

Oberseminar

Ralf Zimmer
Institute for Informatics, LMU München

HuRI, COVID-Simulations, COVID proteomics



Apr 16, 2020



May 12, 2020

Article


A reference map of the human binary protein interactome

<https://doi.org/10.1038/s41586-020-2188-x>

Received: 19 April 2019

Accepted: 14 February 2020

Published online: 8 April 2020

 Check for updates

Katja Luck^{1,2,3,33}, Dae-Kyum Kim^{1,4,5,6,33}, Luke Lambourne^{1,2,3,33}, Kerstin Spirohn^{1,2,3,33},
Bridget E. Begg^{1,2,3}, Wenting Bian^{1,2,3}, Ruth Brignall^{1,2,3}, Tiziana Cafarelli^{1,2,3},
Francisco J. Campos-Laborie^{7,8}, Benoit Charlotoux^{1,2,3}, Dongsic Choi⁹, Atina G. Côté^{1,4,5,6},
Meaghan Daley^{1,2,3}, Steven Deimling¹⁰, Alice Desbuleux^{1,2,3,11}, Amélie Dricot^{1,2,3},
Marinella Gebbia^{1,4,5,6}, Madeleine F. Hardy^{1,2,3}, Nishka Kishore^{1,4,5,6}, Jennifer J. Knapp^{1,4,5,6},
István A. Kovács^{1,12,13}, Irma Lemmens^{14,15}, Miles W. Mee^{4,5,16}, Joseph C. Mellor^{1,4,5,6,17},
Carl Pollis^{1,2,3}, Carles Pons¹⁸, Aaron D. Richardson^{1,2,3}, Sadie Schlabach^{1,2,3}, Bridget Teeking^{1,2,3},
Anupama Yadav^{1,2,3}, Mariana Babor^{1,4,5,6}, Dawit Balcha^{1,2,3}, Omer Basha^{19,20},
Christian Bowman-Colin^{2,3}, Suet-Feung Chin²¹, Soon Gang Choi^{1,2,3}, Claudia Colabella^{22,23},
Georges Coppin^{1,2,3,11}, Cassandra D'Amata¹⁰, David De Ridder^{1,2,3}, Steffi De Rouck^{14,15},
Miquel Duran-Frigola¹⁸, Hanane Ennajaoui^{1,4,5,6}, Florian Goebels^{4,5,16}, Liana Goehring^{2,3},
Anjali Gopal^{1,4,5,6}, Ghazal Haddad^{1,4,5,6}, Elodie Hatchi^{2,3}, Mohamed Helmy^{4,5,16}, Yves Jacob^{24,25},
Yoseph Kassa^{1,2,3}, Serena Landini^{2,3}, Roujia Li^{1,4,5,6}, Natascha van Lieshout^{1,4,5,6},
Andrew MacWilliams^{1,2,3}, Dylan Markey^{1,2,3}, Joseph N. Paulson^{26,27,28},
Sudharshan Rangarajan^{1,2,3}, John Rasla^{1,2,3}, Ashyad Rayhan^{1,4,5,6}, Thomas Rolland^{1,2,3},
Adriana San-Miguel^{1,2,3}, Yun Shen^{1,2,3}, Dayag Sheykhkarimli^{1,4,5,6}, Gloria M. Sheynkman^{1,2,3},
Eyal Simonovsky^{19,20}, Murat Taşan^{1,4,5,6,16}, Alexander Tejada^{1,2,3}, Vincent Tropepe¹⁰,
Jean-Claude Twizere¹¹, Yang Wang^{1,2,3}, Robert J. Weatheritt⁴, Jochen Weile^{1,4,5,6,16}, Yu Xia^{1,29},
Xinping Yang^{1,2,3}, Esti Yeger-Lotem^{19,20}, Quan Zhong^{1,2,3,30}, Patrick Aloy^{18,31}, Gary D. Bader^{4,5,16},
Javier De Las Rivas^{7,8}, Suzanne Gaudet^{1,2,3}, Tong Hao^{1,2,3}, Janusz Rak⁹, Jan Tavernier^{14,15},
David E. Hill^{1,2,3}, Marc Vidal^{1,2}, Frederick P. Roth^{1,4,5,6,16,32} & Michael A. Calderwood^{1,2,3}

Global insights into cellular organization and genome function require comprehensive understanding of the interactome networks that mediate genotype–phenotype relationships^{1,2}. Here we present a human ‘all-by-all’ reference interactome map of human binary protein interactions, or ‘HuRI’. With approximately 53,000 protein–protein interactions, HuRI has approximately four times as many such interactions as there are high-quality curated interactions from small-scale studies. The integration of HuRI with genome³, transcriptome⁴ and proteome⁵ data enables cellular function to be studied within most physiological or pathological cellular contexts. We demonstrate the utility of HuRI in identifying the specific subcellular roles of protein–protein interactions. Inferred tissue-specific networks reveal general principles for the formation of cellular context-specific functions and elucidate potential molecular mechanisms that might underlie tissue-specific phenotypes of Mendelian diseases. HuRI is a systematic proteome-wide reference that links genomic variation to phenotypic outcomes.

Background

1. Vidal, M., Cusick, M. E. & Barabási, A.-L. Interactome networks and human disease. *Cell* **144**, 986–998 (2011).
2. Rolland, T. et al. A proteome-scale map of the human interactome network. *Cell* **159**, 1212–1226 (2014).
3. Amberger, J. S., Bocchini, C. A., Schiettecatte, F., Scott, A. F. & Hamosh, A. OMIM.org: Online Mendelian Inheritance in Man (OMIM®), an online catalog of human genes and genetic disorders. *Nucleic Acids Res.* **43**, D789–D798 (2015).
4. Melé, M. et al. The human transcriptome across tissues and individuals. *Science* **348**, 660–665 (2015).
5. Thul, P. J. et al. A subcellular map of the human proteome. *Science* **356**, eaal3321 (2017).
6. Lek, M. et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536**, 285–291 (2016).
7. The FANTOM Consortium and the RIKEN PMI and CLST (DGT). A promoter-level mammalian expression atlas. *Nature* **507**, 462–470 (2014).
8. Regev, A. et al. The human cell atlas. *eLife* **6**, e27041 (2017).
9. Lander, E. S. et al. Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921 (2001).
10. Venter, J. C. et al. The sequence of the human genome. *Science* **291**, 1304–1351 (2001).
11. Wan, C. et al. Panorama of ancient metazoan macromolecular complexes. *Nature* **525**, 339–344 (2015).

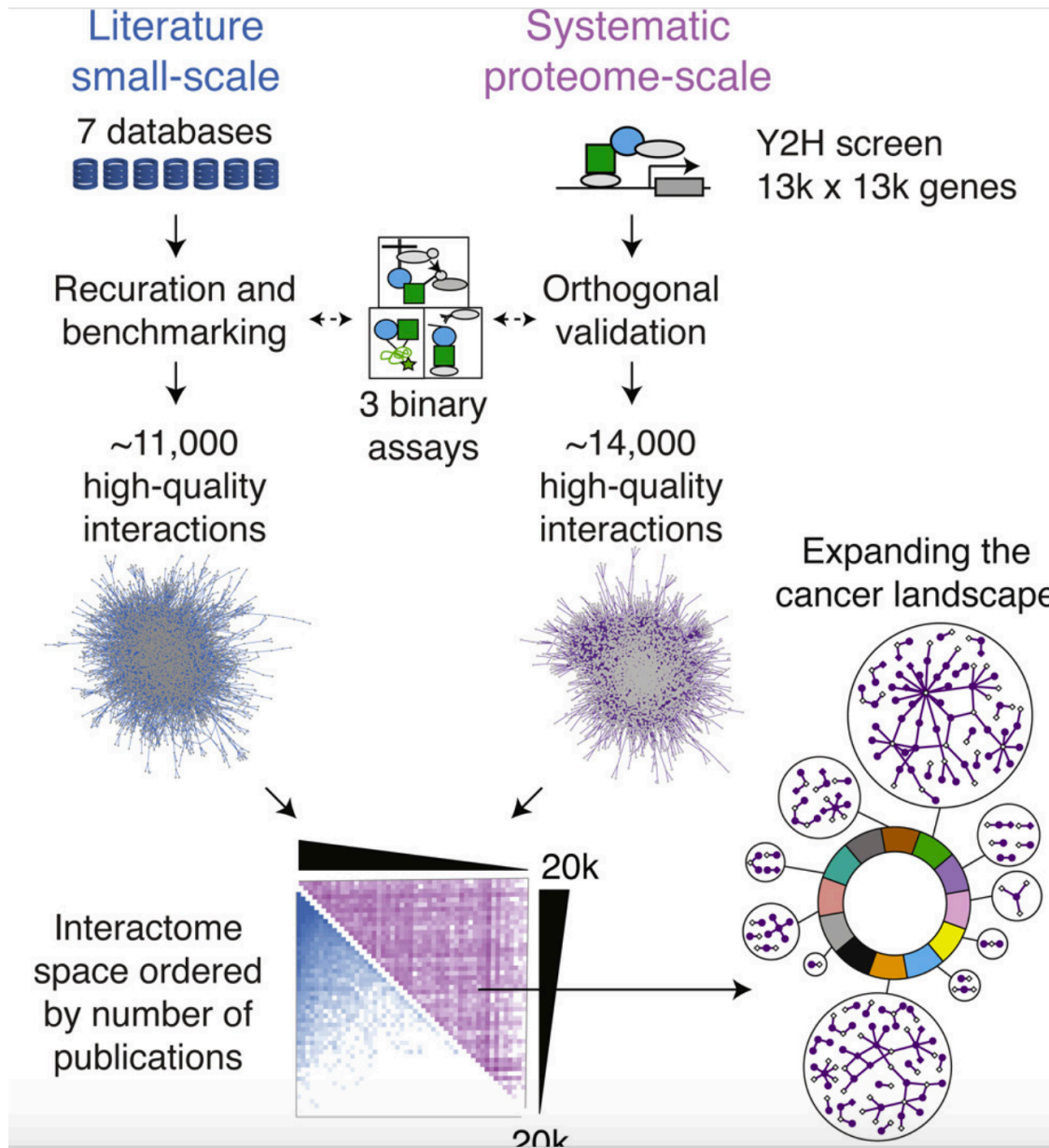


A Proteome-Scale Map of the Human Interactome Network

Thomas Rolland,^{1,2,19} Murat Taşan,^{1,3,4,5,19} Benoit Charloteaux,^{1,2,19} Samuel J. Pevzner,^{1,2,6,7,19} Quan Zhong,^{1,2,8,19} Nidhi Sahni,^{1,2,19} Song Yi,^{1,2,19} Irma Lemmens,⁹ Celia Fontanillo,¹⁰ Roberto Mosca,¹¹ Atanas Kamburov,^{1,2} Susan D. Ghiassian,^{1,12} Xinping Yang,^{1,2} Lila Ghamsari,^{1,2} Dawit Balcha,^{1,2} Bridget E. Begg,^{1,2} Pascal Braun,^{1,2} Marc Brehme,^{1,2} Martin P. Broly,^{1,2} Anne-Ruxandra Carvunis,^{1,2} Dan Convery-Zupan,^{1,2} Roser Corominas,¹³ Jasmin Coulombe-Huntington,^{1,14} Elizabeth Dann,^{1,2} Matija Dreze,^{1,2} Amélie Dricot,^{1,2} Changyu Fan,^{1,2} Eric Franzosa,^{1,14} Fana Gebreab,^{1,2} Bryan J. Gutierrez,^{1,2} Madeleine F. Hardy,^{1,2} Mike Jin,^{1,2} Shuli Kang,¹³ Ruth Kiros,^{1,2} Guan Ning Lin,¹³ Katja Luck,^{1,2} Andrew MacWilliams,^{1,2} Jörg Menche,^{1,12} Ryan R. Murray,^{1,2} Alexandre Palagi,^{1,2} Matthew M. Poulin,^{1,2} Xavier Rambout,^{1,2,15} John Rasla,^{1,2} Patrick Reichert,^{1,2} Viviana Romero,^{1,2} Elien Ruysinck,⁹ Julie M. Sahalie,^{1,2} Annemarie Scholz,^{1,2} Akash A. Shah,^{1,2} Amitabh Sharma,^{1,12} Yun Shen,^{1,2} Kerstin Spirohn,^{1,2} Stanley Tam,^{1,2} Alexander O. Tejada,^{1,2} Shelly A. Trigg,^{1,2} Jean-Claude Twizere,^{1,2,15} Kerwin Vega,^{1,2} Jennifer Walsh,^{1,2} Michael E. Cusick,^{1,2} Yu Xia,^{1,14} Albert-László Barabási,^{1,12,16} Lilia M. Iakoucheva,¹³ Patrick Aloy,^{11,17} Javier De Las Rivas,¹⁰ Jan Tavernier,⁹ Michael A. Calderwood,^{1,2,20} David E. Hill,^{1,2,20} Tong Hao,^{1,2,20} Frederick P. Roth,^{1,3,4,5,18,*} and Marc Vidal^{1,2,*}

Rolland T, Taşan M, Charloteaux B, et al. A proteome-scale map of the human interactome network. *Cell*. 2014;159(5):1212-1226.
doi:10.1016/j.cell.2014.10.050

Lit-BM (2014)



Summary

Just as reference genome sequences revolutionized human genetics, reference maps of interactome networks will be critical to fully understand genotype-phenotype relationships. Here, we describe a systematic map of ~14,000 high-quality human binary protein-protein interactions. At equal quality, this map is ~30% larger than what is available from small-scale studies published in the literature in the last few decades. While currently available information is highly biased and only covers a relatively small portion of the proteome, our systematic map appears strikingly more homogeneous, revealing a “broader” human interactome network than currently appreciated. The map also uncovers significant interconnectivity between known and candidate cancer gene products, providing unbiased evidence for an expanded functional cancer landscape, while demonstrating how high-quality interactome models will help “connect the dots” of the genomic revolution.

