

H2V, a database for human genes and proteins in response to SARS-CoV-2, SARS-CoV, and MERS-CoV infection

Nan Zhou

Guangzhou Huiai Hospital <https://orcid.org/0000-0002-0393-316X>

Jinku Bao (✉ baojinku@scu.edu.cn)

Sichuan University

Yuping Ning

Guangzhou Huiai Hospital

Database

Keywords: database, coronavirus, protein, gene

Posted Date: December 1st, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-61338/v2>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published on January 7th, 2021. See the published version at <https://doi.org/10.1186/s12859-020-03935-2>.

Abstract

The ongoing COVID-19 pandemic in the world is caused by SARS-CoV-2, a new coronavirus first discovered in the end of 2019. It has led to more than 50 million confirmed cases and more than 1 million deaths across 219 countries by 11 November 2020, according to the WHO statistics. SARS-CoV-2, SARS-CoV, and MERS-CoV are alike. They are highly pathogenic, and they threaten public health, impair economy, and inflict long-term impacts on the society. No drug or vaccine has been approved as a cure for these viruses. The efforts to develop antiviral measures are hampered by the insufficient understanding of how the human body responds to viral infections at the cellular and molecular levels. In this study, journal articles and transcriptomic and proteomic data that survey coronavirus infections were collected. Response genes and proteins were then identified via differential analyses which compared the gene/protein between the infected sample and control. A database, H2V, was finally created for human genes/proteins responding to SARS-CoV-2, SARS-CoV, and MERS-CoV infection. H2V provides the molecular information about human response to infection. It can be a powerful tool to discover cellular pathways and processes relevant for viral pathogenesis to identify potential drug targets. It is expected to speed up the process of developing antiviral agents and to shed light on the preparation for potential coronavirus emergency in the future.

Background

Coronaviruses are single-stranded RNA viruses, and some can cross the species barrier to cause deadly and infectious respiratory disease in humans [1]. A novel coronavirus that causes viral pneumonia was reported in December 2019 [2]. The virus, now known as SARS-CoV-2, is commonly asymptomatic and contagious prior to symptom onset [3]. These characteristics contribute to the difficulty of containing the virus. As a result, SARS-CoV-2 spread rapidly in the world and caused the ongoing COVID-19 pandemic.

The last two coronavirus disease epidemics were severe acute respiratory syndrome (SARS) in 2002-2003 and Middle East respiratory syndrome (MERS) from 2012 [4]. With a case fatality rate of ~10%, the SARS-related coronavirus (SARS-CoV) infected 8098 people and caused 774 deaths; the MERS-related coronavirus (MERS-CoV) has a higher mortality rate of ~34%, and it has resulted in ~2500 confirmed cases and ~900 deaths so far [5]. The average case fatality rate of COVID-19 is ~2%, though the risk of serious complications and mortality increases dramatically at later ages [6]. The death rate is < 0.1 in children, but it increases to 10% or more in older people [7]. In terms of the absolute number of cases and deaths, COVID-19 is more severe than the previous two outbreaks. As of 11 November 2020, > 50 million confirmed cases and > 1 million deaths have been reported to WHO ([https:// www.who.int](https://www.who.int)) worldwide. It is urgent for the world to unite to find effective ways to bring the COVID-19 crisis to an end.

SARS-CoV-2, SARS-CoV and MERS-CoV are beta-coronaviruses and are able to cause serious health consequences in humans. Two other beta-coronaviruses, HCoV-OC43 and HKU1, are also able to infect human beings but only cause self-limiting flu-like illness [8]. Even though the world has been repeatedly suffered from coronavirus outbreaks, there is no clinically effective prophylactics or therapeutics

available. The clinical management of COVID-19, as well as SARS and MERS, is largely limited to infection prevention and supportive care. This amplifies the need to develop therapies to treat coronavirus diseases.

The life cycle of coronavirus includes several key steps: viral entry, genomic RNA replication, mRNA translation, protein processing, and virion assembly and release [9]. The interplay between host cell and virus at the viral entry stage has been well documented. To enter the human cell, both SARS-CoV-2 and SARS-CoV bind their S proteins to the cell surface receptor ACE2, angiotensin-converting enzyme 2 [10]. MERS-CoV enters the human cell via binding another receptor, dipeptidyl peptidase 4 (DPP4) [4]. Hoffmann and colleagues have also proved that the binding of SARS-CoV-2 S protein to human ACE2 additionally depends on TMPRSS2 and showed that the cell entry of SARS-CoV-2 can be blocked by serine protease inhibitor camostat mesylate [11]. More details about the interplay between human and virus at other life cycle stages remain to be revealed. There is no doubt that the human body responds to virus infection and the response can be detected at the molecular level by genome- and proteome-wide measurements.

Although SARS-CoV suddenly disappeared in the Summer of 2003, MERS-CoV is occasionally saw and SARS-CoV-2 keeps spreading rapidly in some parts of the world. It is getting worse that the 2020 Winter wave of COVID-19 has forced new lockdowns in some European cities. To pull life back to normal track, specific drugs for COVID-19 are urgently required but unavailable yet. Also, there is no cure for SARS and MERS, indicating our understanding of these dangerous coronaviruses is very limited. As long as the knowledge of cellular responses to viral infections is essential to establish therapeutics, we identified human proteins and genes that respond to SARS-CoV-2, SARS-CoV and MERS-CoV infections and then developed the H2V database in the present study.

Construction And Content

Data collection

In the study, response proteins/genes of human to virus infection are defined as differentially expressed genes (DEGs), proteins that participant in human-virus protein-protein interactions (PPIs), differentially expressed proteins (DEPs), differentially phosphorylated proteins (DPPs), differentially translated proteins (DTPs), differentially ubiquitinated proteins (DUPs), and disease severity associated proteins (SAPs).

We used the Bing search engine (<https://www.bing.com>), NCBI resources (<https://www.ncbi.nlm.nih.gov/>), and Proteome Xchange database (<http://www.proteomexchange.org/>) to search studies of SARS-CoV-2, SARS-CoV, and MERS-CoV infection. With respect to the definition of response gene/protein, studies were classified into types of DEG, PPI, DEP, DPP, DTP, DUP and SAP. For each study type, three independent studies per virus were selected. If the number of available studies was less than three, what we could find would be used. Since we focused on the dynamic change of response gene/protein over time post infection, studies with time-course survey were selected with priority. Only if there were insufficient studies, would that without time-course examination be selected. After study

selection, journal articles of the selected studies were collected, and then the information about response genes and proteins were extracted from the main text and supplementary materials of the article. When such information was not available in the journal article, raw data of the selected studies were downloaded from public repositories and subsequently analyzed. The selected studies ([12–24]) and corresponding strategies to identify response genes and proteins were summarized in Table 1.

