

SARSCOVIDB—A New Platform for the Analysis of the Molecular Impact of SARS-CoV-2 Viral Infection

Rafael Lopes da Rosa, Tung Sheng Yang, Emanuela Fernanda Tureta, Laura Rascovetzki Saciloto de Oliveira, Amanda Naiara Silva Moraes, Juliana Miranda Tatara, Renata Pereira Costa, Júlia Spier Borges, Camila Innocente Alves, Markus Berger, Jorge Almeida Guimarães, Lucélia Santi,* and Walter Orlando Beys-da-Silva*

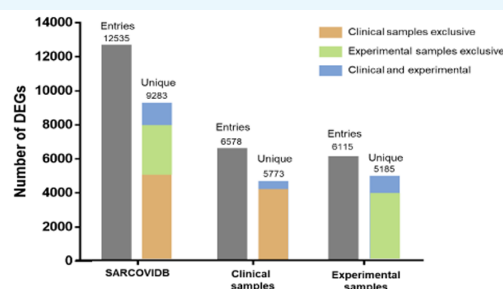
Cite This: *ACS Omega* 2021, 6, 3238–3243

Read Online

ACCESS |

Metrics & More

Article Recommendations



ABSTRACT: The COVID-19 pandemic caused by the new coronavirus (SARS-CoV-2) has become a global emergency issue for public health. This threat has led to an acceleration in related research and, consequently, an unprecedented volume of clinical and experimental data that include changes in gene expression resulting from infection. The SARS-CoV-2 infection database (SARSCOVIDB: <https://sarscovidb.org/>) was created to mitigate the difficulties related to this scenario. The SARSCOVIDB is an online platform that aims to integrate all differential gene expression data, at messenger RNA and protein levels, helping to speed up analysis and research on the molecular impact of COVID-19. The database can be searched from different experimental perspectives and presents all related information from published data, such as viral strains, hosts, methodological approaches (proteomics or transcriptomics), genes/proteins, and samples (clinical or experimental). All information was taken from 24 articles related to analyses of differential gene expression out of 5,554 COVID-19/SARS-CoV-2-related articles published so far. The database features 12,535 genes whose expression has been identified as altered due to SARS-CoV-2 infection. Thus, the SARSCOVIDB is a new resource to support the health workers and the scientific community in understanding the pathogenesis and molecular impact caused by SARS-CoV-2.

INTRODUCTION

In December 2019, a new strain of coronavirus associated with severe acute respiratory syndrome was identified as SARS-CoV-2.¹ This is an encapsulated, single-stranded RNA virus from the *Coronaviridae* family, presenting high virulence and generating a significant global impact.² The COVID-19 pandemic caused by SARS-CoV-2 had as its epicenter the city of Wuhan, China, and in a short time became a serious public health problem worldwide.³ As of end-October, it is present in more than 200 countries, accounting for 45 million cases.⁴ In the USA, the country with the highest number of deaths, more than 8 million cases and 220,000 deaths have been recorded.⁴ The same dramatic outcome of the pandemic can be noticed in Brazil, the country with the second highest

number of COVID-19 deaths, with more than 5 million cases and 150,000 deaths.⁵

The increasing proliferation of SARS-CoV-2-infection cases and the lack of specific therapeutics and vaccines have caused concern among public authorities and international agencies, resulting in the mobilization of the scientific community to understand the disease and clinical outcomes aiming to improve treatments and find ways to prevent COVID-19