Highlights

- Network map of 14,000 high-quality binary human protein-protein interactions
- Uniform coverage of the interactome space overlooked by other approaches
- This map allows detection of functional relationships beyond previous knowledge
- Unbiased network-based evidence for an expanded cancer landscape

HuRI (2020)

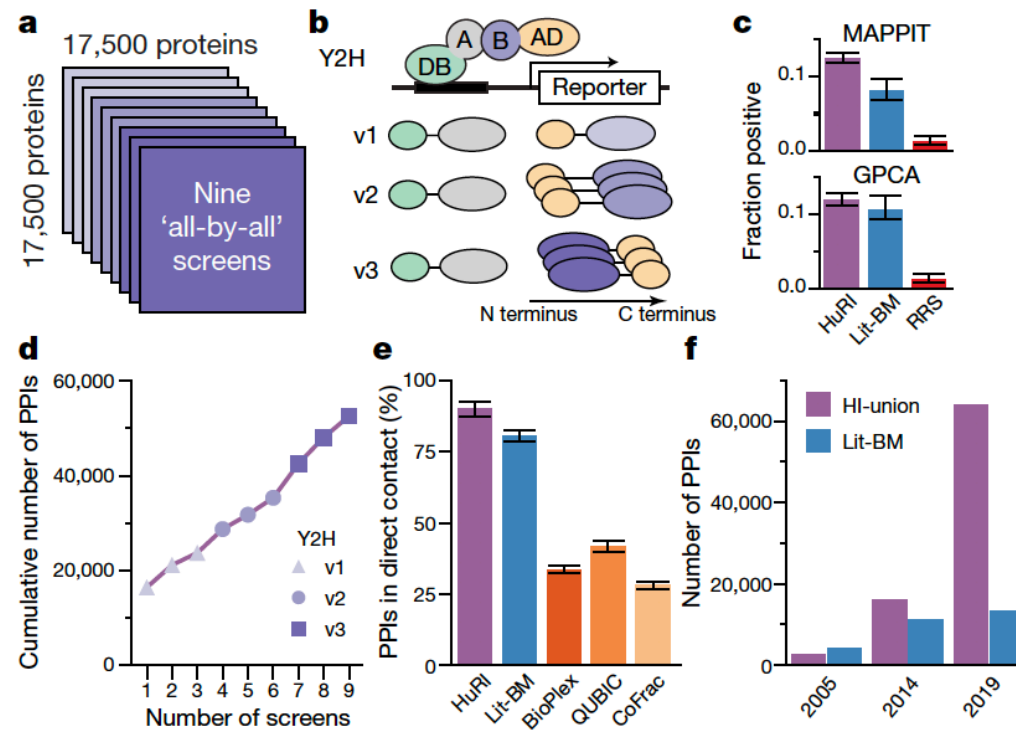


Fig. 1 | Generation of a reference interactome map using a panel of binary assays. **a**, Overview of HuRI generation. **b**, Schematic of the Y2H assay versions. **c**, Experimental validation. Lit-BM, literature-curated binary PPIs with multiple evidence; RRS, random protein pairs. Error bars are 68.3% Bayesian confidence interval. MAPPIT: $n = 2,281, 383$ and 475 ; GPCA: $n = 1,639, 382$ and 465 (from left to right). **d**, Number of PPIs detected with each additional screen. **e**, Fraction of direct contact pairs among five PPI networks. Error bar is standard error of proportion. $n = 121, 410, 1,169, 584$ and $1,211$ PPIs (from left to right). **f**, Number of PPIs identified over time from screening at the Center for Cancer Systems Biology (CCSB) and Lit-BM.

To map the reference interactome, we performed nine screens of Space III, followed by pairwise verification by quadruplicate retesting and sequence confirmation. PPIs verified by two orthogonal binary PPI assays, MAPPIT¹⁹ and GPCA²⁰, were recovered at rates on par with high-confidence binary PPIs from the literature (each having at least two pieces of experimental evidence, with at least one from a binary assay type; Lit-BM) over a large range of score thresholds (Fig. 1c, Extended Data Fig. 1g, h, Supplementary Table 8). Each additional screen identified new PPIs and proteins, with the largest gains obtained by switching assay versions (Fig. 1d, Extended Data Fig. 1i). The dataset, versioned HI-III-20 (Human Interactome obtained from screening Space III, published in 2020), contains 52,569 verified PPIs involving 8,275 proteins (Supplementary Table 9). Although our knowledge of the interactome remains incomplete, we refer to HI-III-20 as a reference map of the human binary protein interactome (HuRI) given its systematic nature, extensive coverage, and scale.

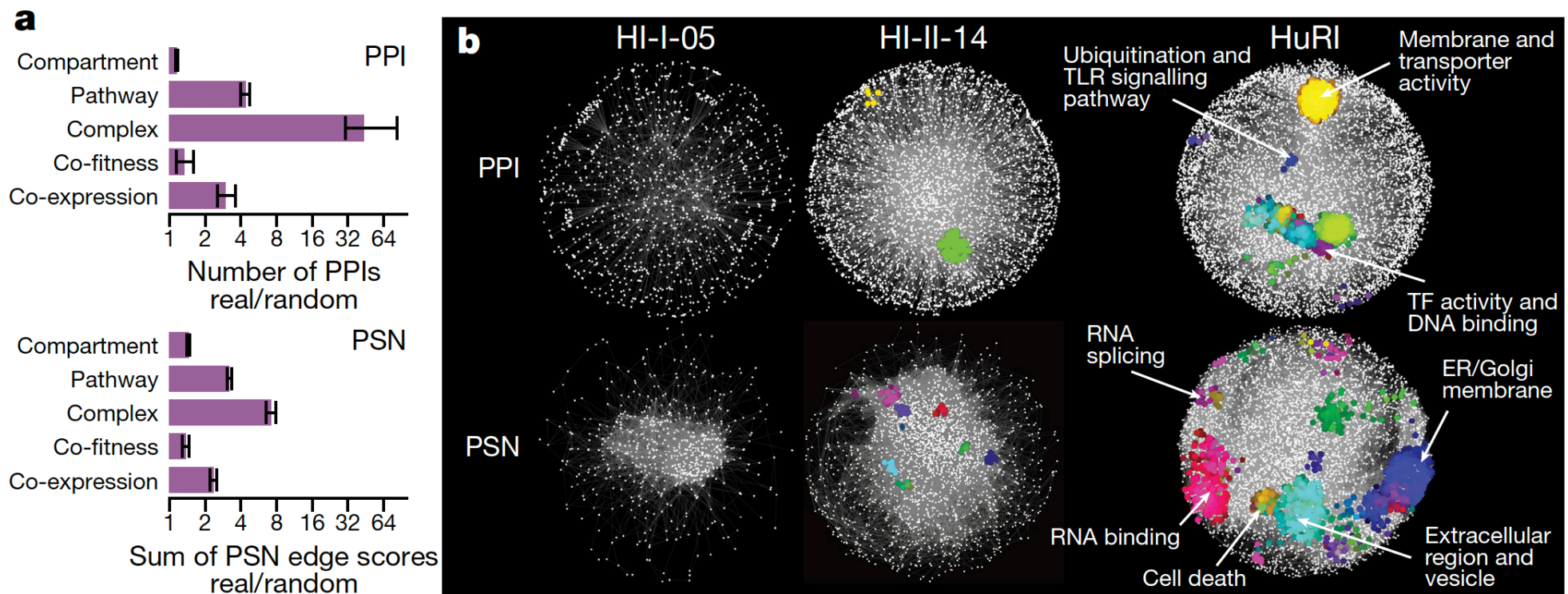


Fig. 2 | Complementary functional relationships in HuRI between genes.
a, Enrichment of HuRI and its profile similarity network (PSN) for protein pairs with shared functional annotation, showing mean and 95% interval of 100 random networks. **b**, Functional modules in HuRI and its PSN and in previously published interactome maps from CCSB.

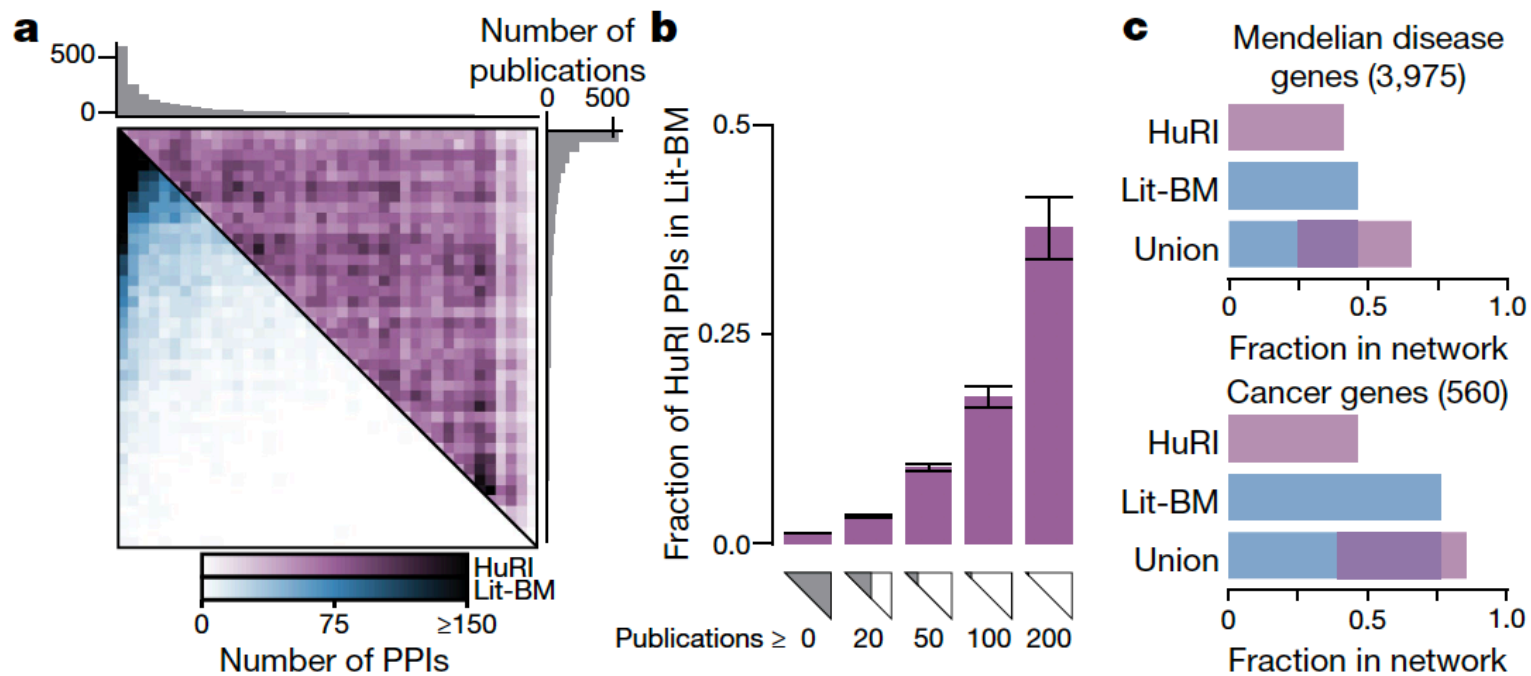
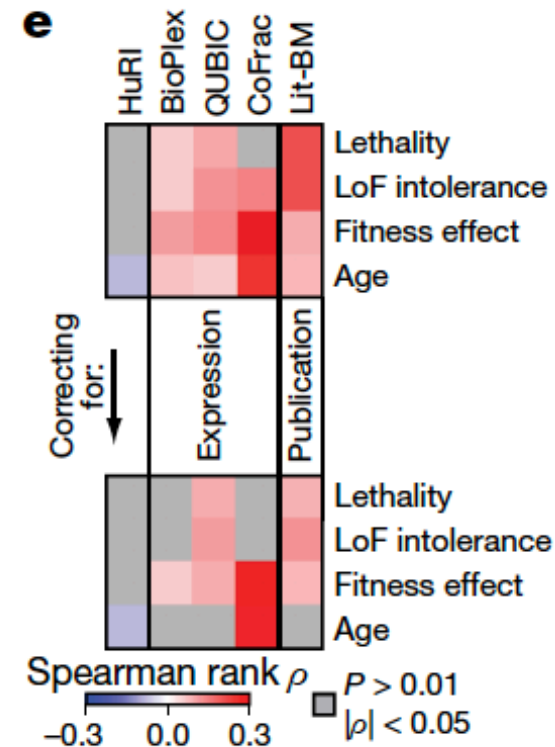
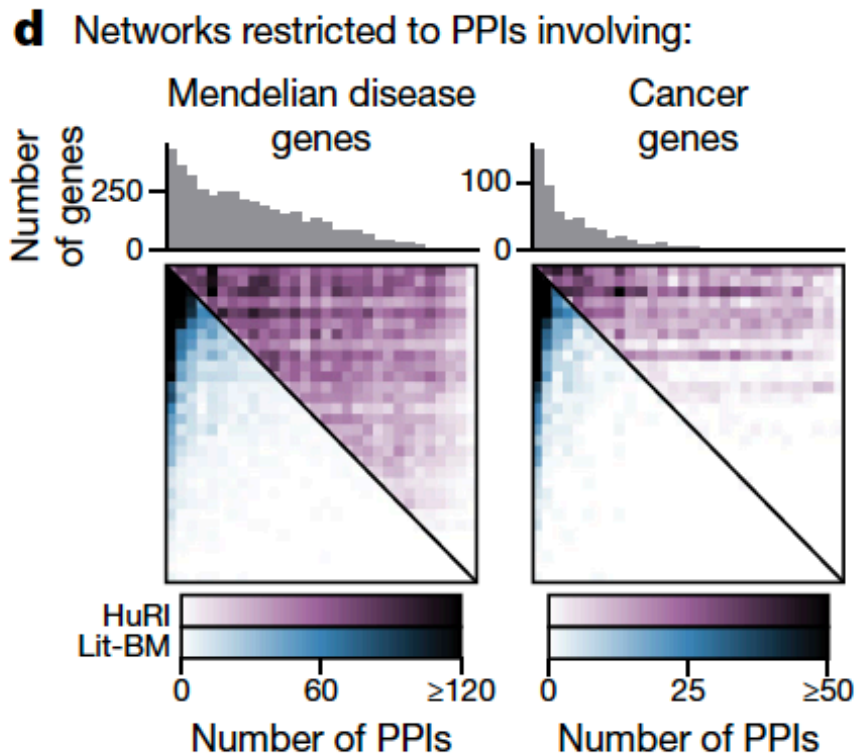