Genome assemblies MN985325.1, NC_004718.3 and NC_019843.3 from the NCBI database (<https://www.ncbi.nlm.nih.gov/>) were respectively used to annotate SARS-CoV-2, SARS-CoV and MERS-CoV genes. Drug information was collected from the DrugBank database [25]. Post processing of data was performed by R (<https://www.r-project.org/>) and Python (<https://python.org/>).

Implementation

H2V was developed with mainstream web development techniques. The user interface was developed with HTML5, CSS3, and JavaScript. Bootstrap v4 (<https://getbootstrap.com/>) was used for layout design. DataTables (<https://datatables.net/>) was used to organize data in table on the web page. Cytoscape.js was used for network visualization [26]. Plotly (<https://plotly.com/>) was used to create interactive plots. PHP (<https://www.php.net/>), python (https://www.python.org) and bash scripts were used for server-side development. SQLite (<https://www.sqlite.org/>) database was used to manage data. NCBI's sequence viewer (<https://www.ncbi.nlm.nih.gov/projects/sviewer/>) was embedded on the web page to show viral genome. PANTHER API was used for pathway enrichment analysis [27]. Drug information is not stored in H2V, instead it is automatically retrieved on request from the DrugBank database via UniProt's REST API [28]. H2V is deployed in an Amazon AWS host running Ubuntu 16.04.

Utility And Discussion

Statistics of H2V data

Due to the variation in the availability of studies, H2V datasets vary in viruses. As shown in Table 2, genes/proteins that respond to SARS-CoV-2 infection exist in seven datasets, namely DEGs, PPIs, DEPs, DPPs, DTPs, DUPs and SAPs. In comparison, genes/proteins that respond to SARS-CoV and MERS-CoV infections exist in three (DEGs, PPIs and DEPs) and two (DEGs and PPIs) datasets, respectively. There are DEGs datasets for the response to infections of all viruses. There are 9321 human genes responding to MERS-CoV infection, while the number of genes in response to SARS-CoV infection drops to 2249 and then to as low as 1395 for SARS-CoV-2 infection. PPIs datasets are also available for the response to infections of all viruses. There are 1581, 1150, and 296 interaction pairs between human proteins and the respective protein of SARS-CoV-2, SARS-CoV and MERS-CoV. DEPs datasets are available for the response to SARS-CoV-2 and SARS-CoV infections; the number of human proteins in response to infections of the two viruses are 253 and 66, respectively. Datasets DPPs, DTPs, DUPs and SAPs are only available for the response to SARS-CoV-2 infection, and the number of response proteins are 2198 (5046 phosphorylation sites), 232, 516 (730 ubiquitination sites) and 610.

To know whether there are common proteins participating in different processes in response to SARS-CoV-2 infection, the intersection of DEPs, DPPs, DTPs and DUPs were analyzed. Figure 1a shows that both expression and translation of 11 proteins change dramatically upon infection, that both phosphorylation and ubiquitination of 180 proteins change remarkably upon infection, and that one protein experiences noticeable change in expression, phosphorylation, translation and ubiquitination. We then used Venn diagrams to analyze genes/proteins that are common in response to different virus infections. This would help to elucidate the fundamental mechanisms of viral pathogenesis. Figure 1b shows that 130 common genes encounter significant difference in expression upon infection. Figure 1c shows that 62 human proteins are able to interact with all of the three viruses.

Overview of H2V

As shown in Figure 2a, there is a navigation bar and a search box in the header on the web page. The search box accepts queries from the user and then tries to match anything that looks like a gene or protein. The navigation bar provides access to all resources in the database. The SARS2 drop-down menu links to response genes/proteins to SARS-CoV-2 infection. Similarly, the SARS1 and MERS drop-down menus link to response genes/proteins to SARS-CoV-1 and MERS-CoV infections, respectively. Under the Utilities drop-down menu, useful utilities, such as the link to download data from or upload data to H2V, are provided. On the page that lists response genes/proteins, the genes/proteins are shown in a table in rows, with additional information about the gene/protein shown in columns (Figure 2b). The Score column in the table indicates the reliability of the gene/protein, and the score was calculated as the number of studies in which the gene/protein was identified [29]. The genes/proteins in the table are clickable. Clicking on a gene/protein will link to another page showing details of how the gene/protein responds to virus infection. This page comes with two fantastic features: one is to examine changes of the gene/protein at different timepoints post infection (Figure 2c), and the other is to discover known drugs that target the gene/protein. For PPIs, an embedded sequence viewer, as shown in Figure 2d, is provided for easy inspection of the gene/protein annotation in the viral genome. In addition, PPIs can also be visualized as an interaction network on the page (Figure 2e).

Application cases

To facilitate rapid drug discovery to treat COVID-19 at the pandemic time, H2V provides a drug finder which can be used to find drugs of a given protein based on the UniProt accession number. The found drugs, with DrugBank identifiers, will then be displayed on the lower part of the same page. For example, searching Q9BYF1 will find a few drugs, including Chloroquine and Hydroxychloroquine (Figure 3a).

To help users establish a concrete perception of how all genes/proteins change dynamically over time post infection, H2V provides a utility named Data animation for this purpose. On the page, a setting panel is provided to select data for animation. For example, Figure 3b shows the setting to animate DPPs in response to SARS-CoV-2 infection. The results (Figure 3c and 3d) of this example demonstrate that more human proteins are differentially phosphorylated at 24 h than at the very beginning after the infection of

SARS-CoV-2. This indicates that the human body responds to SARS-CoV-2 infection by continuously rewiring cellular pathways.

H2V can be used to analyze integrated findings from different studies. Figure 4 shows an example of using the Enrichment analysis utility to analyze enriched pathways of DPPs that respond to SARS-CoV-2 infection. DPPs identified in at least two studies were analyzed at first (also referred to as analysis 1). After setting parameters on the left in Figure 4a, the analysis was implemented by clicking the button at the bottom. When the analysis was completed, the input DPPs for analysis were listed on the right in Figure 4a, and the result was shown in Figure 4b. It shows that seven pathways were enriched, including FAS signaling pathway, p38 MAPK pathway, and PDGF signaling pathway. It is supposed that findings repeated by independent studies would be more reliable than the unrepeatable ones, so the same analysis (also referred to as analysis 2) was performed for DPPs identified in at least one study. This time, more pathways were enriched, and the top seven pathways were shown in Figure 4c. The comparison shows that the top two pathways in analysis 1 were not in the top seven pathways of analysis 2. This indicates that the inclusion of DPPs of low confidence could distort the analysis result. H2V can be used to remove confounding factors to acquire reliable biological inferences.