Received: November 23, 2020

Accepted: January 13, 2021

Published: January 21, 2021



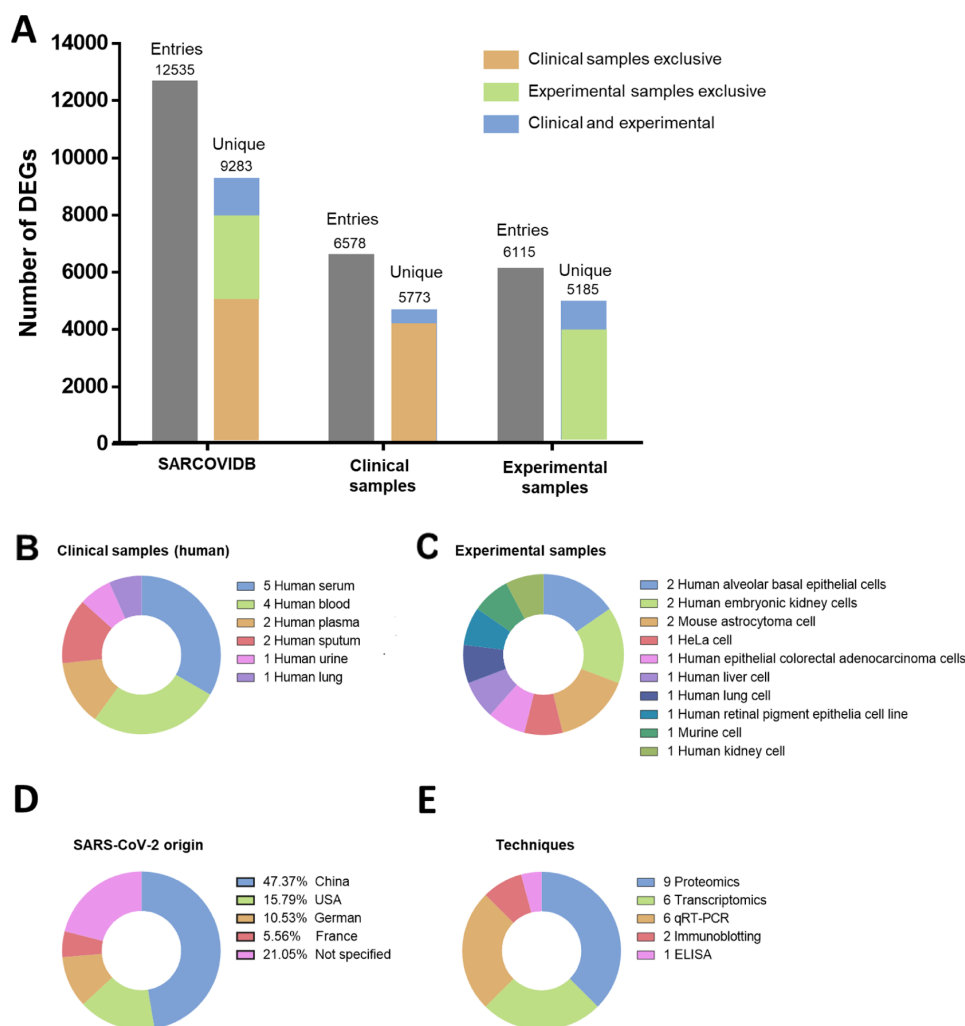


Figure 1. SARSCOVID summary of contents. (A) Number of DEGs identified in clinical and experimental samples; (B) number of articles containing clinical samples; (C) number of articles containing experimental samples; (D) number of articles according to the reported origin of SARS-CoV-2 strain; (E) number of articles according to the methodological approach.

cases. This current scenario has generated significant changes in the field, such as a rapid and unprecedented increase in the number of scientific articles being published. It was possible through facilitated submission processes and preprint publications, without peer evaluation, to disseminate articles with very speculative results, such as computational predictions.⁶