Fig. 3 | Unbiased proteome coverage of HuRI reveals uncharted network neighbourhoods of disease-related genes. a, Heat maps of PPI counts, proteins ordered by number of publications. **b**, Fraction of HuRI PPIs in Lit-BM, for increasing values of the minimum number of publications per protein. Error bar is standard error of proportion, $n = 52,569$ to 170 PPIs (from left to right). **c**, Fraction of genes with at least one PPI for biomedically interesting genes. **d**, As



in **a**, but restricted to PPIs involving genes from the indicated gene sets. **e**, Correlation between degree and variables of interest, before (top) and after (bottom) correcting for the technical confounding factors. $n = 13,441-53,704$ PPIs per network, P value determined by two-tailed permutation test. LoF, loss of function.



- Henriques, R. *et al. Nat. Methods* **7**, 339–340 (2010).
- Wolter, S. *et al. Nat. Methods* **9**, 1040–1041 (2012).
- Baddeley, D., Cannell, M.B. & Soeller, C. *Microsc. Microanal.* **16**, 64–72 (2010).
- Gehlenborg, N. & Wong, B. *Nat. Methods* **9**, 851 (2012).

mentha: a resource for browsing integrated protein-interaction networks

To the Editor: Systems-level approaches require access to comprehensive genome-wide and proteome-wide databases. A comprehensive resource that archives all published protein-protein interactions (PPIs) is not available. In fact, primary PPI databases capture only a fraction of published data.

This dispersion of information has motivated projects such as the Agile Protein Interaction DataAnalyzer (APID), the Protein Interaction Network Analysis (PINA) platform, iRefWeb, Michigan Molecular Interactions (MiMI) and the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING), which offer wider coverage of PPI information by integrating heterogeneously curated data. The difficulty of combining annotations from heterogeneous efforts, however, consistently hampers the integration of data extracted from databases that adopt different curation policies; one consequence of laborious integration procedures is that updates are infrequent.

reliability score can be used to filter the PPI network of interest from

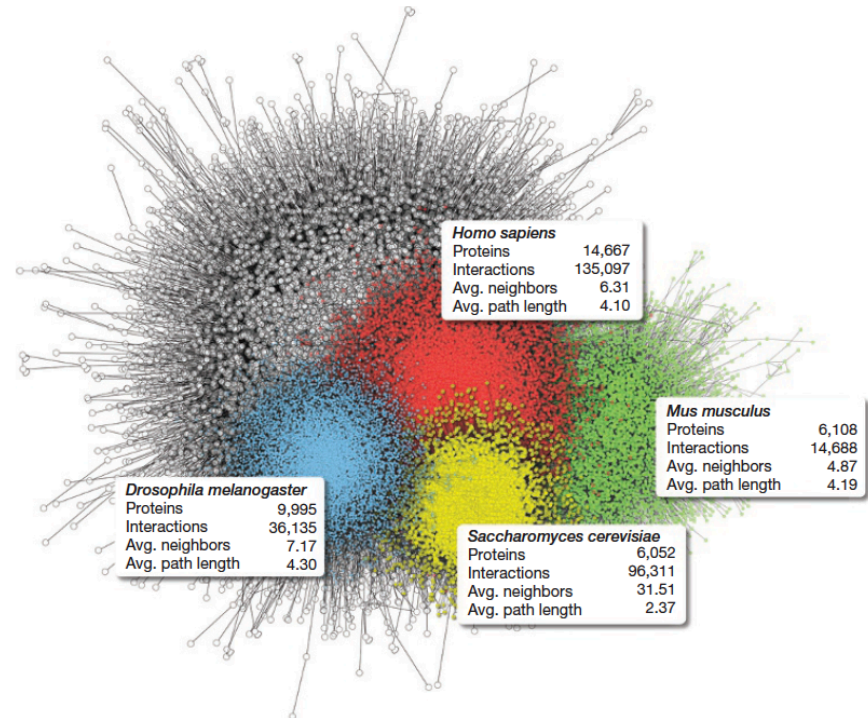


Figure 1 | mentha's interactomes. The gray graph illustrates mentha's "All" interactome. The colored graphs report the interactomes of *Homo sapiens* and three model organisms. The insets report the number of proteins, interactions and some topological characteristics. mentha offers graph analysis tools to extract subnetworks and paths, optionally identifying enzymatic interactions.



Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-020-2188-x>.

HuRI Networks - Luck et al. (2020)

Description:

This set of networks corresponds to all human reference interactome (HuRI)-related networks generated and analyzed in [Luck et al. \(2020\)](#).

The network labelled HuRI corresponds to the full HuRI network.

































The network labelled HI_union corresponds to the union of all PPIs identified in systematic screens at the Center for Cancer Systems Biology (CCSB) and includes HuRI.

The network named CORUM_HI_BP corresponds to the protein complexes integrated with HuRI and BioPlex.

The network labelled EV_network corresponds to the network shown in Fig. 4b.

All networks with a tissue name not preceded by CG correspond to inferred tissue PPI networks from HuRI around tissue-preferentially expressed genes/proteins (Fig. 5c).

Networks in set HuRI Networks - Luck et al. (2020)

✓	Network Name		Ref.	Disease	Tissue	Nodes	Edges
	<input type="text"/>			<input type="text"/>	<input type="text"/>		
✓	esophagus_muscularis					64	62
✓	CG_whole_blood					36	36
✓	heart_atrial_appendage					418	452
✓	heart_left_ventricle					735	836
✓	kidney_cortex					713	859
✓	liver					988	1296
✓	lung					175	162
✓	minor_salivary_gland					329	351
✓	muscle_skeletal					1260	1728
✓	nerve_tibial					123	117
✓	ovary					125	112
✓	pancreas					782	935
✓	CG_brain_basal_ganglia					659	888
✓	pituitary					1125	1512
✓	prostate					153	139
✓	HuRI					8275	52569

Total Items: 49

STRING



Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P, Jensen LJ, von Mering C.

STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets.

Nucleic Acids Res. 2019 Jan; 47:D607-613.

Welcome to STRING

Protein-Protein Interaction Networks

Functional Enrichment Analysis

ORGANISMS

5090

PROTEINS

24.6 mio

INTERACTIONS

>2000 mio

SEARCH

STRING

[Search](#)[Download](#)[Help](#)[My Data](#)

DOWNLOAD

FILES TOO LARGE?

Some of the files below can be made smaller (prior to download), by restricting the data to one organism of interest. Choose an organism here:

[UPDATE](#)














INTERACTION DATA

File	Description	Access
protein.links.v11.0.txt.gz (43.3 Gb)	protein network data (scored links between proteins)	
protein.links.detailed.v11.0.txt.gz (67.6 Gb)	protein network data (incl. subscores per channel)	
protein.links.full.v11.0.txt.gz (72.5 Gb)	protein network data (incl. distinction: direct vs. interologs)	
protein.actions.v11.0.txt.gz (12.9 Gb)	interaction types for protein links	
COG.links.v11.0.txt.gz (288.3 Mb)	association scores between orthologous groups	
COG.links.detailed.v11.0.txt.gz (366.7 Mb)	association scores (incl. subscores per channel)	








STRING

ACCESSORY DATA

File	Description	Access
protein.info.v11.0.txt.gz (571.6 Mb)	list of STRING proteins incl. their display names and descriptions	
protein.sequences.v11.0.fa.gz (5.5 Gb)	sequences of the proteins in STRING (can be used as a blast db)	
COG.mappings.v11.0.txt.gz (381.4 Mb)	orthologous groups (COGs,NOGs,KOGs,...) and their proteins	
species.mappings.v11.0.txt.gz (45.1 Mb)	presence / absence of orthologous groups in species	
species.v11.0.txt (352.9 Kb)	organisms in STRING	
species.tree.v11.0.txt (111.6 Kb)	STRING tree of species	
protein.aliases.v11.0.txt.gz (1.7 Gb)	aliases for STRING proteins: locus names, accessions, descriptions...	
protein.homology.v11.0.txt.gz (8.3 Gb)	SW alignment scores between proteins within each STRING species	
protein.homology_best.v11.0.txt.gz (286.7 Gb)	SW alignment scores for best hits between proteins across species	
clusters.proteins.v11.0.txt.gz (5.6 Gb)	hierarchical STRING clusters and their proteins	
clusters.info.v11.0.txt.gz (103.2 Mb)	hierarchical STRING clusters annotations	
clusters.tree.v11.0.txt.gz (16.3 Mb)	hierarchical STRING clusters tree (represented as child-parent relationship)	
mapping_files (download directory)	separate identifier mapping files, for several frequently used name_spaces...	

STRING

FULL DATABASE DUMPS

File	Description	Access
database.schema.v11.0.pdf (121.8 Kb)	STRING database schema	
items_schema.v11.0.sql.gz (13 Gb)	full database, part I: the players (proteins, species, COGs,...)	
network_schema.v11.0.sql.gz (82.3 Gb)	full database, part II: the networks (nodes, edges, scores,...)	
evidence_schema.v11.0.sql.gz (13.2 Gb)	full database, part III: interaction evidence (but: excluding license-restricted data)	
homology_schema.v11.0.sql.gz (2700 Gb)	full database, part IV: homology data (all-against-all SIMAP similarity searches)	

nature

Subscribe

NEWS FEATURE · 02 APRIL 2020 · CORRECTION [03 APRIL 2020](#)

Special report: The simulations driving the world's response to COVID-19

How epidemiologists rushed to model the coronavirus pandemic.

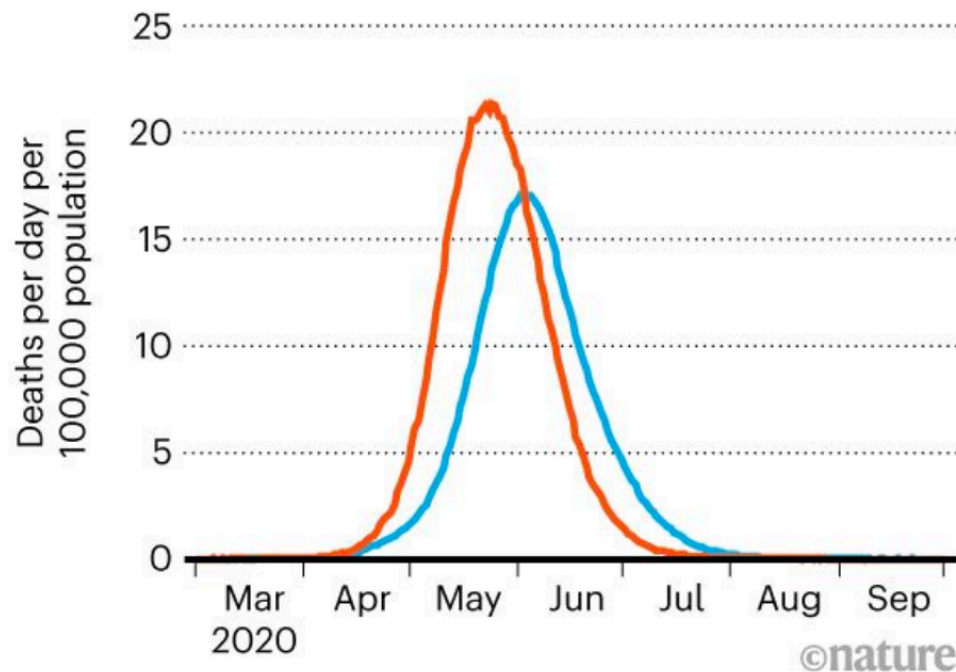
David Adam

COVID-19 Simulations

SIMULATION SHOCK

A model by Imperial College London in mid-March predicted a total of more than 500,000 UK deaths from COVID-19, and more than 2.2 million in the United States if no action was taken to stop the virus spreading in those countries.