Conclusions

We have developed H2V, the first database for human proteins and genes that respond to SARS-CoV-2, SARS-CoV, and MERS-CoV infection. The database will help to understand the cellular details of how the human body responds to coronavirus infections. H2V can also be used as a platform to analyze rewired pathways by combining findings of independent studies. This can be helpful to find key targets with the potential to treat coronavirus diseases. We have to acknowledge that the present release of our database may omit some data which should be included, we will keep updating the database and offer missing data in future releases. In summary, the database will help to design effective and specific drugs and preventive vaccines targeting SARS-CoV-2, SARS-CoV and MERS-CoV.

Abbreviations

SARS-CoV-2: Severe acute respiratory syndrome coronavirus 2

SARS-CoV: Severe acute respiratory syndrome coronavirus

MERS-CoV: Middle East respiratory syndrome-related coronavirus

COVID-19: Coronavirus disease 2019

WHO: World Health Organization

ACE2: Angiotensin-converting enzyme 2

DPP4: Dipeptidyl peptidase 4

DEGs: Differentially expressed genes

PPIs: Protein-protein interactions

DEPs: Differentially expressed proteins

DPPs: Differentially phosphorylated proteins

DTPs: Differentially translated proteins

SAPs: Severity associated proteins

NCBI: National center for biotechnology information

HTML: Hypertext markup language

CSS: Cascading style sheets

REST: Representational state transfer

API: Application programming interface

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

All data generated or analyzed during this study are included in this published article.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by: Basic and Applied Basic Research of Guangzhou Municipal Basic Research Plan, Guangzhou Municipal Psychiatric Disease Clinical Transformation Laboratory [201805010009]; Guangzhou Municipal Key Discipline in Medicine (2017-2019); Key Laboratory for Innovation Platform Plan, Science and Technology Program of Guangzhou, China.

Authors' contributions

N.Z. and J.B. collected and analyzed data. N.Z. developed the website and wrote manuscript draft. J.B. and Y.N. conceived and supervised the study. All author took roles in reviewing and revising the manuscript.

Acknowledgements

We thank the authors who generated the raw data that have been used in this study.

References

1. Weiss SR, Navas-Martin S. Coronavirus pathogenesis and the emerging pathogen severe acute respiratory syndrome coronavirus. *Microbiol Mol Biol Rev.* 2005;69:635–64. doi:10.1128/MMBR.69.4.635-664.2005.
2. Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N Engl J Med.* 2020;382:727–33. doi:10.1056/NEJMoa2001017.
3. Bai Y, Yao L, Wei T, Tian F, Jin D-Y, Chen L, et al. Presumed Asymptomatic Carrier Transmission of COVID-19. *JAMA.* 2020;323:1406–7. doi:10.1001/jama.2020.2565.
4. Zhou N, Zhang Y, Zhang J-C, Feng L, Bao J-K. The receptor binding domain of MERS-CoV: The dawn of vaccine and treatment development. *J Formos Med Assoc.* 2014;113:143–7. doi:https://doi.org/10.1016/j.jfma.2013.11.006.
5. Walls AC, Park Y-J, Tortorici MA, Wall A, McGuire AT, Velesler D. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell.* 2020;181:281-292.e6. doi:https://doi.org/10.1016/j.cell.2020.02.058.
6. Promislow DEL. A Geroscience Perspective on COVID-19 Mortality. *Journals Gerontol Ser A.* 2020;75:e30–3. doi:10.1093/gerona/glaa094.
7. Verity R, Okell LC, Dorigatti I, Winskill P, Whittaker C, Imai N, et al. Estimates of the severity of coronavirus disease 2019: a model-based analysis. *Lancet Infect Dis.* 2020;20:669–77. doi:10.1016/S1473-3099(20)30243-7.
8. Liu DX, Liang JQ, Fung TS. Human Coronavirus-229E, -OC43, -NL63, and -HKU1. *Ref Modul Life Sci.* 2020;B978-0-12-809633-8.21501-X. doi:10.1016/B978-0-12-809633-8.21501-X.
9. Dai L, Zheng T, Xu K, Han Y, Xu L, Huang E, et al. A Universal Design of Betacoronavirus Vaccines against COVID-19, MERS, and SARS. *Cell.* 2020. doi:https://doi.org/10.1016/j.cell.2020.06.035.
10. Zhou P, Yang X-L, Wang X-G, Hu B, Zhang L, Zhang W, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature.* 2020;579:270–3. doi:10.1038/s41586-020-2012-7.
11. Hoffmann M, Kleine-Weber H, Schroeder S, Krüger N, Herrler T, Erichsen S, et al. SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell.* 2020;181:271-280.e8. doi:https://doi.org/10.1016/j.cell.2020.02.052.