Understanding of the molecular aspects associated with COVID-19 and the search for understanding the complex response of the host after viral infection have been gaining space.^{7–9} Approaches that evaluate viral infections from the perspective of the molecular impact on the host contribute to the understanding of the disease, which is important for outlining potential antiviral strategies.¹⁰ Analysis of differentially expressed genes (DEGs), which allow *in silico* characterization of the molecular response and the impact of infection, can be applied. Some databases contain a collection of expression data, but their results are mostly obtained by text mining with automatic and semiautomatic approaches, which

may lead to nonaccurate data being deposited.¹¹ Furthermore, the results may present redundancy and ambiguity, and a postanalysis is necessary for data conference.¹² In the same way, the need for bioinformaticians to extract raw data in many of the available databases is another important limitation of its use; it can be especially limiting for medical workers without bioinformatics skills who are facing the COVID-19 impact in real time and need information promptly. This scenario hampers the access of the medical community and also other nonspecialized scientists who may need this information in their research. Recently, our group developed the ZIKAVID, a database based on gathering all up-to-date gene expression data generated after Zika virus (ZIKV) infection, containing different experimental approaches, hosts and strains, and other related information.¹²

In this work, we present the SARS-CoV-2 virus infection database (SARSCOVIDB: <https://sarscovidb.org/>), a public database containing all DEGs identified in SARS-CoV-2 infection and COVID-19 samples, manually developed and

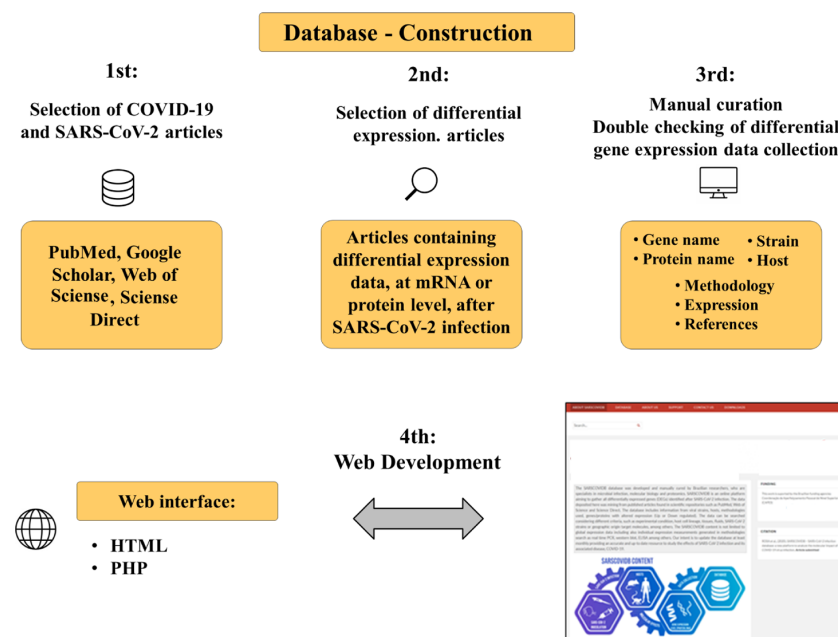


Figure 2. Database construction steps.

checked, with a friendly interface that is easy to navigate. This database will help researchers worldwide, and general users, to speed up the research and understanding of the molecular impact of COVID-19 and possible clinical outcomes.

RESULTS AND DISCUSSION

The outbreak of COVID-19 worldwide, linked to the lack of efficient treatments and approved vaccines, triggered a great effort by the scientific community and governments toward research involving SARS-CoV-2 and the potential comorbidities associated in humans.¹³ For this reason, the SARSCOVVIDB was created, comprising all data from DEGs identified after SARS-CoV-2 infection to date.

The database was initially built by searching the specific terms “COVID” and “SARS-CoV-2”, with a manual double-check for differential expression of genes or proteins after SARS-CoV-2 infection, regardless of the host. To increase the search and the user interface, all data were categorized according to experimental approach, as described above, including the reference article. The SARSCOVVIDB is a database to exclusively gather DEGs after SARS-CoV-2 infection and thus is an important resource to be explored by the scientific community and COVID-19 medical workers in this topic of urgent need. So far, the SARSCOVVIDB contains 12,535 differential DEG entries and 9,283 unique genes. These data comprise different experimental approaches with distinct objectives, the majority being obtained from clinical samples (Figure 1A). Thus, users can easily consult the most frequently reported host models and compare specific gene sets and other information from all published studies with differential expression after SARS-CoV-2 infection. This can facilitate the planning of new experiments, accelerating understanding of the data available and contributing to accelerate understanding of the molecular impact of COVID-19. These expressive numbers in such a short time span of only a few months also highlight the great efforts made by the