— United Kingdom — United States



1. Ferguson, N. M. *et al.* Preprint at Spiral <https://doi.org/10.25561/77482> (2020).

2. Klepac, P. *et al.* Preprint at medRxiv <https://doi.org/10.1101/2020.02.16.20023754> (2020).

3. Ferguson, N. M. *et al. Nature* **437**, 209–214 (2005).

[PubMed](#) [Article](#) [Google Scholar](#)

4. Ferguson, N. M. *et al. Nature* **442**, 448–452 (2006).

[PubMed](#) [Article](#) [Google Scholar](#)

Coronavirus models: the basics

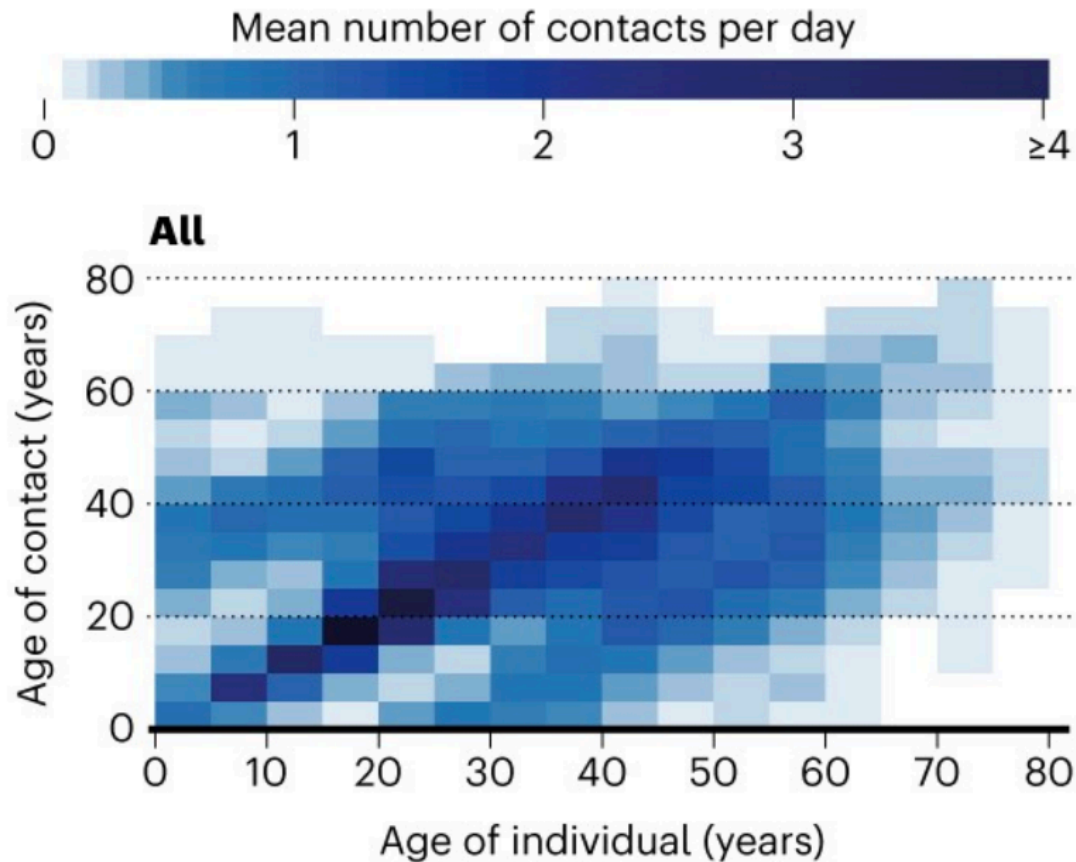
Many of the models simulating how diseases spread are unique to individual academic groups that have been developing them for years. But the mathematical principles are similar. They are based around trying to understand how people move between three main states, and how quickly: individuals are either susceptible (S) to the virus; have become infected (I); and then either recover (R) or die. The R group is presumed to be immune to the virus, so can no longer pass on the infection. People with natural immunity would also belong to this group.

The simplest SIR models make basic assumptions, such as that everyone has the same chance of catching the virus from an infected person because the population is perfectly and evenly mixed, and that people with the disease are all equally infectious until they die or recover. More-advanced models, which make the quantitative predictions policymakers need during an emerging pandemic, subdivide people into smaller groups – by age, sex, health status, employment, number of contacts, and so on – to set who meets whom, when and in which places (see ‘Measuring social mixing’).

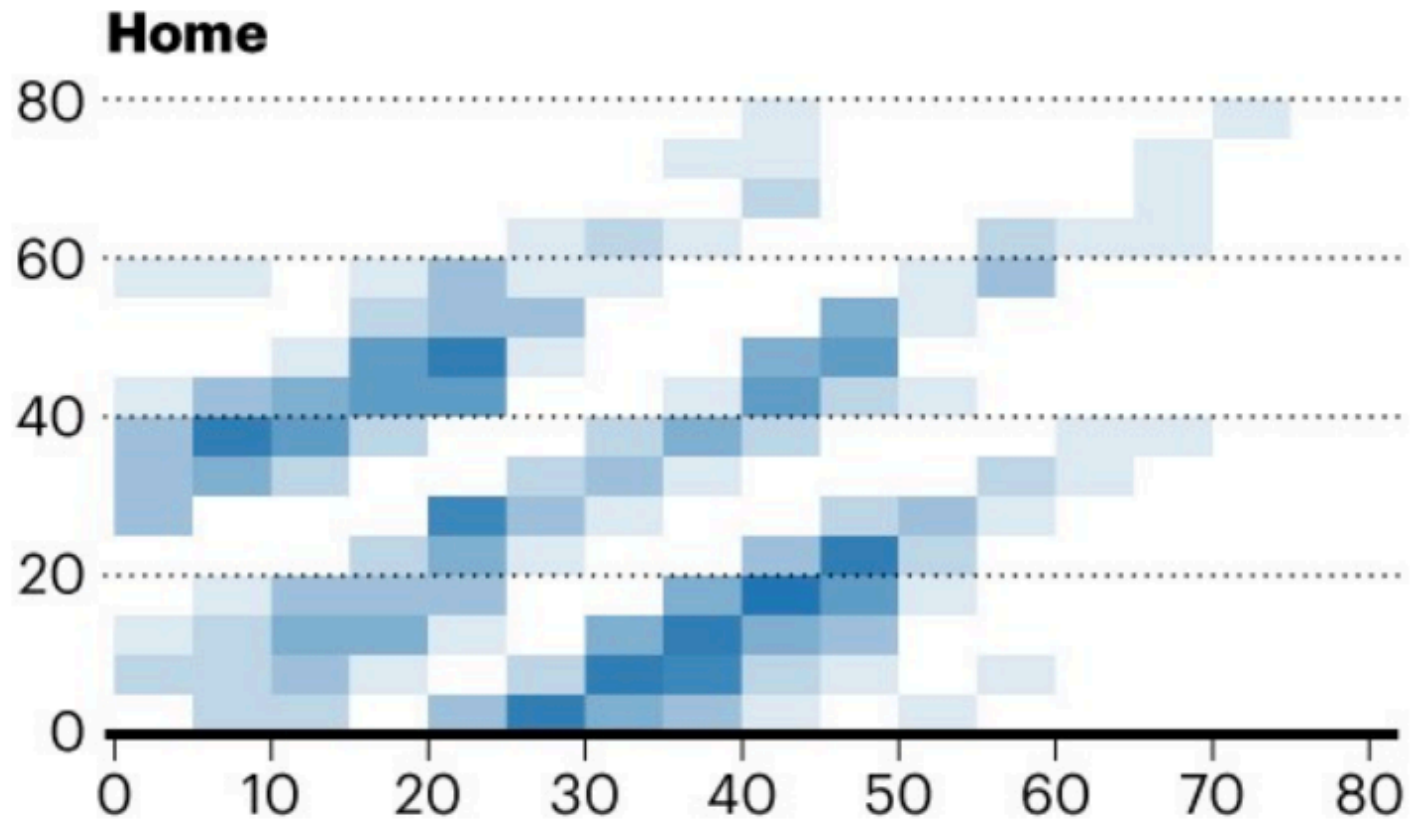
COVID-19 Simulations

MEASURING SOCIAL MIXING

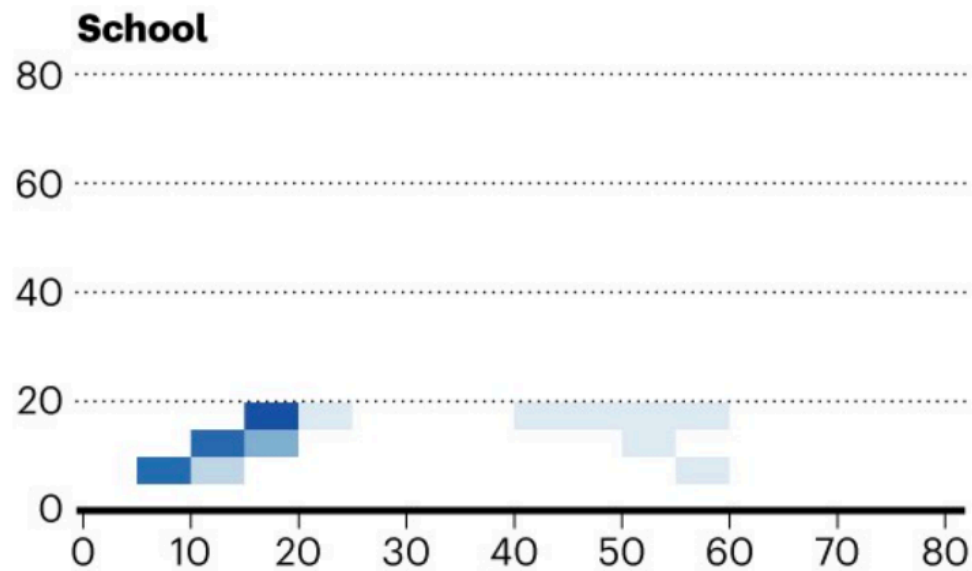
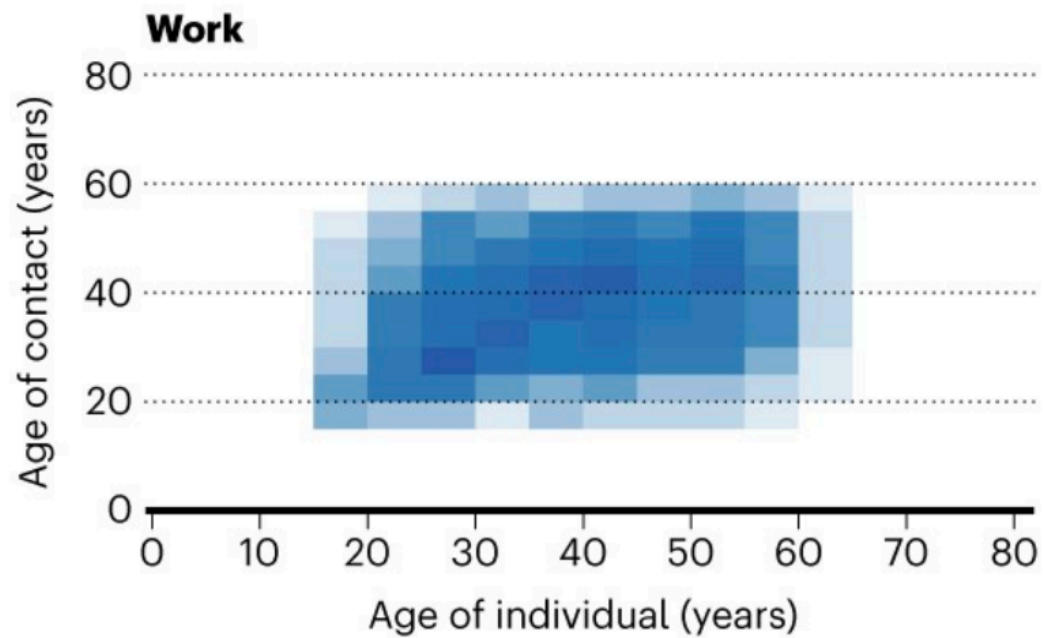
A model from the London School of Hygiene and Tropical Medicine, UK, lays out the average number of social contacts per day in China in the absence of virus-containment measures, by age group and location. This feeds simulations of how quickly a virus spreads.



COVID-19 Simulations



COVID-19 Simulations



Which model to choose?

The Imperial team has used both agent-based and equation-based models in this pandemic. The 16 March simulations that the team ran to inform the UK government's COVID-19 response used an agent-based model built in 2005 to see what would happen in Thailand if H5N1 avian flu mutated to a version that could spread easily between people³. (In 2006, the same model was used to study how the United Kingdom and the United States might mitigate the impact of a lethal flu pandemic⁴.) [Ferguson told *Nature* in 2005](#) that collecting detailed data on Thailand's population was harder than writing the programming code for the model. That code was not released when his team's projections on the coronavirus pandemic were first made public, but the team is working with Microsoft to tidy up the code and make it available, Ferguson says.