12. Stukalov A, Girault V, Grass V, Bergant V, Karayel O, Urban C, et al. Multi-level proteomics reveals host-perturbation strategies of SARS-CoV-2 and SARS-CoV. *bioRxiv*. 2020;:2020.06.17.156455. doi:10.1101/2020.06.17.156455.
13. Lamers MM, Beumer J, van der Vaart J, Knoops K, Puschhof J, Breugem TI, et al. SARS-CoV-2 productively infects human gut enterocytes. *Science* (80-). 2020;369:50 LP – 54. doi:10.1126/science.abc1669.
14. Yoshikawa T, Hill TE, Yoshikawa N, Popov VL, Galindo CL, Garner HR, et al. Dynamic Innate Immune Responses of Human Bronchial Epithelial Cells to Severe Acute Respiratory Syndrome-Associated Coronavirus Infection. *PLoS One*. 2010;5:e8729. <https://doi.org/10.1371/journal.pone.0008729>.
15. Jiang X-S, Tang L-Y, Dai J, Zhou H, Li S-J, Xia Q-C, et al. Quantitative Analysis of Severe Acute Respiratory Syndrome (SARS)-associated Coronavirus-infected Cells Using Proteomic Approaches. *Mol & Cell Proteomics*. 2005;4:902 LP – 913. doi:10.1074/mcp.M400112-MCP200.
16. Zhang X, Chu H, Wen L, Shuai H, Yang D, Wang Y, et al. Competing endogenous RNA network profiling reveals novel host dependency factors required for MERS-CoV propagation. *Emerg Microbes Infect*. 2020;9:733–46. doi:10.1080/22221751.2020.1738277.
17. Gordon DE, Jang GM, Bouhaddou M, Xu J, Obernier K, White KM, et al. A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature*. 2020. doi:10.1038/s41586-020-2286-9.
18. Gordon DE, Hiatt J, Bouhaddou M, Rezelj V V, Ulferts S, Braberg H, et al. Comparative host-coronavirus protein interaction networks reveal pan-viral disease mechanisms. *Science* (80-). 2020;:eabe9403. doi:10.1126/science.abe9403.
19. Bojkova D, Klann K, Koch B, Widera M, Krause D, Ciesek S, et al. Proteomics of SARS-CoV-2-infected host cells reveals therapy targets. *Nature*. 2020. doi:10.1038/s41586-020-2332-7.
20. Bouhaddou M, Memon D, Meyer B, White KM, Rezelj V V, Marrero MC, et al. The Global Phosphorylation Landscape of SARS-CoV-2 Infection. *Cell*. 2020. doi:<https://doi.org/10.1016/j.cell.2020.06.034>.
21. Klann K, Bojkova D, Tascher G, Ciesek S, Münch C, Cinatl J. Growth Factor Receptor Signaling Inhibition Prevents SARS-CoV-2 Replication. *Mol Cell*. 2020;80:164-174.e4. doi:10.1016/j.molcel.2020.08.006.
22. Shen B, Yi X, Sun Y, Bi X, Du J, Zhang C, et al. Proteomic and Metabolomic Characterization of COVID-19 Patient Sera. *Cell*. 2020. doi:<https://doi.org/10.1016/j.cell.2020.05.032>.
23. Li Y, Wang Y, Liu H, Sun W, Ding B, Zhao Y, et al. Urine Proteome of COVID-19 Patients. *medRxiv*. 2020;:2020.05.02.20088666. doi:10.1101/2020.05.02.20088666.
24. Mitchell HD, Eisfeld AJ, Sims AC, McDermott JE, Matzke MM, Webb-Robertson B-JM, et al. A Network Integration Approach to Predict Conserved Regulators Related to Pathogenicity of Influenza and SARS-CoV Respiratory Viruses. *PLoS One*. 2013;8:e69374. <https://doi.org/10.1371/journal.pone.0069374>.
25. Wishart DS, Feunang YD, Guo AC, Lo EJ, Marcu A, Grant JR, et al. DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res*. 2017;46:D1074–82. doi:10.1093/nar/gkx1037.

26. Franz M, Lopes CT, Huck G, Dong Y, Sumer O, Bader GD. Cytoscape.js: a graph theory library for visualisation and analysis. *Bioinformatics*. 2015;:btv557.
27. Mi H, Muruganujan A, Ebert D, Huang X, Thomas PD. PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. *Nucleic Acids Res*. 2019;47:D419–26. doi:10.1093/nar/gky1038.
28. Consortium TU. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res*. 2018;47:D506–15. doi:10.1093/nar/gky1049.
29. Zhou N, Bao J. FerrDb: a manually curated resource for regulators and markers of ferroptosis and ferroptosis-disease associations. *Database (Oxford)*. 2020;2020. doi:10.1093/database/baaa021.

Tables

Table 1. Studies and strategies to identify response genes/proteins.

Study¹	Type	Strategy
SARS-CoV-2		
SRA: SRP257667	DEG	b
doi: 10.1101/2020.06.17.156455	DEG	a
PMID: 32358202	DEG	c
PMID: 32353859	PPI	a
doi: 10.1101/2020.06.17.156455	PPI	a
PMID: 33060197	PPI	a
PMID: 32408336	DEP	a
doi: 10.1101/2020.06.17.156455	DEP	a
PMID: 32645325	DPP	a
doi: 10.1101/2020.06.17.156455	DPP	a
PMID: 32877642	DPP	a
PMID: 32408336	DTP	a
doi: 10.1101/2020.06.17.156455	DUP	a
PMID: 32492406	SAP	a
doi: 10.1101/2020.05.02.20088666	SAP	a
SARS-CoV		
PMID: 32358202	DEG	c
PMID: 23935999	DEG	d
PMID: 20090954	DEG	d
PMID: 33060197	PPI	a
doi: 10.1101/2020.06.17.156455	PPI	a
PMID: 15784933	DEP	a
MERS-CoV		
PMID: 32223537	DEG	a
GEO: GSE81909	DEG	d
GEO: GSE79458	DEG	d
PMID: 33060197	PPI	a

¹ If PMID is not available, alternative database accession is used.

a: Response genes/proteins were extracted from the journal article.

b: Response genes/proteins were identified from RNA-seq data using RaNA-seq, with $p < 0.05$ and $|\log_2(\text{fold change})| > 1$ at any timepoint post infection.

c: Response genes/proteins were identified from read counts from GEO using DESeq2, with $p < 0.05$ and $|\log_2(\text{fold change})| > 1$ at any timepoint post infection.

d: Response genes/proteins were identified from expression matrix from GEO using limma, with $p < 0.05$ and $|\log_2(\text{fold change})| > 1$ at any timepoint post infection.

Table2. Statistics of data in H2V.

Virus	Dataset	Number of entries
SARS-CoV-2	DEGs	1395
	PPIs	1581
	DEPs	253
	DPPs	5046
	DTPs	232
	DUPs	730
	SAPs	610
SARS-CoV	DEGs	2249
	PPIs	1150
	DEPs	66
MERS-CoV	DEGs	9321
	PPIs	296