scientific community to study SARS-CoV-2 infection and COVID-19. The ZIKAVID covered the same kind of data but on ZIKV; however, this generated a smaller number of clinical samples of ZIKV-infected patients,¹² even though the epidemic also reached a global threat status¹⁴ and caused great concern worldwide in 2015–2017.

Regarding the samples used, most are peripheral blood samples (five serum, four blood, and two plasma samples) and one lung sample (Figure 1B). On the other hand, the experimental samples used different cell lines, most of them from humans (Figure 1C). As observed, many scientific data have been generated using cell lines and other models but there are still no robust animal models that faithfully replicate the pathogenesis of SARS-CoV-2.¹⁵ Thus, the lack of comprehensive studies comparing different hosts and their responses to infection made the survey of data gathered in the SARSCOVVIDB an important alternative to planning future experiments regarding the choice of a more meaningful experimental model.

The SARSCOVVIDB comprises data from SARS-CoV-2 isolates from different geographic regions (Figure 1D). However, studies have still given little importance to the impact that the mutations suffered by strains in different regions may have on disease dynamics and virulence, as occurred with ZIKV for instance.¹⁶ Interestingly, almost 50% of the SARS-CoV-2 isolates or clinical samples studied are from China, followed by America (around 16%) and Europe (around 16%). It is important to highlight that depending on the origin, viral isolates of SARS-CoV-2 can lead to different pathological impacts and mortality rates, as previously suggested.¹⁷ The most recent example of a virus that had a differential clinical impact, depending on the origin, was ZIKV, where Brazilian isolates were strongly associated with severe neurological data.¹⁸ Thus, the SARSCOVVIDB contributes to the study of the pathology on the origin of the virus, once it is

possible to promptly cross-reference expression data in similar models using viruses from different sources.

Currently, there are various databases gathering genomic, transcriptomic, and proteomic data on viruses and their impact on the host.^{19,20} Several databases were developed to obtain host–pathogen interaction data (gene expression and protein interaction), such as VirHostNet 2.0,²¹ Virhostome,²² the HIV-1 human interaction database,²³ the Gene Expression Omnibus,²⁴ and the Virus Pathogen Resource (ViPR).¹⁹ On SARS-CoV-2, there are already two databases: Coronascape,²⁵ available at Metascape database, and The COVID-19 Drug and Gene Set Library,²⁶ which contain a collection of drug and gene sets related to COVID-19. Although not the first SARS-CoV-2 database, SARSCOVIDB is certainly the first to devote itself exclusively to this specific focus on differential expression after the SARS-CoV-2 infection. Most of the data generated by other databases are automated by concentrating generic information, which can result in inaccurate information.²⁷ The SARSCOVIDB proposes to fill this gap in a simple, organized, and objective way by curating the data manually, making it reliable. Furthermore, the user, who does not need a bioinformatic background or skills, can query data from different experimental perspectives, such as the methodology used, host, virus strain, and so forth.

CONCLUSIONS

The SARSCOVIDB is the first database to gather all data to date from differential expression analyses after SARS-CoV-2 infection, with manual checking of the data to give more accurate and faster analysis. Users do not need to have a background in bioinformatics because the user-friendly and simple interface enables all search possibilities to be explored. This allows the user to cross-reference data for the best understanding and analysis of the deposited data. Finally, the SARSCOVIDB is a promising tool for supporting scientists and medical professionals carrying out research and analysis on the molecular mechanisms of SARS-CoV-2 infection.