COVID-19 Simulations

On 26 March, Ferguson and his team released global projections of the impact of COVID-19 that uses the simpler equation-based approach⁵. It divides people into four groups: S, E, I and R, where 'E' refers to those who have been exposed, but who are not yet infectious. "They give broadly similar overall numbers," says epidemiologist Azra Ghani, who is also in the Imperial group. For instance, the global projections suggest that, had the United States taken no action against the virus, it would have seen 2.18 million deaths. By comparison, the earlier agent-based simulation, run using the same assumptions about mortality rate and reproduction number, estimated 2.2 million US deaths¹.

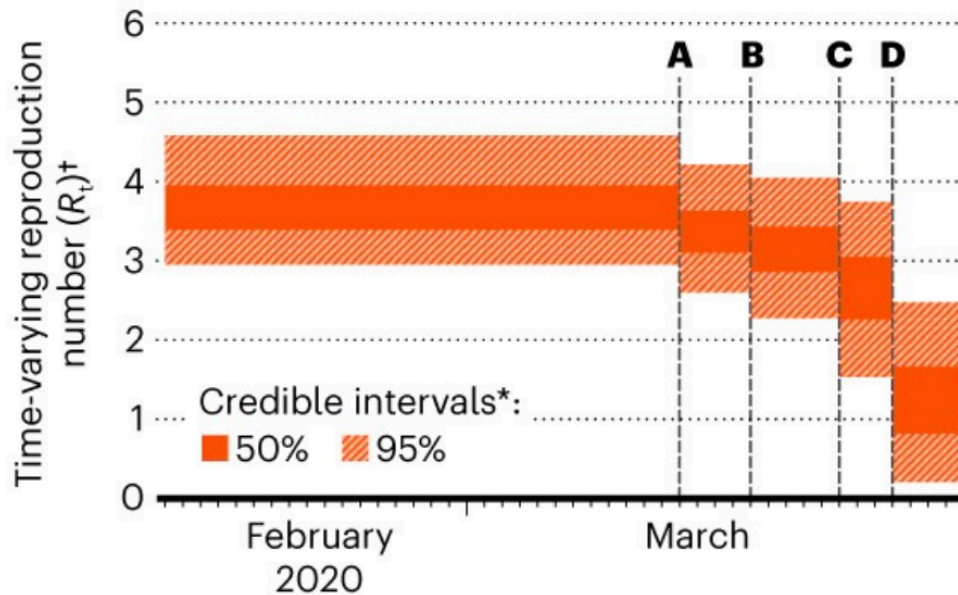
The different kinds of model have their own strengths and weaknesses, says Vittoria Colizza, a modeller at the Pierre Louis Institute of Epidemiology and Public Health in Paris, who is advising the French government during the current emergency. "It depends on the question you want to ask," she says.

COVID-19 Simulations

LOCKDOWNS KEEP INFECTIONS AT BAY

UK interventions reduced the virus's effective reproduction number — the average number of people an infected person passes the disease to — from almost four to around one, a model from Imperial College London says.

A: Self-isolation **B:** Social distancing **C:** School closure
D: Public events banned and complete lockdown



*Bayesian statistics: interval within which unobserved parameter falls, with particular probability.

† R_t : average number of infections, at time t , per infected individual over the course of their infection. If R_t is maintained at <1 , new infections decrease, resulting in control of the epidemic.

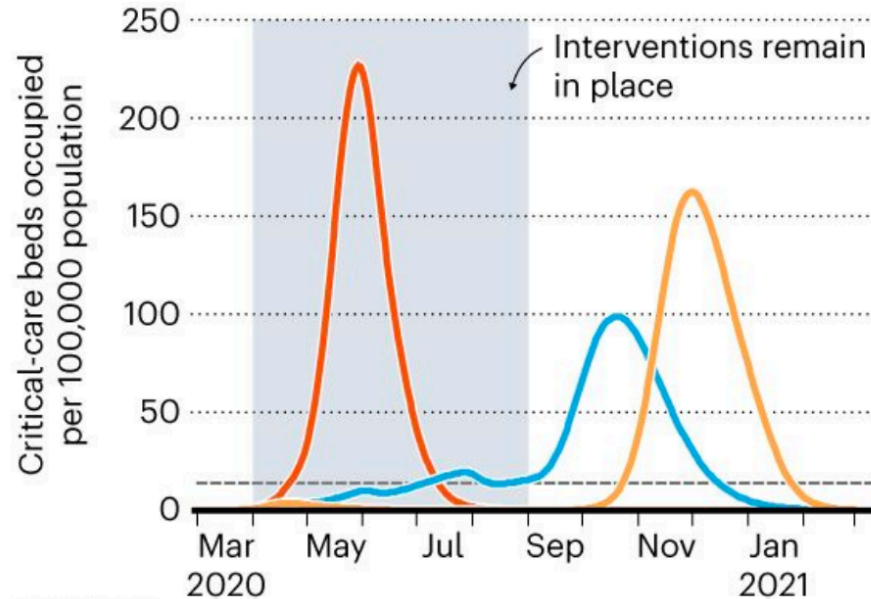
©nature

COVID-19 Simulations

A SECOND WAVE

In the United States, implementing measures to contain the virus could stop people with COVID-19 from immediately overwhelming the country's critical-care hospital-bed capacity, a simulation from Imperial College London suggests. But a second wave of the pandemic might be expected later in the year.

- Estimated critical-care bed capacity
- Do nothing
- Case isolation, household quarantine and general social distancing
- School and university closure, case isolation and general social distancing



©nature

COVID-19 Simulations

1. Ferguson, N. M. *et al.* Preprint at Spiral <https://doi.org/10.25561/77482> (2020).

2. Klepac, P. *et al.* Preprint at medRxiv <https://doi.org/10.1101/2020.02.16.20023754> (2020).

3. Ferguson, N. M. *et al.* *Nature* **437**, 209–214 (2005).
[PubMed](#) [Article](#) [Google Scholar](#)

4. Ferguson, N. M. *et al.* *Nature* **442**, 448–452 (2006).
[PubMed](#) [Article](#) [Google Scholar](#)

5. Walker, P. G. T. *et al.* Preprint at Spiral <https://go.nature.com/2yqz47x> (2020).

6. Funk, S. *et al.* *PLoS Comput. Biol.* **15**, e1006785 (2019).
[PubMed](#) [Article](#) [Google Scholar](#)

7. Prem, K. *et al.* *Lancet Public Health* [https://doi.org/10.1016/S2468-2667\(20\)30073-6](https://doi.org/10.1016/S2468-2667(20)30073-6) (2020).
[Article](#) [Google Scholar](#)

8. Huang, C. *et al.* *Lancet* **395**, 497–506 (2020).
[PubMed](#) [Article](#) [Google Scholar](#)

9. Flaxman, S. *et al.* Preprint at Spiral <https://doi.org/10.25561/77731> (2020).

10. Lourenço, J. *et al.* Preprint at medRxiv <https://doi.org/10.1101/2020.03.24.20042291> (2020).

11. Zhang, J. *et al.* Preprint at medRxiv <https://doi.org/10.1101/2020.03.19.20039107> (2020).



Article

Proteomics of SARS-CoV-2-infected host cells reveals therapy targets

<https://doi.org/10.1038/s41586-020-2332-7>

Received: 27 February 2020

Accepted: 6 May 2020

Published online: 14 May 2020

Denisa Bojkova^{1,7}, Kevin Klann^{2,7}, Benjamin Koch^{3,7}, Marek Widera¹, David Krause², Sandra Ciesek^{1,4}, Jindrich Cinatl¹✉ & Christian Münch^{2,5,6}✉

A novel coronavirus was recently discovered and termed SARS-CoV-2. Human infection can cause coronavirus disease 2019 (COVID-19), which has been rapidly spreading around the globe^{1,2}. SARS-CoV-2 shows some similarities to other coronaviruses. However, treatment options and a cellular understanding of SARS-CoV-2 infection are lacking. Here we identify the host cell pathways modulated by SARS-CoV-2 infection and show that inhibition of these pathways prevent viral replication in human cells. We established a human cell culture model for infection with SARS-CoV-2 clinical isolate. Employing this system, we determined the SARS-CoV-2 infection profile by translato³ and proteome proteomics at different times after infection. These analyses revealed that SARS-CoV-2 reshapes central cellular pathways, such as translation, splicing, carbon metabolism and nucleic acid metabolism. Small molecule inhibitors targeting these pathways prevented viral replication in cells. Our results reveal the cellular infection profile of SARS-CoV-2 and led to the identification of drugs inhibiting viral replication. We anticipate our results to guide efforts to understand the molecular mechanisms underlying host cell modulation upon SARS-CoV-2 infection. Furthermore, our findings provide insight for the development of therapy options for COVID-19.



Article

Proteomics of SARS-CoV-2-infected host cells reveals therapy targets

<https://doi.org/10.1038/s41586-020-2332-7>

Received: 27 February 2020

Accepted: 6 May 2020

Published online: 14 May 2020

Denisa Bojkova^{1,7}, Kevin Klann^{2,7}, Benjamin Koch^{3,7}, Marek Widera¹, David Krause², Sandra Ciesek^{1,4}, Jindrich Cinatl¹ & Christian Münch^{2,5,6}

A novel coronavirus was recently discovered and termed SARS-CoV-2. Human infection can cause coronavirus disease 2019 (COVID-19), which has been rapidly

Received
27 February 2020

Accepted
06 May 2020

Published
14 May 2020

to other
inding of
ways modulated
prevent viral
el for infection
ned the

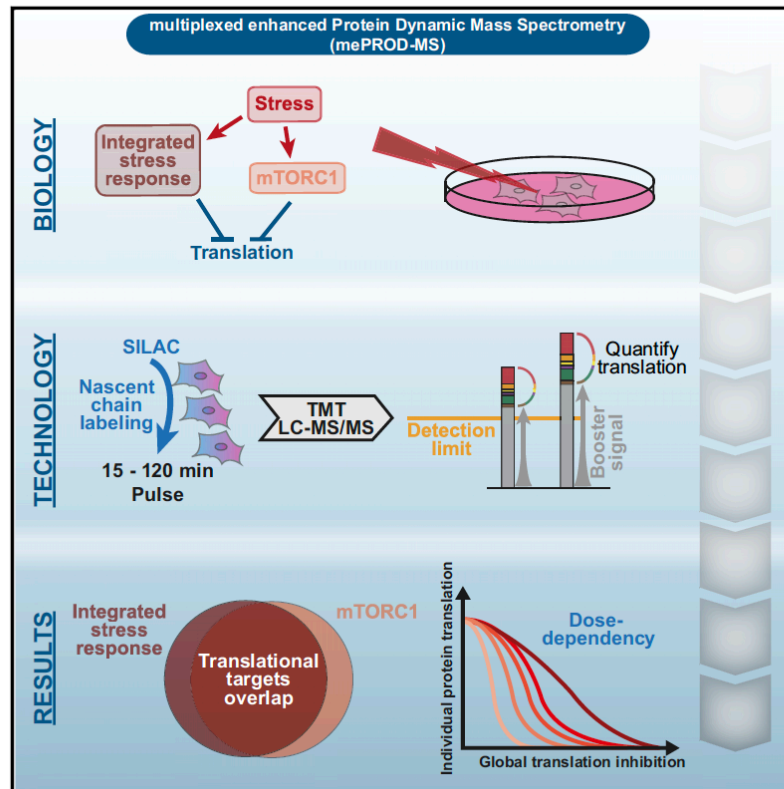
SARS-CoV-2 infection profile by translato³ and proteome proteomics at different times after infection. These analyses revealed that SARS-CoV-2 reshapes central cellular pathways, such as translation, splicing, carbon metabolism and nucleic acid metabolism. Small molecule inhibitors targeting these pathways prevented viral replication in cells. Our results reveal the cellular infection profile of SARS-CoV-2 and led to the identification of drugs inhibiting viral replication. We anticipate our results to guide efforts to understand the molecular mechanisms underlying host cell modulation upon SARS-CoV-2 infection. Furthermore, our findings provide insight for the development of therapy options for COVID-19.



Molecular Cell

Functional Translatome Proteomics Reveal Converging and Dose-Dependent Regulation by mTORC1 and eIF2 α

Graphical Abstract



Authors

Kevin Klann, Georg Tascher,
Christian Münch

Correspondence

ch.muench@em.uni-frankfurt.de

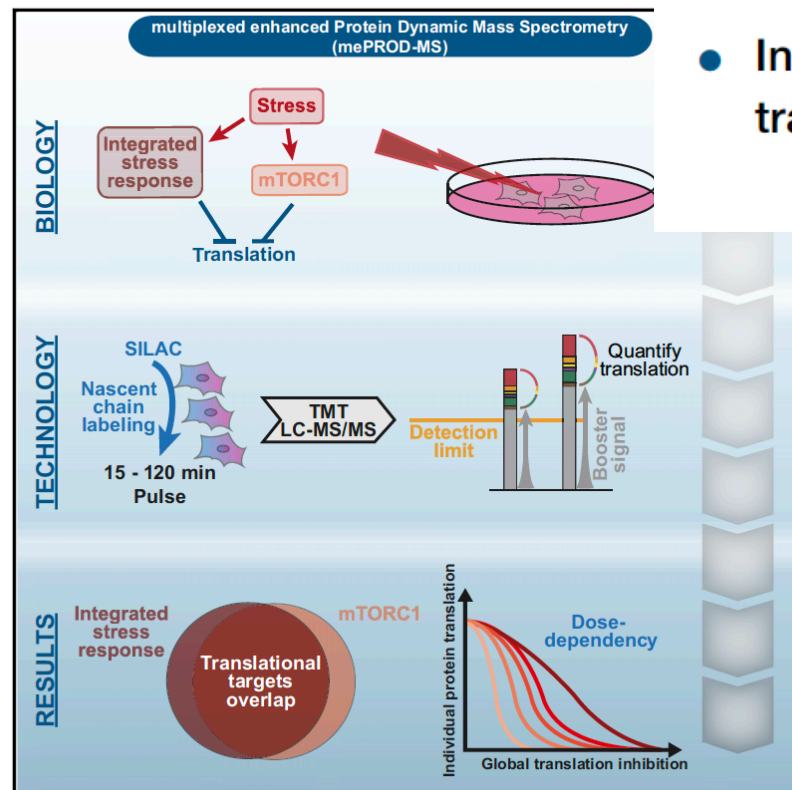
In Brief

Klann et al. developed a proteomics method that uses signal amplification to enable acute translation measurements. Employing this method, they reveal that the integrated stress response and mTORC1 inhibit translation of the same set of proteins. Target specificity is defined by global translation attenuation in a dose-dependent manner.