Figures

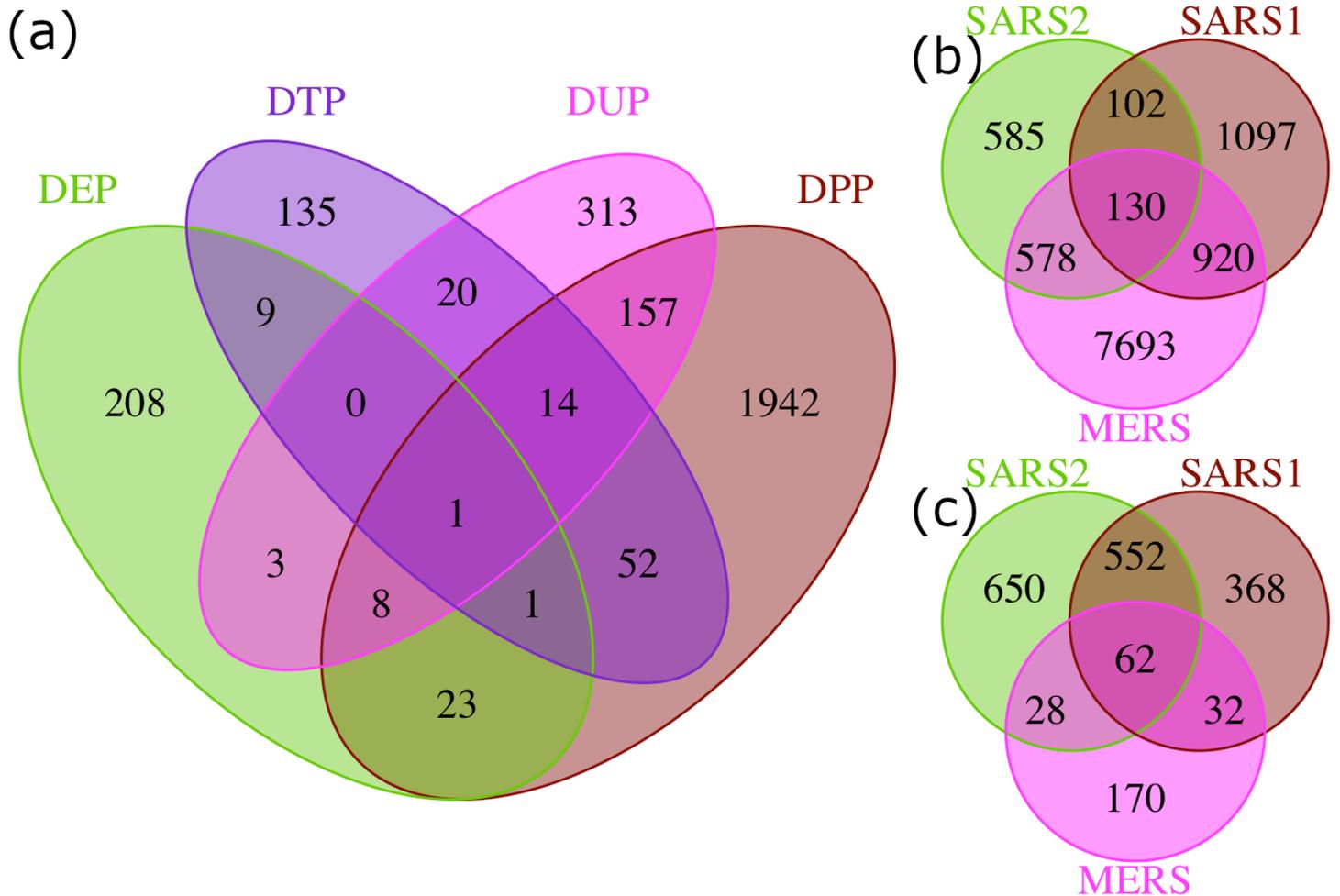


Figure 1

Venn diagrams showing intersections of response genes/proteins. (a) The intersection of DEPs, DPPs, DTPs, and DUPs in response to SARS-CoV-2 infection. (b) The intersection of DEGs in response to SARS-CoV-2, SARS-CoV, and MERS-CoV infection. (c) The intersection of PPIs in response to SARS-CoV-2, SARS-CoV, and MERS-CoV infection. SARS2 denotes SARS-CoV-2; SARS1 denotes SARS-CoV; MERS denotes MERS-CoV.

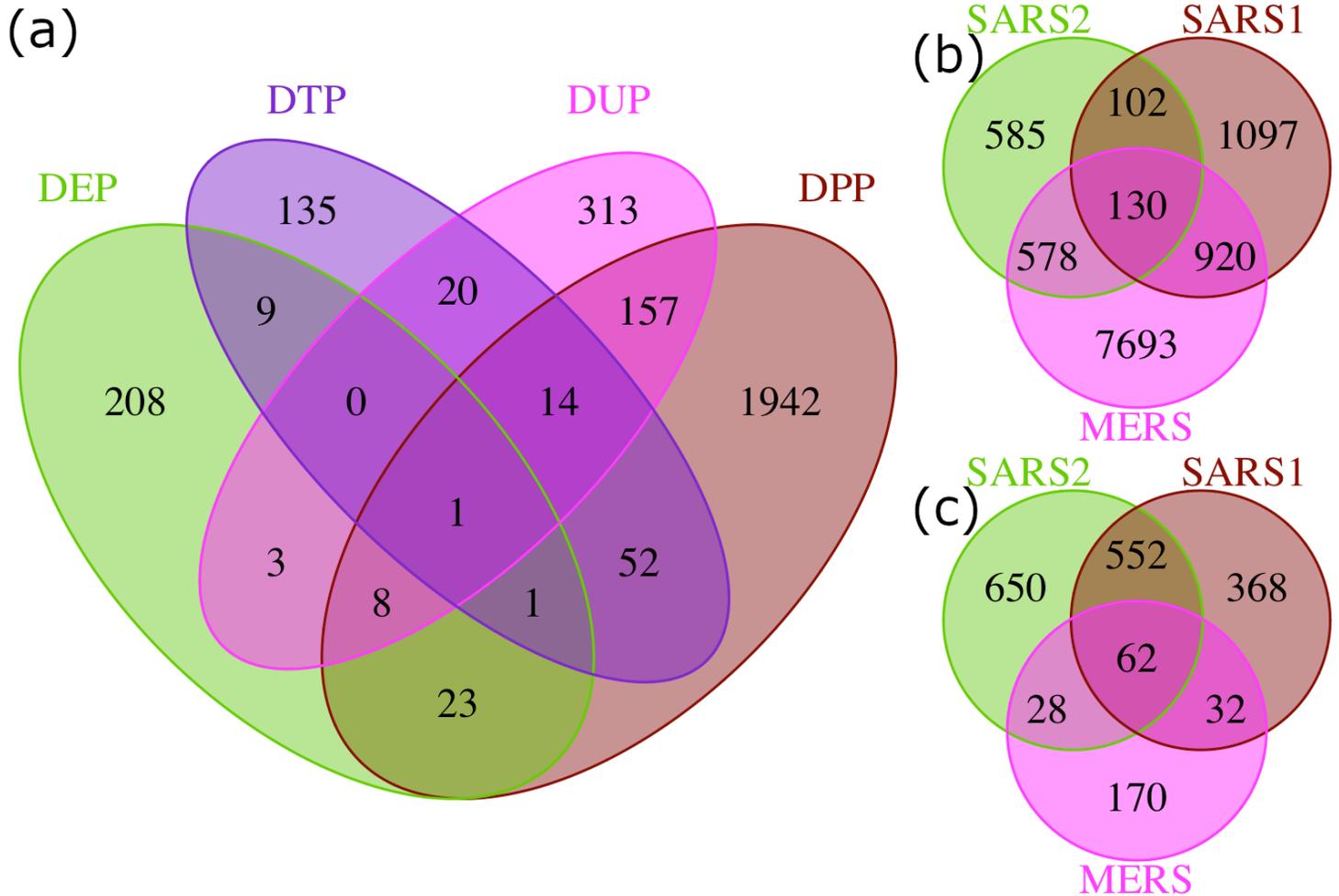


Figure 1

Venn diagrams showing intersections of response genes/proteins. (a) The intersection of DEPs, DPPs, DTPs, and DUPs in response to SARS-CoV-2 infection. (b) The intersection of DEGs in response to SARS-CoV-2, SARS-CoV, and MERS-CoV infection. (c) The intersection of PPIs in response to SARS-CoV-2, SARS-CoV, and MERS-CoV infection. SARS2 denotes SARS-CoV-2; SARS1 denotes SARS-CoV; MERS denotes MERS-CoV.