MATERIALS AND METHODS

Original Article Selection. The SARSCOVIDB (available at: <https://sarscovidb.org/>) comprises differential gene expression measurements, at mRNA and protein levels, built through four main steps (Figure 2). The first step was a manual search by manual text mining to find all articles available in PubMed, Web of Science, Google Scholar, and ScienceDirect databases containing the terms “SARS-CoV-2” and “COVID-19”. Only accepted/published manuscripts were used as the source of DEGs in SARS-CoV2 infection. This search retrieved 5554 articles, from December 2019 to date. The second step comprised a manual check of abstracts to select articles containing only differential gene expression measurements after SARS-CoV-2 infection. All references were double-checked by two independent individuals, resulting in the selection of 24 articles.

Data Collection and Related Information. The data collected from selected papers were checked and organized, comprising a list of DEGs at mRNA and/or protein level identified after SARS-CoV-2 infection. Other information was also collected, such as the type of study (*in vivo*, *in vitro*, and clinical), methodological approach (transcriptomic, proteomic, qRT-PCR, and immunoblotting), viral strain, hosts (clinical or animal model and cell culture), and expression status (see

Table 1). The SARSCOVIDB also contains information from data deposited in a repository or database. The database will be

Table 1. General Information Stored in the SARSCOVIDB

SARSCOVIDB ID	reference identification of the entry in the SARSCOVIDB.
gene name	official gene name for each stored entry.
protein name	official protein name for each stored entry.
host/sample	host model used to generate SARS-COV-2 infection data.
sample	description of the sample, such as if it is a human clinical sample (blood, urine, serum, etc.) or experimental sample (cell line or animal tissue tested <i>in vitro</i>).
virus reference	SARS-CoV-2 strain reference or geographic origin of the SARS-CoV-2 viral isolate or clinical sample.
methods	methods used to measure gene expression at the level of RNA and/or protein after SARS-CoV-2 infection.
expression	indicates the identified gene expression status (up- and/or downregulated) in response to SARS-CoV-2 infection.
article	article title used as a data source.
DOI	direct link to the original reference where the data came from.
type of sample	indicates if the data were generated using clinical or experimental samples.

updated at least monthly by considering the availability of new data from published articles.

Webpage Construction and User Interface. The last step in the SARSCOVIDB was webpage development (Figure 1). MySQL v5.0, PHP v.5.2.99, and HTML10 were used to build the database and the graphical user interface. The data were stored and maintained by the relational database management system (RDBMS) server, with the SQL language used for data management. The search can be customized easily by the user, combining specific proteins with hosts or selecting a specific viral strain or geographic origin. The database also provides step-by-step links to instruct the user how to search. A list containing all genes deposited in the SARSCOVIDB is available through direct download (<https://sarscovidb.org/download/>).

AUTHOR INFORMATION

Corresponding Authors

Lucélia Santi – Programa de Pós-Graduação em Biologia Celular e Molecular, Universidade Federal do Rio Grande do Sul, Porto Alegre, Rio Grande do Sul 91501-970, Brasil; Faculdade de Farmácia, Universidade Federal do Rio Grande do Sul, Porto Alegre, Rio Grande do Sul 90610-000, Brasil; Phone: +55 (51) 3308-5304; Email: lucelia.santi@ufrgs.br

Walter Orlando Beys-da-Silva – Programa de Pós-Graduação em Biologia Celular e Molecular, Universidade Federal do Rio Grande do Sul, Porto Alegre, Rio Grande do Sul 91501-970, Brasil; Faculdade de Farmácia, Universidade Federal do Rio Grande do Sul, Porto Alegre, Rio Grande do Sul 90610-000, Brasil; orcid.org/0000-0001-8329-7839; Phone: +55 (51) 3308-5354; Email: walter.beys@ufrgs.br

Authors

Rafael Lopes da Rosa – Programa de Pós-Graduação em Biologia Celular e Molecular, Universidade Federal do Rio Grande do Sul, Porto Alegre, Rio Grande do Sul 91501-970, Brasil