Molecular Cell

Functional Translatome of Converging and Dose-Dependent mTORC1 and eIF2 α

Graphical Abstract



Highlights

- A proteomics method quantifies nascent protein chains minutes after labeling
- The integrated stress response and mTORC1 share translation targets
- Individual protein translation levels are driven by global translation status

In Brief

Klann et al. developed a proteomics method that uses signal amplification to enable acute translation measurements. Employing this method, they reveal that the integrated stress response and mTORC1 inhibit translation of the same set of proteins. Target specificity is defined by global translation attenuation in a dose-dependent manner.

- **Proteomics + COVID**

- **Bojkova, D., Klann, K., Koch, B. *et al.***

- Proteomics of SARS-CoV-2-infected host cells reveals therapy targets. *Nature* (2020).**

- <https://doi.org/10.1038/s41586-020-2332-7>

- **Klann K, Tascher G, Münch C.,**

- Functional Translatome Proteomics Reveal Converging and Dose-Dependent Regulation by mTORC1 and eIF2 α**
Mol Cell. 2020 Feb 20;77(4):913-925.e4.

- doi: 10.1016/j.molcel.2019.11.010. Epub 2019 Dec 4.**

- **Schwanhäusser, B., Busse, D., Li, N., Dittmar, G.,**

- Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M. (2011).**

- Global quantification of mammalian gene expression control.**

- Nature 473, 337–342.**

Article

Proteomics of SARS-CoV-2-infected host cells reveals therapy targets

<https://doi.org/10.1038/s41586-020-2332-7>

Received: 27 February 2020

Accepted: 6 May 2020

Published online: 14 May 2020

Denisa Bojkova^{1,7}, Kevin Klann^{2,7}, Benjamin Koch^{3,7}, Marek Widera¹, David Krause², Sandra Ciesek^{1,4}, Jindrich Cinatl¹✉ & Christian Münch^{2,5,6}✉

A novel coronavirus was recently discovered and termed SARS-CoV-2. Human infection can cause coronavirus disease 2019 (COVID-19), which has been rapidly spreading around the globe^{1,2}. SARS-CoV-2 shows some similarities to other coronaviruses. However, treatment options and a cellular understanding of SARS-CoV-2 infection are lacking. Here we identify the host cell pathways modulated by SARS-CoV-2 infection and show that inhibition of these pathways prevent viral replication in human cells. We established a human cell culture model for infection with SARS-CoV-2 clinical isolate. Employing this system, we determined the SARS-CoV-2 infection profile by translato³ and proteome proteomics at different times after infection. These analyses revealed that SARS-CoV-2 reshapes central cellular pathways, such as translation, splicing, carbon metabolism and nucleic acid metabolism. Small molecule inhibitors targeting these pathways prevented viral replication in cells. Our results reveal the cellular infection profile of SARS-CoV-2 and led to the identification of drugs inhibiting viral replication. We anticipate our results to guide efforts to understand the molecular mechanisms underlying host cell modulation upon SARS-CoV-2 infection. Furthermore, our findings provide insight for the development of therapy options for COVID-19.



SARS Proteomics

Abstract

A novel coronavirus was recently discovered and termed SARS-CoV-2. Human infection can cause coronavirus disease 2019 (COVID-19), which has been rapidly spreading around the globe^{1,2}. SARS-CoV-2 shows some similarities to other coronaviruses. However, treatment options and a cellular understanding of SARS-CoV-2 infection are lacking. Here we identify the host cell pathways modulated by SARS-CoV-2 infection and show that inhibition of these pathways prevent viral replication in human cells. We established a human cell culture model for infection with SARS-CoV-2 clinical isolate. Employing this system, we determined the SARS-CoV-2 infection profile by translato³ and proteome proteomics at different times after infection. These analyses revealed that SARS-CoV-2 reshapes central cellular pathways, such as translation, splicing, carbon metabolism and nucleic acid metabolism. Small molecule inhibitors targeting these pathways prevented viral replication in cells. Our results reveal the cellular infection profile of SARS-CoV-2 and led to the identification of drugs inhibiting viral replication. We anticipate our results to guide efforts to understand the molecular mechanisms underlying host cell modulation upon SARS-CoV-2 infection. Furthermore, our findings provide insight for the development of therapy options for COVID-19.



SARS Proteomics

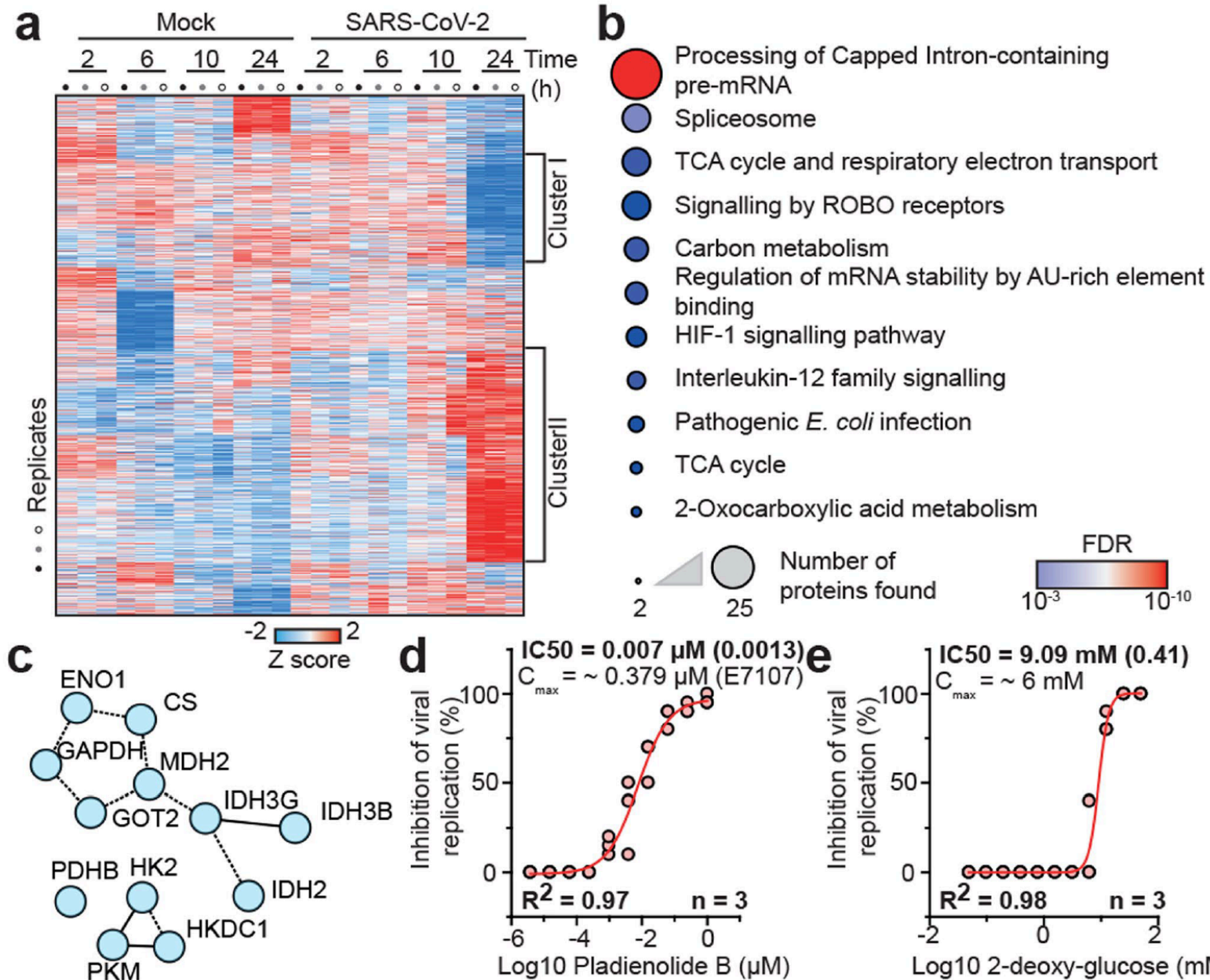
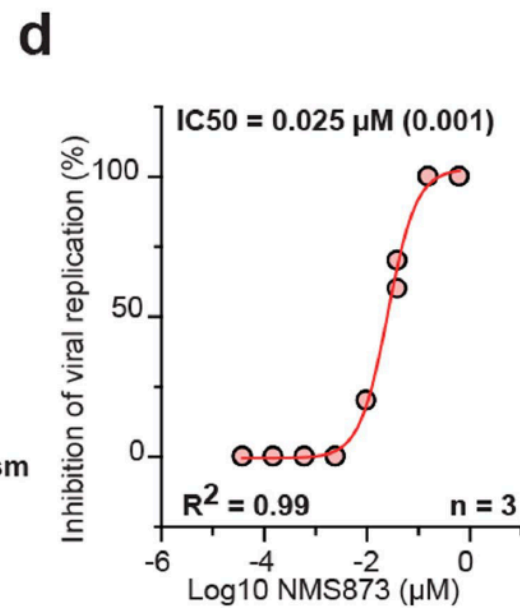
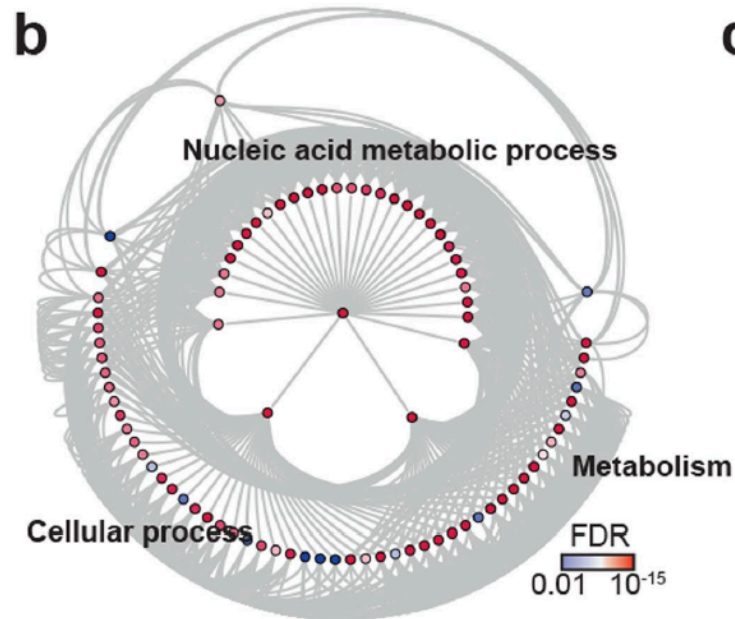
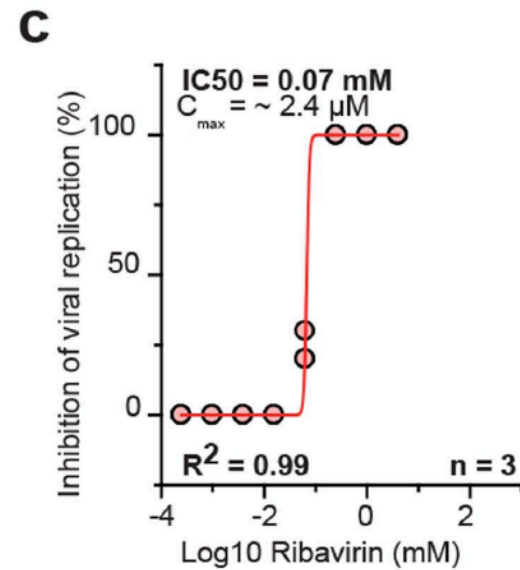
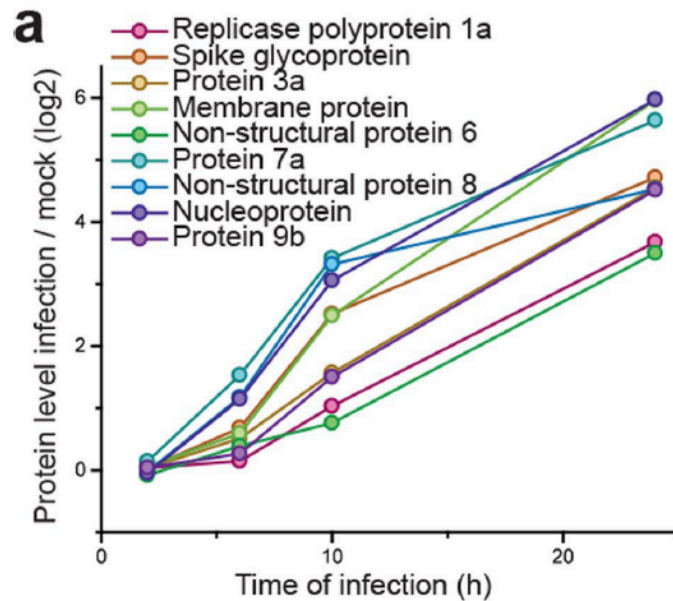


Fig. 3 | SARS-CoV-2 infection profiling reveals cellular pathways essential for replication. a, Patterns of protein levels across all samples. Shown are proteins tested significant (two-sided, unpaired t-test with equal variance assumed, $P < 0.05$, $n = 3$) in at least one infected sample compared to corresponding control. Data was standardized using Z scoring before row-wise clustering and plotting. **b**, Reactome pathway analysis of protein network created from cluster II (**a**, see Table S4). Pathway results are shown with number of proteins found in dataset and computed FDR for pathway enrichment. **c**, Functional interaction network of proteins found annotated to carbon metabolism in Reactome pathway analysis. Lines indicate functional interaction. **d**, **e**, Antiviral assay showing inhibition of viral replication in dependency of pladienolide B (**d**, $n = 3$) and 2-deoxy-glucose (**e**, $n = 3$) concentration. Each data point indicates a biological replicate and red line shows dose response curve fit. R^2 and IC_{50} values were computed from the curve fit and s.d. of IC_{50} is indicated in brackets. All n numbers represent independent biological samples.