(b)

Gene symbol	HGNC ID	Protein AC	Protein ID	Score
CYP1A1	HGNC:2595	P04798	CP1A1_HUMAN	2
FST	HGNC:3971	P19883	FST_HUMAN	2
LINGO1	HGNC:21205	Q96FE5	LIGO1_HUMAN	2

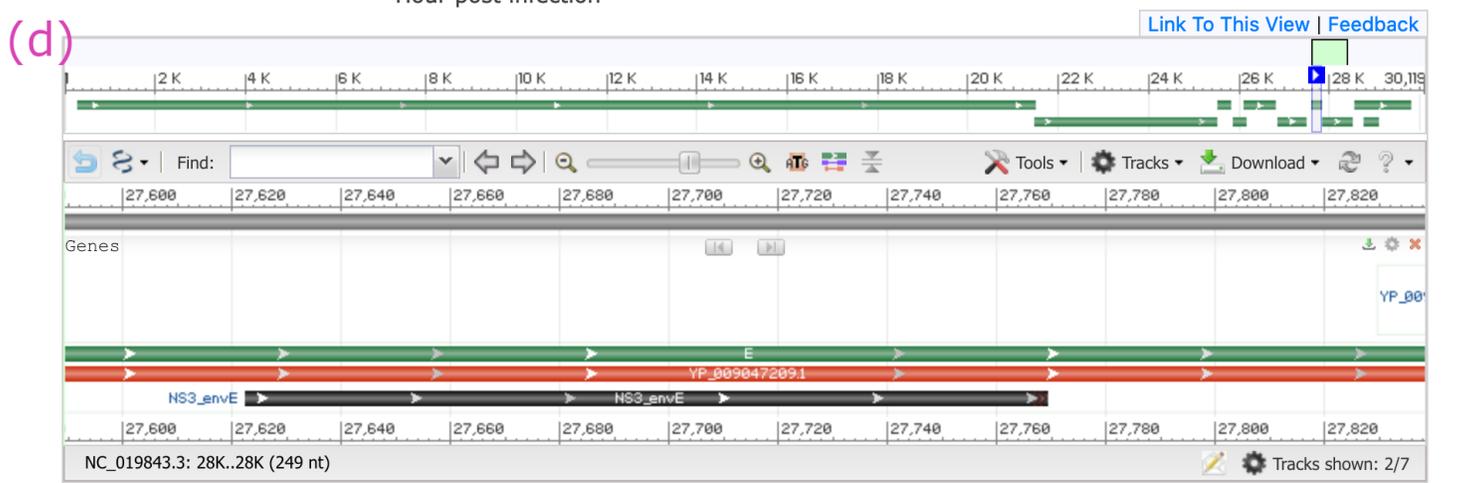
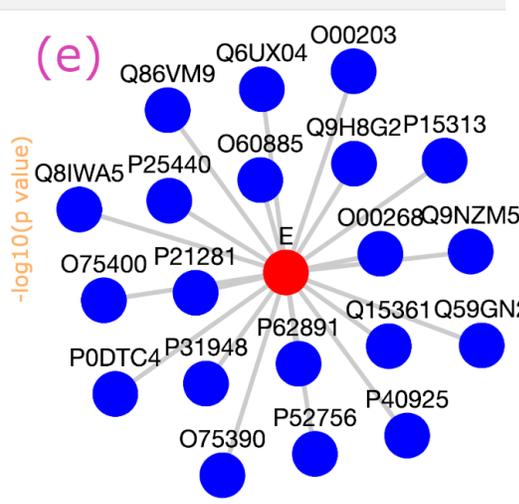


Figure 2

H2V overview. (a) Header. (b) An example data table. (c) An example temporal profile. The x axis shows timepoints, the left y axis shows the log₂-transformed fold change, and the right y axis shows the negation of log₁₀-transformed p value. (d) The embedded NCBI sequence viewer. (e) An example visualization of a subset of PPIs. Red node denotes viral protein; blue nodes denote human proteins.