Tung Sheng Yang – Faculdade de Farmácia, Universidade Federal do Rio Grande do Sul, Porto Alegre, Rio Grande do Sul 90610-000, Brasil

Emanuela Fernanda Tureta – Programa de Pós-Graduação em Biologia Celular e Molecular, Universidade Federal do Rio Grande do Sul, Porto Alegre, Rio Grande do Sul 91501-970, Brasil; Faculdade de Farmácia, Universidade Federal do Rio Grande do Sul, Porto Alegre, Rio Grande do Sul 90610-000, Brasil

Laura Rascovetzki Saciloto de Oliveira – Faculdade de Farmácia, Universidade Federal do Rio Grande do Sul, Porto Alegre, Rio Grande do Sul 90610-000, Brasil

Amanda Naiara Silva Moraes – Programa de Pós-Graduação em Biologia Celular e Molecular, Universidade Federal do Rio Grande do Sul, Porto Alegre, Rio Grande do Sul 91501-970, Brasil

Juliana Miranda Tatará – Programa de Pós-Graduação em Biologia Celular e Molecular, Universidade Federal do Rio Grande do Sul, Porto Alegre, Rio Grande do Sul 91501-970, Brasil

Renata Pereira Costa – Faculdade de Farmácia, Universidade Federal do Rio Grande do Sul, Porto Alegre, Rio Grande do Sul 90610-000, Brasil

Júlia Spier Borges – Faculdade de Farmácia, Universidade Federal do Rio Grande do Sul, Porto Alegre, Rio Grande do Sul 90610-000, Brasil

Camila Innocente Alves – Faculdade de Farmácia, Universidade Federal do Rio Grande do Sul, Porto Alegre, Rio Grande do Sul 90610-000, Brasil

Markus Berger – Centro de Pesquisa Experimental, Hospital de Clínicas de Porto Alegre, Porto Alegre, Rio Grande do Sul 90035-903, Brasil

Jorge Almeida Guimarães – Centro de Pesquisa Experimental, Hospital de Clínicas de Porto Alegre, Porto Alegre, Rio Grande do Sul 90035-903, Brasil

Complete contact information is available at:
<https://pubs.acs.org/10.1021/acsomega.0c05701>

Author Contributions

R.L.d.R.: methodology, web design, data collection and curation, formal analysis, investigation, and writing—original draft; T.S.Y.: methodology, data collection and curation, formal analysis, investigation, and writing—original draft; E.F.T., L.R.S.d.O., A.N.S.M., J.M.T., R.P.C., J.S.B., and C.I.A.: methodology and data collection and curation; M.B. and J.A.G.: writing—review and editing; and L.S.: investigation and writing—review and editing. W.O.B.-d.-S.: conceptualization, methodology, investigation, supervision, and writing—review and editing.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors would like to thank the Brazilian agency Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) for financial support.

REFERENCES

(1) WHO, Statement on the second meeting of the International Health Regulations (2005) Emergency Committee regarding the outbreak of novel coronavirus (2019-nCoV). Available in: [https://www.who.int/news-room/detail/30-01-2020-statement-on-the-second-meeting-of-the-international-health-regulations-\(2005\)-emergency-committee-regarding-the-outbreak-of-novel-coronavirus-\(2019-ncov\)](https://www.who.int/news-room/detail/30-01-2020-statement-on-the-second-meeting-of-the-international-health-regulations-(2005)-emergency-committee-regarding-the-outbreak-of-novel-coronavirus-(2019-ncov)); World Health Organization, 2020.

(2) Wrapp, D.; Wang, N.; Corbett, K. S.; Goldsmith, J. A.; Hsieh, C.-L.; Abiona, O.; Graham, B. S.; McLellan, J. S. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* **2020**, *367*, 1260–1263.

