.H., Navia, A.W., Tang, Y., Bromley, J.D., Lotfy, P., Sloan, M., Laird, H., Williams, H.B., et al. (2021). Impaired local intrinsic immunity to SARS-CoV-2 infection in severe COVID-19. *Cell* **184**, 4713–4733.e22.
- Mathew, D., Giles, J.R., Baxter, A.E., Oldridge, D.A., Greenplate, A.R., Wu, J.E., Alanio, C., Kuri-Cervantes, L., Pampena, M.B., D'Andrea, K., et al.; UPenn COVID Processing Unit (2020). Deep immune profiling of COVID-19 patients reveals distinct immunotypes with therapeutic implications. *Science* **369**, eabc8511.
- Sadarangani, M., Marchant, A., and Kollmann, T.R. (2021). Immunological mechanisms of vaccine-induced protection against COVID-19 in humans. *Nat. Rev. Immunol.* **21**, 475–484.
- Kuhlmann, C., Mayer, C.K., Claassen, M., Mampanga, T., Burgers, W.A., Keeton, R., Riou, C., Sutherland, A.D., Suliman, T., Shaw, M.L., and Preiser, W. (2022). Breakthrough infections with SARS-CoV-2 omicron despite mRNA vaccine booster dose. *Lancet* **399**, 625–626.
- Kanji, J.N., Zelyas, N., MacDonald, C., Pabbaraju, K., Khan, M.N., Prasad, A., Hu, J., Diggle, M., Berenger, B.M., and Tipples, G. (2021). False negative rate of COVID-19 PCR testing: a discordant testing analysis. *Virology* **18**, 13.
- Shao, W., Guo, T., Toussaint, N.C., Xue, P., Wagner, U., Li, L., Champi, K., Zhu, Y., Wu, J., Buljan, M., et al. (2019). Comparative analysis of mRNA and protein degradation in prostate tissues indicates high stability of proteins. *Nat. Commun.* **10**, 2524.
- Xiao, Q., Zhang, F., Xu, L., Yue, L., Kon, O.L., Zhu, Y., and Guo, T. (2021). High-throughput proteomics and AI for cancer biomarker discovery. *Adv. Drug Deliv. Rev.* **176**, 113844.