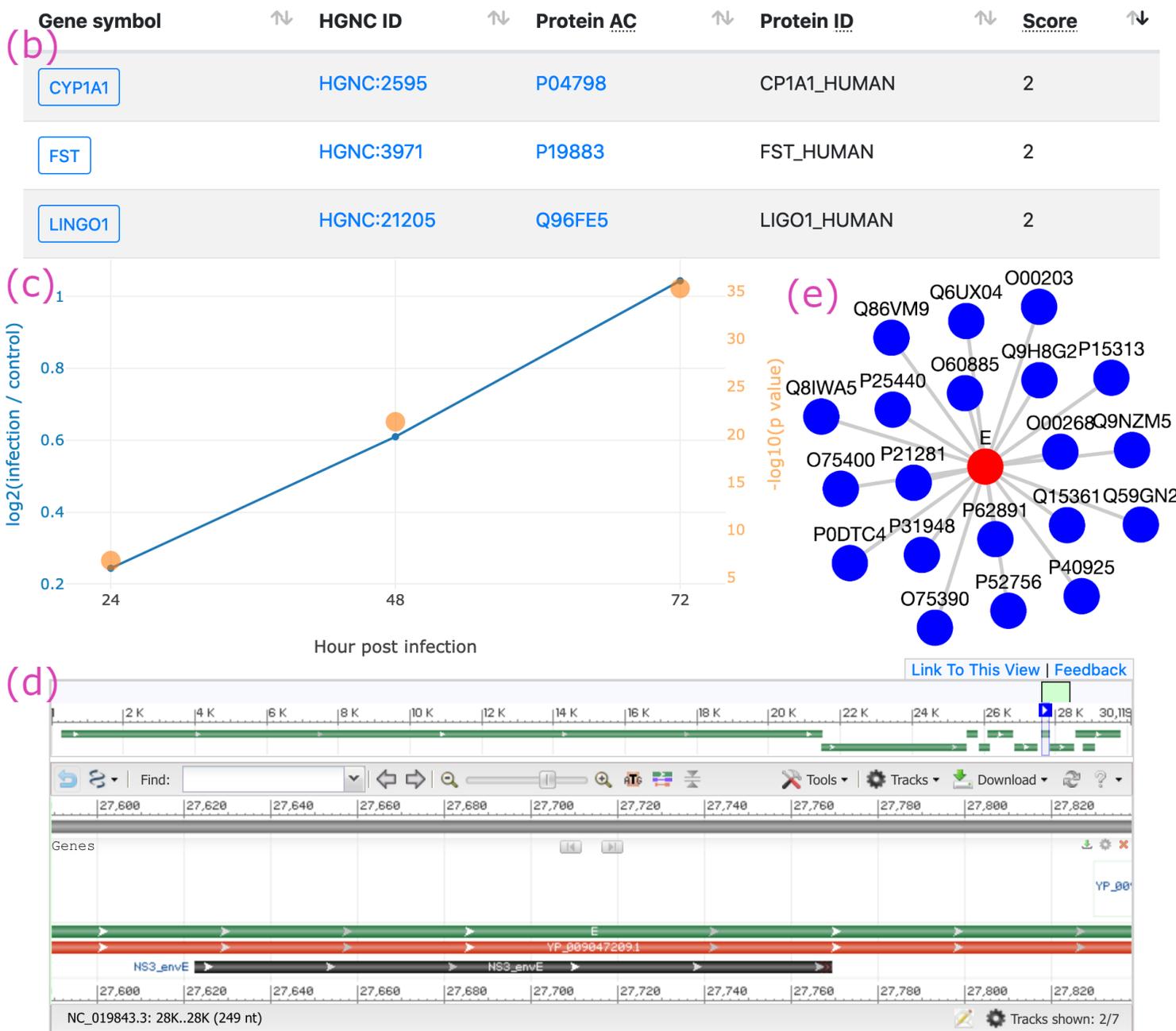


Figure 2

H2V overview. (a) Header. (b) An example data table. (c) An example temporal profile. The x axis shows timepoints, the left y axis shows the log₂-transformed fold change, and the right y axis shows the negation of log₁₀-transformed p value. (d) The embedded NCBI sequence viewer. (e) An example visualization of a subset of PPIs. Red node denotes viral protein; blue nodes denote human proteins.

Uniprot Accession Find Reset

Finish loading data. Drugs will be shown as following. If you do not see anything, no drug exists.

DB00608: Chloroquine

(a)

DB01611: Hydroxychloroquine

DB15643: N-(2-Aminoethyl)-1-aziridineethanamine

DB05203: SPP1148

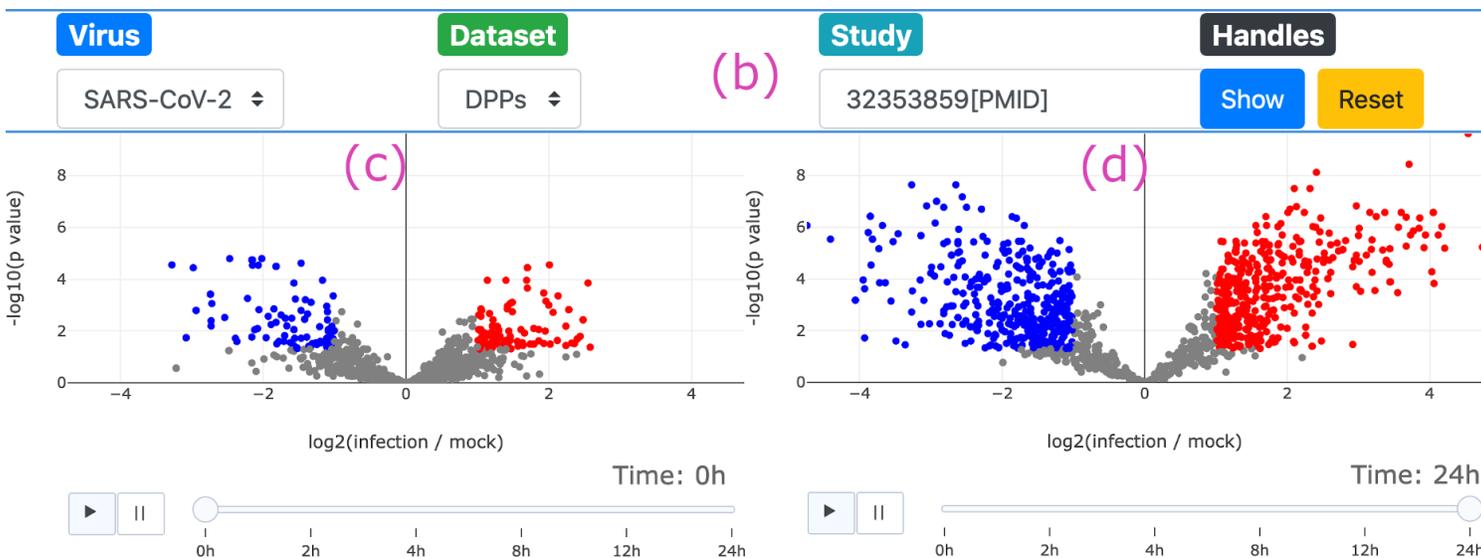


Figure 3

Utilities and examples. (a) Drug finder. (b) Setting panel for data animation. The phosphorylation profile of all proteins at the beginning (c) and at the end (d). Of note, the animation can be played on the web page. For (c) and (d), the x axis shows the \log_2 -transformed fold change, the y axis shows the negation of \log_{10} -transformed p value, blue points denote proteins with down-regulated phosphorylation upon infection, red points denote proteins with up-regulated phosphorylation upon infection, and gray points denote proteins without significant change in phosphorylation upon infection.

Uniprot Accession Find Reset

Finish loading data. Drugs will be shown as following. If you do not see anything, no drug exists.