SARS Proteomics: Concept



TMT data: Methods

Mass spectrometry data analysis

Raw files were analyzed using Proteome Discoverer (PD) 2.4 software (ThermoFisher Scientific). Spectra were selected using default settings and database searches performed using SequestHT node in PD. Database searches were performed against trypsin digested Homo Sapiens SwissProt database, SARS-CoV-2 database (Uniprot pre-release) and FASTA files of common contaminants (`contaminants.fasta` provided with MaxQuant) for quality control. Fixed modifications were set as TMT6 at the N-terminus and carbamidomethyl at cysteine residues. One search node was set up to search with TMT6 (K) and methionine oxidation as static modifications to search for light peptides and one search node was set up with TMT6+K8 (K, +237.177), Arg10 (R, +10.008) and methionine oxidation as static modifications to identify heavy peptides. Searches were performed using Sequest HT. After search, posterior error probabilities were calculated and PSMs filtered using Percolator using default settings. Consensus Workflow for reporter ion quantification was performed with default settings, except the minimal signal-to-noise ratio was set to 5. Results were then exported to Excel files for further processing. For proteome quantification all PSMs were summed intensity normalized, followed by IRS³² and TMM³³ normalization and peptides corresponding to a given UniProt Accession were summed including all modification states.

Mass spectrometry data analysis

For translome measurements, Excel files were processed in Python, as previously described³. Python 3.6 was used together with the following packages: pandas 0.23.4³⁴, numpy 1.15.4³⁵, scipy 1.3.0. Excel files with normalized PSM data were read in and each channel was normalized to the lowest channel based on total intensity. For each peptide sequence, all possible modification states containing a heavy label were extracted and the intensities for each channel were averaged between all modified peptides. Baseline subtraction was performed by subtracting the measured intensities for the non-SILAC-labeled sample from all other values. Negative intensities were treated as zero. The heavy label incorporation at the protein level was calculated by summing the intensities of all peptide sequences belonging to one unique protein accession. These values were combined with the standard protein output of PD 2.4 to add annotation data to the master protein accessions.

Mass spectrometry data analysis

Hierarchical clustering and profile comparison

Hierarchical cluster analysis and comparison with viral protein profiles for all samples was performed using Perseus³⁶ software package (version 1.6.5.0) after centering and scaling of data (Z scores). K-means pre-processing was performed with a cluster number of 12 and a maximum of 10 iterations. For the comparison of profiles, the viral profiles were Z scored and averaged to generate reference profile. Profiles of all proteins were compared to the reference (Pearson), distances and False discovery rates computed.

Network analysis

For network analysis, Cytoscape 3.7.1³⁷ software was used with BiNGO 3.0.3³⁸ plugin for GO term analysis, EnrichmentMap 3.1.0³⁹ and ReactomeFI 6.1.0⁴⁰. For GO-term analyses, gene sets were extracted from data as indicated using fold change and significance cut-offs.

Statistical analysis

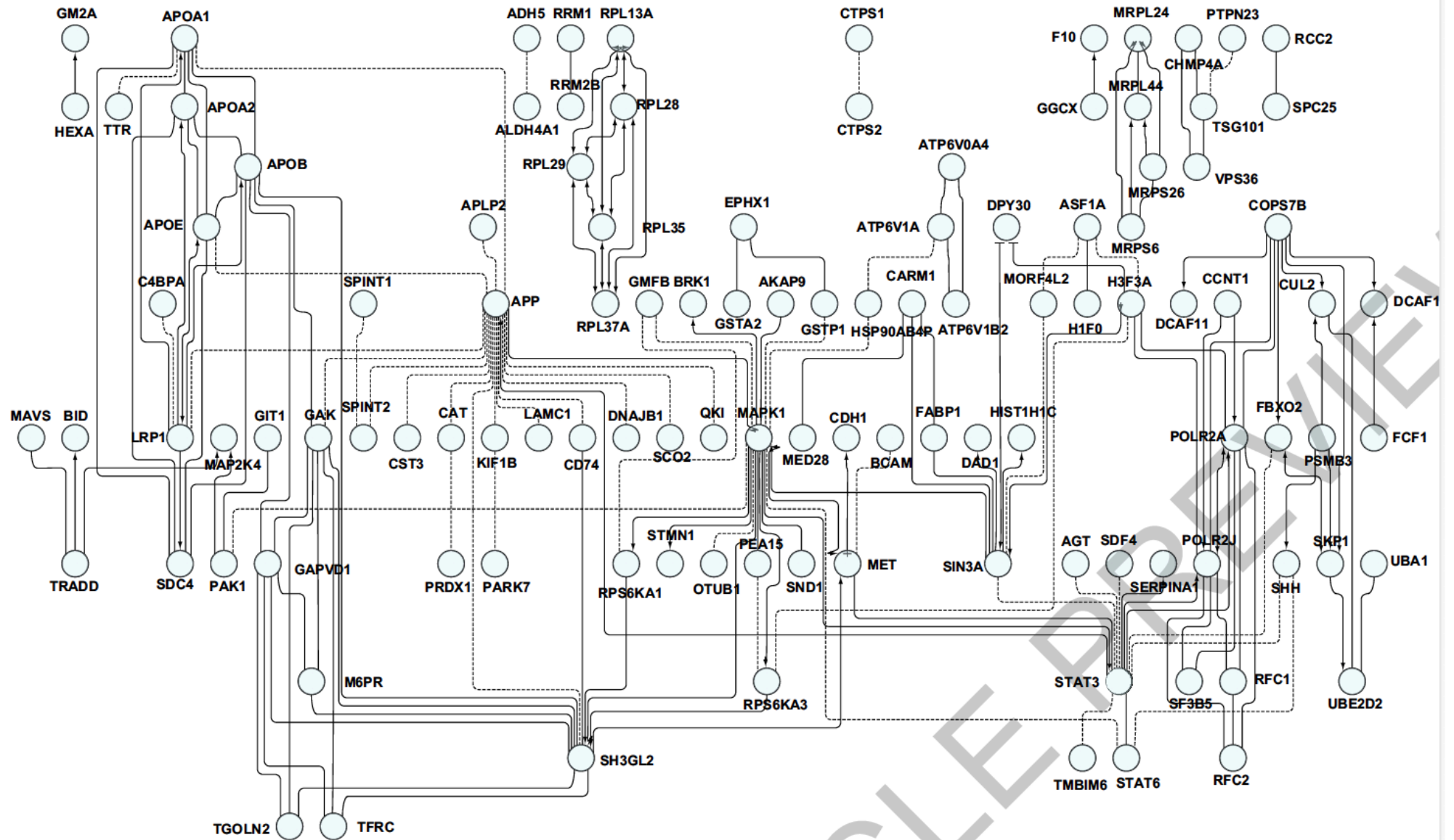
Significance was, unless stated otherwise, tested using unpaired two-sided students t-tests with equal variance assumed. Statistical analysis was performed using OriginPro 2020 analysis software. For network and GO analysis all statistical computations were performed by the corresponding packages.

Data availability

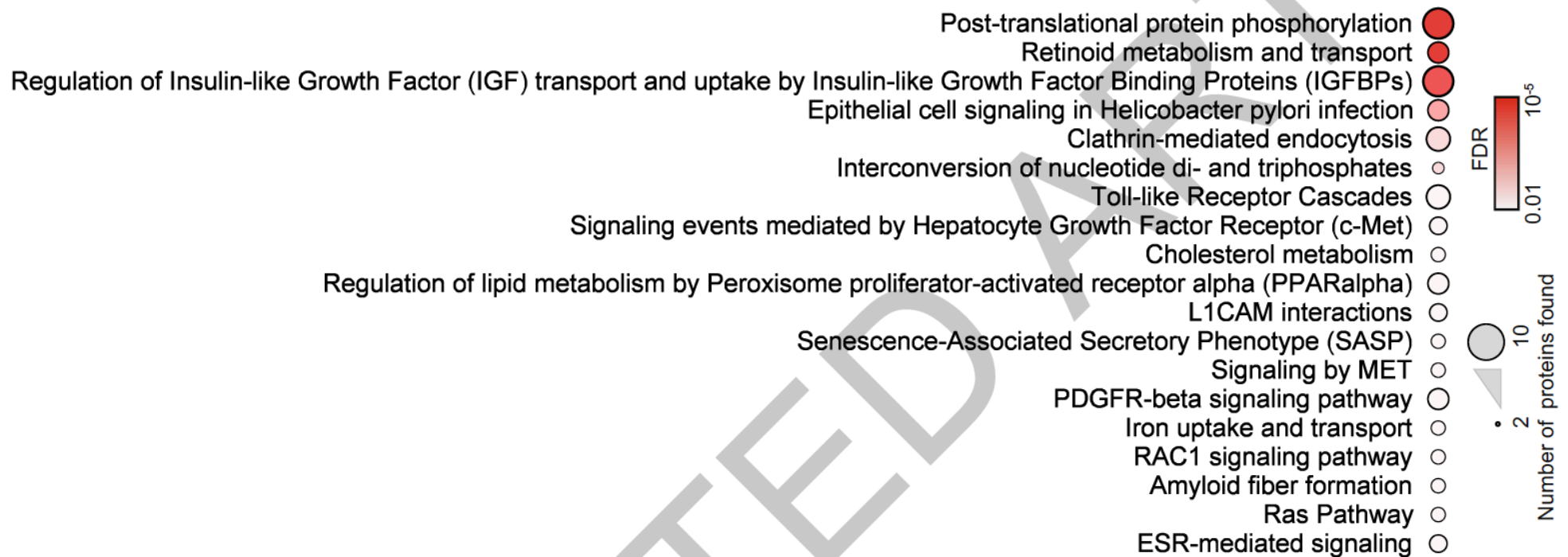
The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE⁴¹ partner repository with the dataset identifier PXD017710. We furthermore created a webpage (<http://corona.papers.biochem2.com/>) visualizing the presented data for easy access of the published data.