(3) WHO, WHO Statement regarding cluster of pneumonia cases in Wuhan, China. Available in: <https://www.who.int/china/news/detail/09-01-2020-who-statement-regarding-cluster-of-pneumonia-cases-in-wuhan-china>; World Health Organization, 2020.

(4) JHU CSSE. COVID-19 Dashboard by the Center for Systems Science and Engineering (CSSE). Available in: <https://coronavirus.jhu.edu/map.html>; John Hopkins University, 2020.

(5) CONASS. Conselho Nacional de Secretários de Saúde. Painel COVID-19. Available in: <https://www.conass.org.br/>, 2020.

(6) Horbach, S. P. J. M. Pandemic Publishing: Medical journals drastically speed up their publication process for Covid-19. *Quant. Sci. Studies* **2020**, *1*, 1056–1067.

(7) Bojkova, D.; Klann, K.; Koch, B.; Widera, M.; Krause, D.; Ciesek, S.; Cinatl, J.; Münch, C. Proteomics of SARS-CoV-2-infected host cells reveals therapy targets. *Nature* **2020**, *583*, 469–472.

(8) Ouyang, Y.; Yin, J.; Wang, W.; Shi, H.; Shi, Y.; Xu, B.; Qiao, L.; Feng, Y.; Pang, L.; Wei, F.; Guo, X.; Jin, R.; Chen, D. Downregulated gene expression spectrum and immune responses changed during the disease progression in patients with COVID-19. *Clin. Infect. Dis.* **2020**, *71*, 2052.

(9) Shen, B.; Yi, X.; Sun, Y.; Bi, X.; Du, J.; Zhang, C.; Quan, S.; Zhang, F.; Sun, R.; Qian, L.; et al. Proteomic and metabolomic characterization of COVID-19 patient sera. *Cell* **2020**, *182*, 59–72.

(10) Yang, S.; Fu, C.; Lian, X.; Dong, X.; Zhang, Z. Understanding human-virus protein-protein interactions using a human protein complex-based analysis framework. *mSystems* **2019**, *4*, No. e00303-18.

(11) Przybyla, P.; Shardlow, M.; Aubin, S.; Bossy, R.; Eckart de Castilho, R.; Piperidis, S.; McNaught, J.; Ananiadou, S. Text mining resources for the life sciences. *Database* **2016**, *2016*, baw145.

(12) Rosa, R. L.; Santi, L.; Berger, M.; Tureta, E. F.; Quincozes-Santos, A.; Souza, D. O.; Guimarães, J. A.; Beys-da-Silva, W. O. ZIKAVID—Zika virus infection database: a new platform to analyze the molecular impact of Zika virus infection. *J. NeuroVirol.* **2019**, *26*, 77–83.

(13) Kambhampati, S. B. S.; Vaishya, R.; Vaish, A. Unprecedented surge in publications related to COVID-19 in the first three months of pandemic: A bibliometric analytic report. *J. Clin. Orthop. Trauma* **2020**, *11*, S304–S306.

(14) PAHO, WHO announces a Public Health Emergency of International Concern; Pan American Health Organization, 2016. Available in: https://www.paho.org/hq/index.php?option=com_content&view=article&id=11640:2016-who-statement-on-1st-meeting-ihc-2005-emergency-committee-on-zika-virus&Itemid=135(=en.

(15) Neerukonda, S. N.; Katneni, U. A Review on SARS-CoV-2 virology, pathophysiology, animal models, and anti-viral interventions. *Pathogens* **2020**, *9*, 426.

(16) Esser-Nobis, K.; Aarberg, L. D.; Roby, J. A.; Fairgrieve, M. R.; Green, R.; Gale, M., Jr. Comparative analysis of African and Asian lineage-derived Zika virus strains reveals differences in activation of and sensitivity to antiviral innate immunity. *J. Virol.* **2019**, *93*, No. e00640.