DB00608: Chloroquine

(a)

DB01611: Hydroxychloroquine

DB15643: N-(2-Aminoethyl)-1-aziridineethanamine

DB05203: SPP1148

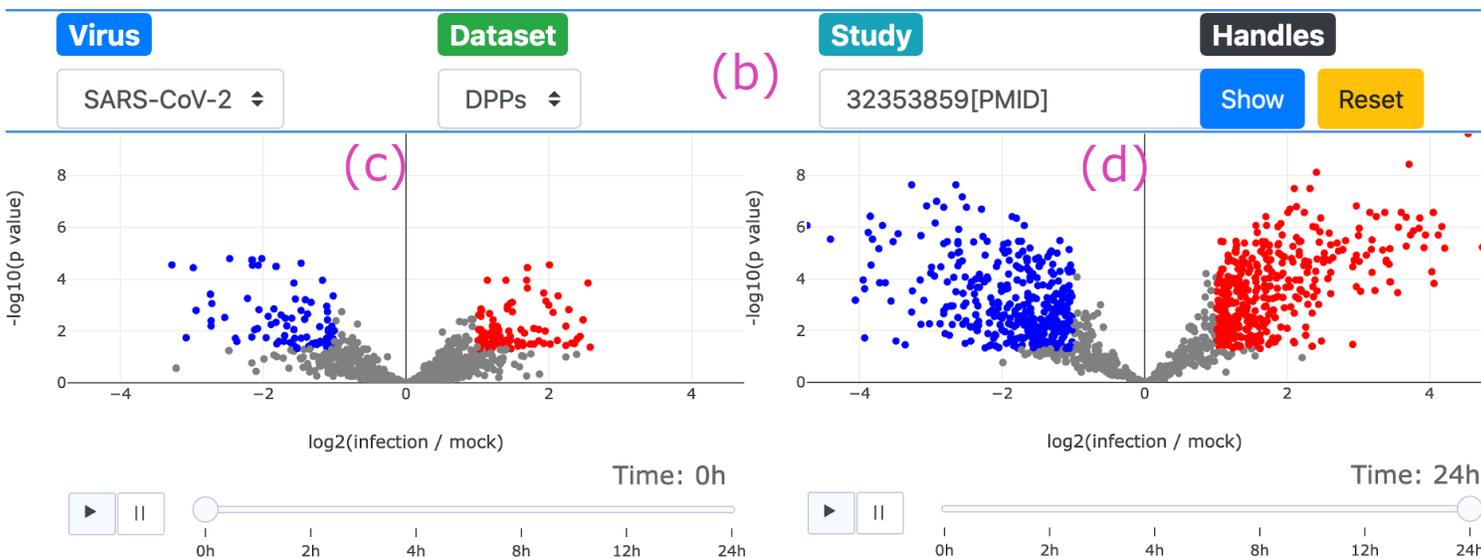


Figure 3

Utilities and examples. (a) Drug finder. (b) Setting panel for data animation. The phosphorylation profile of all proteins at the beginning (c) and at the end (d). Of note, the animation can be played on the web page. For (c) and (d), the x axis shows the log₂-transformed fold change, the y axis shows the negation of log₁₀-transformed p value, blue points denote proteins with down-regulated phosphorylation upon infection, red points denote proteins with up-regulated phosphorylation upon infection, and gray points denote proteins without significant change in phosphorylation upon infection.

(a)

Input

Dataset: SARS-CoV-2 DPPs

Score:

Only genes/proteins identified in ≥ 2 studies will be used.

[Get data and analyze](#)

Analyzed gene/protein list:

- Q9UKV3
- Q09666
- O14617
- Q92974
- Q96BY7
- Q9UIG0
- O95999

(b)

Output

Pathway	PANTHER ID	p value	FDR
FAS signaling pathway	P00020	0.0004473673735765739	0.07336824926655812
p38 MAPK pathway	P05918	0.0007123317871758992	0.05841120654842373
PDGF signaling pathway	P00047	0.012775382075844386	0.6983875534794931
Glycolysis	P00024	0.0158230210906928	0.6487438647184048
CCKR signaling map	P06959	0.02426906638385598	0.7960253773904762
Ras Pathway	P04393	0.03431415447675284	0.9379202223645775
Cytoskeletal regulation by Rho GTPase	P00016	0.045242086466919726	1

(c)

Pathway	PANTHER ID	p value	FDR
PDGF signaling pathway	P00047	8.70480223097153e-8	0.000014275875658793309
CCKR signaling map	P06959	8.599121307698015e-7	0.00007051279472312373
Ras Pathway	P04393	0.0000035424959400913377	0.00019365644472499314
EGF receptor signaling pathway	P00018	0.000027298789826515326	0.0011192503828871283
Gonadotropin-releasing hormone receptor pathway	P06664	0.00007739185912905976	0.00253845297943316
Insulin/IGF pathway-mitogen activated protein kinase kinase/MAP kinase cascade	P00032	0.0001398118091893848	0.0038215227845098507
B cell activation	P00010	0.00035366396039385704	0.008285841357798937

Figure 4

Enrichment analysis. (a) Input panel. (b) Pathway enrichment of DPPs identified in at least two studies. (c) Pathway enrichment of DPPs identified in at least one study (only top seven pathways are shown).

(a)

Input

Dataset: SARS-CoV-2 DPPs

Score:

Only genes/proteins identified in ≥ 2 studies will be used.

[Get data and analyze](#)

Analyzed gene/protein list:

- Q9UKV3
- Q09666
- O14617
- Q92974
- Q96BY7
- Q9UIG0
- O95999

(b)

Output

Pathway	PANTHER ID	p value	FDR
FAS signaling pathway	P00020	0.0004473673735765739	0.07336824926655812
p38 MAPK pathway	P05918	0.0007123317871758992	0.05841120654842373
PDGF signaling pathway	P00047	0.012775382075844386	0.6983875534794931
Glycolysis	P00024	0.0158230210906928	0.6487438647184048
CCKR signaling map	P06959	0.02426906638385598	0.7960253773904762
Ras Pathway	P04393	0.03431415447675284	0.9379202223645775
Cytoskeletal regulation by Rho GTPase	P00016	0.045242086466919726	1

(c)

Pathway	PANTHER ID	p value	FDR
PDGF signaling pathway	P00047	8.70480223097153e-8	0.000014275875658793309
CCKR signaling map	P06959	8.599121307698015e-7	0.00007051279472312373
Ras Pathway	P04393	0.0000035424959400913377	0.00019365644472499314
EGF receptor signaling pathway	P00018	0.000027298789826515326	0.0011192503828871283
Gonadotropin-releasing hormone receptor pathway	P06664	0.00007739185912905976	0.00253845297943316
Insulin/IGF pathway-mitogen activated protein kinase kinase/MAP kinase cascade	P00032	0.0001398118091893848	0.0038215227845098507
B cell activation	P00010	0.00035366396039385704	0.008285841357798937

Figure 4

Enrichment analysis. (a) Input panel. (b) Pathway enrichment of DPPs identified in at least two studies. (c) Pathway enrichment of DPPs identified in at least one study (only top seven pathways are shown).