Decreased during SARS-CoV-2 infection



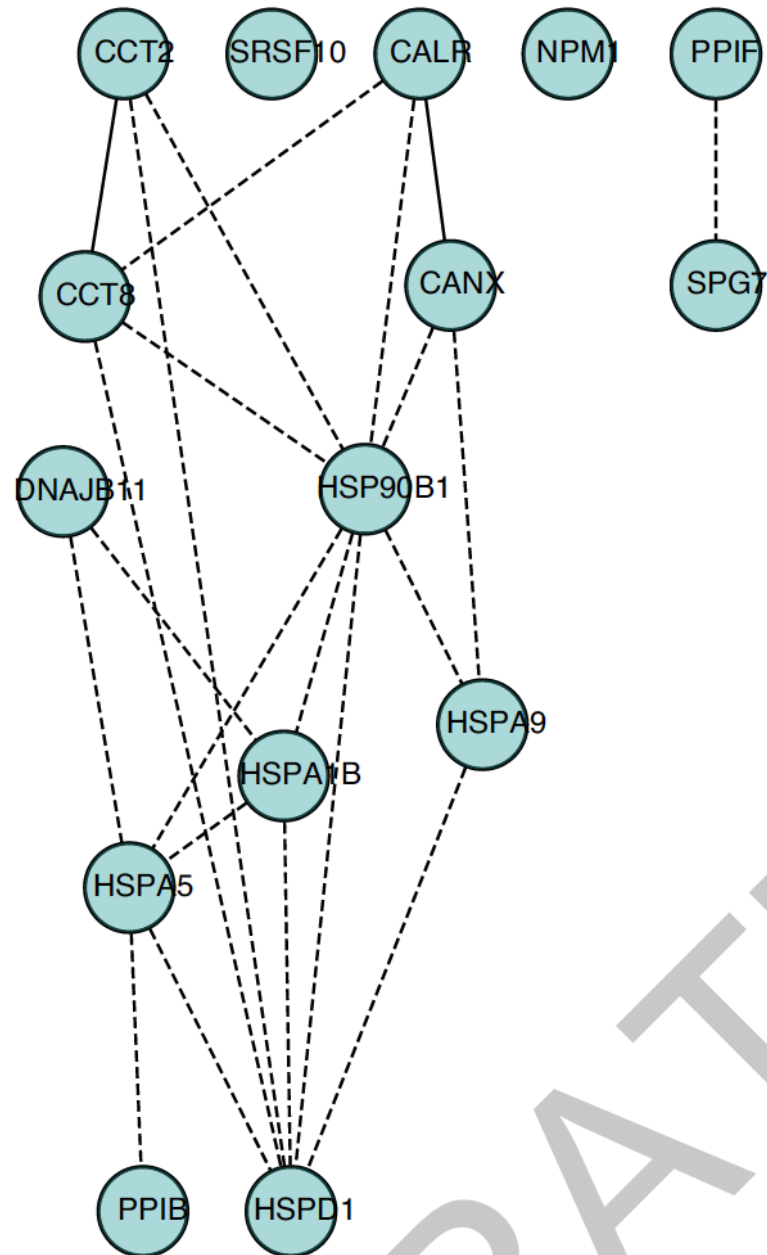
Decreased during SARS-CoV-2 infection

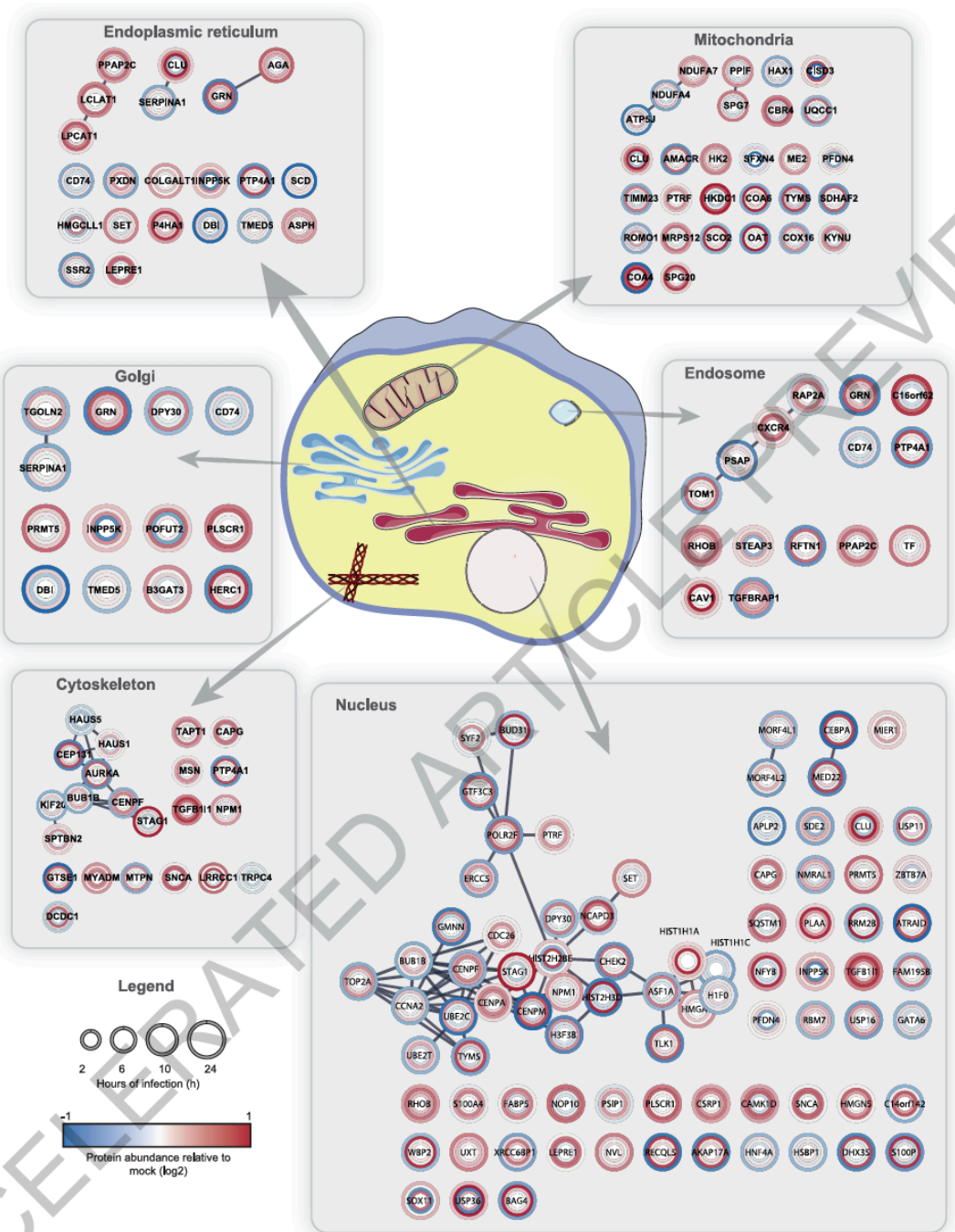


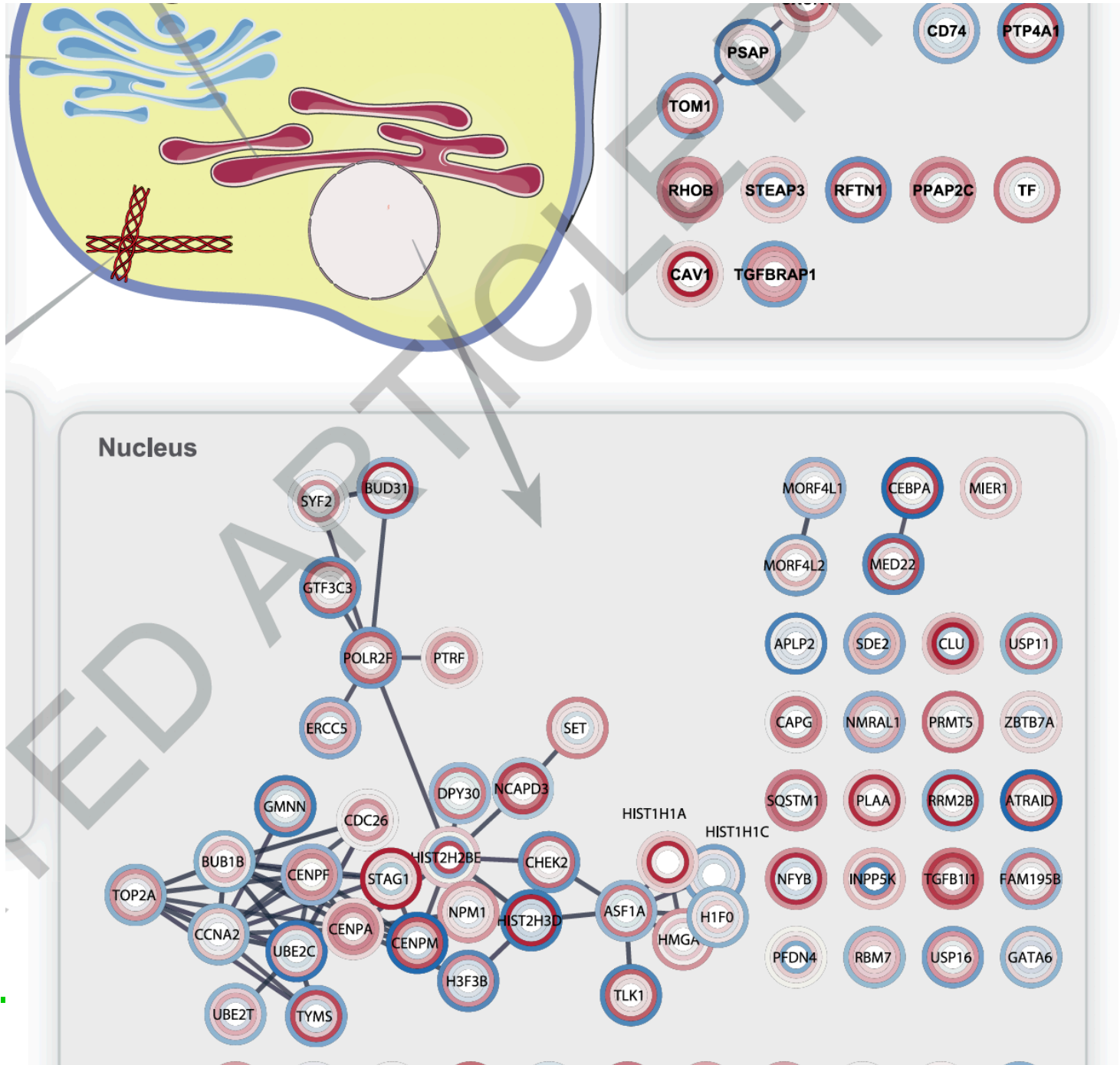
Extended Data Fig. 4 | Network of proteins decreased during SARS-CoV-2 infection. a, Proteins belonging to cluster I in Fig. 3a were used for functional interaction network creation. Lines indicate functional interactions. The network was created using the ReactomeFI plugin in cytoscape, protein names

added in the plugin and the network adjusted by the yFiles Layout algorithm. **b**, ReactomeFI network analysis of proteins downregulated in total protein levels. Circle size represents number of proteins found in the pathway, colour shows FDR for enrichment.

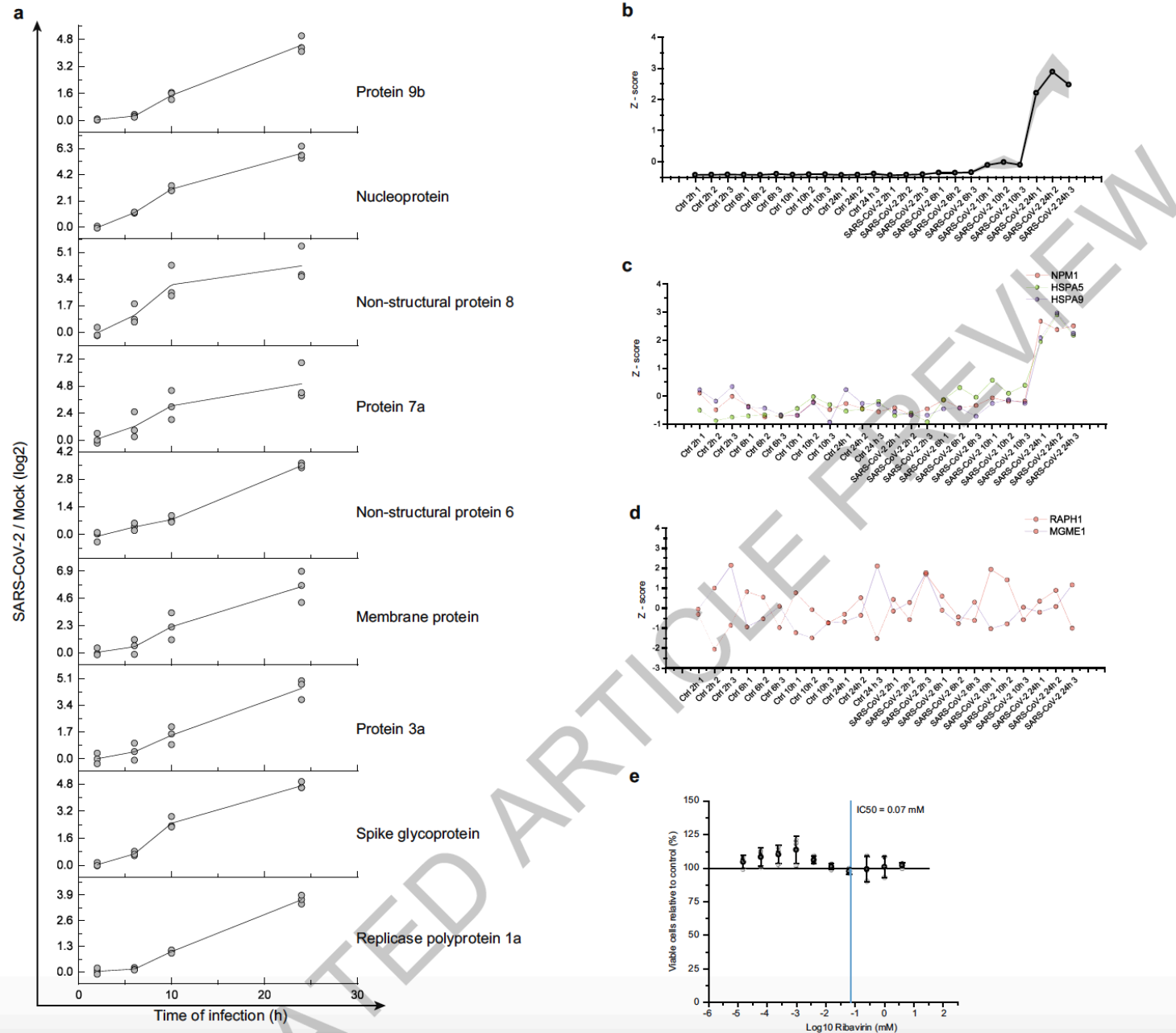
Unfolded protein binding







SARS Proteomics



SARS Proteomics