(17) Islam, M. R.; Hoque, M. N.; Rahman, M. S.; Ul Alam, A. S. M. R.; Akther, M.; Puspo, J. A.; Akter, S.; Sultana, M.; Crandall, K. A.; Hossain, M. A. Genome-wide analysis of SARS-CoV-2 virus strains circulating worldwide implicates heterogeneity. *Sci. Rep.* **2020**, *10*, 14004.

(18) Beys-da-Silva, W. O.; Rosa, R. L.; Santi, L.; Berger, M.; Park, S. K.; Campos, A. R.; Terraciano, P.; Varela, A. P. M.; Teixeira, T. F.; Roehle, P. M.; Quincozes-Santos, A.; Yates, J. R., III; Souza, D. O.; Cirne-Lima, E. O.; Guimarães, J. A. Zika virus infection of human mesenchymal stem cells promotes differential expression of proteins linked to several neurological diseases. *Mol. Neurobiol.* **2018**, *56*, 4708–4717.

(19) Pickett, B. E.; Sadat, E. L.; Zhang, Y.; Noronha, J. M.; Squires, R. B.; Hunt, V.; Liu, M.; Kumar, S.; Zaremba, S.; Gu, Z.; Zhou, L.; Larson, C. N.; Dietrich, J.; Klem, E. B.; Scheuermann, R. H. ViPR: an open bioinformatics database and analysis resource for virology research. *Nucleic Acids Res.* **2012**, *40*, D593–D598.

(20) Pyro, V. S.; Oliveira, F. S.; Morais, D. K.; Cuadros-Orellana, S.; Pais, F. S.-M.; Medeiros, J. D.; Geraldo, J. A.; Gilbert, J.; Volpini, A. C.; Fernandes, G. R. ZIKV – CDB: a collaborative database to guide research linking sncRNAs and Zika virus disease symptoms. *PLoS Neglected Trop. Dis.* **2016**, *10*, No. e0004817.

(21) Guirimand, T.; Delmotte, S.; Navratil, V. VirHostNet 2.0: surfing on the web of virus/host molecular interactions data. *Nucleic Acids Res.* **2015**, *43*, D583–D587.

(22) Rozenblatt-Rosen, O.; Deo, R. C.; Padi, M.; Adelmant, G.; Calderwood, M. A.; Rolland, T.; Grace, M.; Dricot, A.; Askenazi, M.; Tavares, M.; et al. Interpreting cancer genomes using systematic host perturbations by tumour virus proteins. *Nature* **2012**, *487*, 491–495.

(23) Ako-Adjei, D.; Fu, W.; Wallin, C.; Katz, K. S.; Song, G.; Darji, D.; Brister, J. R.; Ptak, R. G.; Pruitt, K. D. HIV-1, human interaction database: current status and new features. *Nucleic Acids Res.* **2014**, *43*, D566–D570.

(24) Clough, E.; Barrett, T. The gene expression omnibus database. In *Statistical Genomics, Methods in Molecular Biology*; Mathé, E., Davis, S., Eds.; Springer: New York, 2016; Vol. 1418, pp 93–110.

(25) Zhou, Y.; Zhou, B.; Pache, L.; Chang, M.; Khodabakhshi, A. H.; Tanaseichuk, O.; Benner, C.; Chandra, S. K. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* **2019**, *10*, 1523.

(26) Kuleshov, M. V.; Stein, D. J.; Clarke, D. J. B.; Kropiwnicki, E.; Jagodnik, K. M.; Bartal, A.; Evangelista, J. E.; Hom, J.; Cheng, M.; Bailey, A.; Zhou, A.; Ferguson, L. B.; Lachmann, A.; Ma'ayan, A. The COVID-19 Gene and Drug Set Library. *Patterns* **2020**, *1*, 100090.

(27) Papanikolaou, N.; Pavlopoulos, G. A.; Theodosiou, T.; Iliopoulos, I. Protein–protein interaction predictions using text mining methods. *Methods* **2015**, *74*, 47–53.