Article

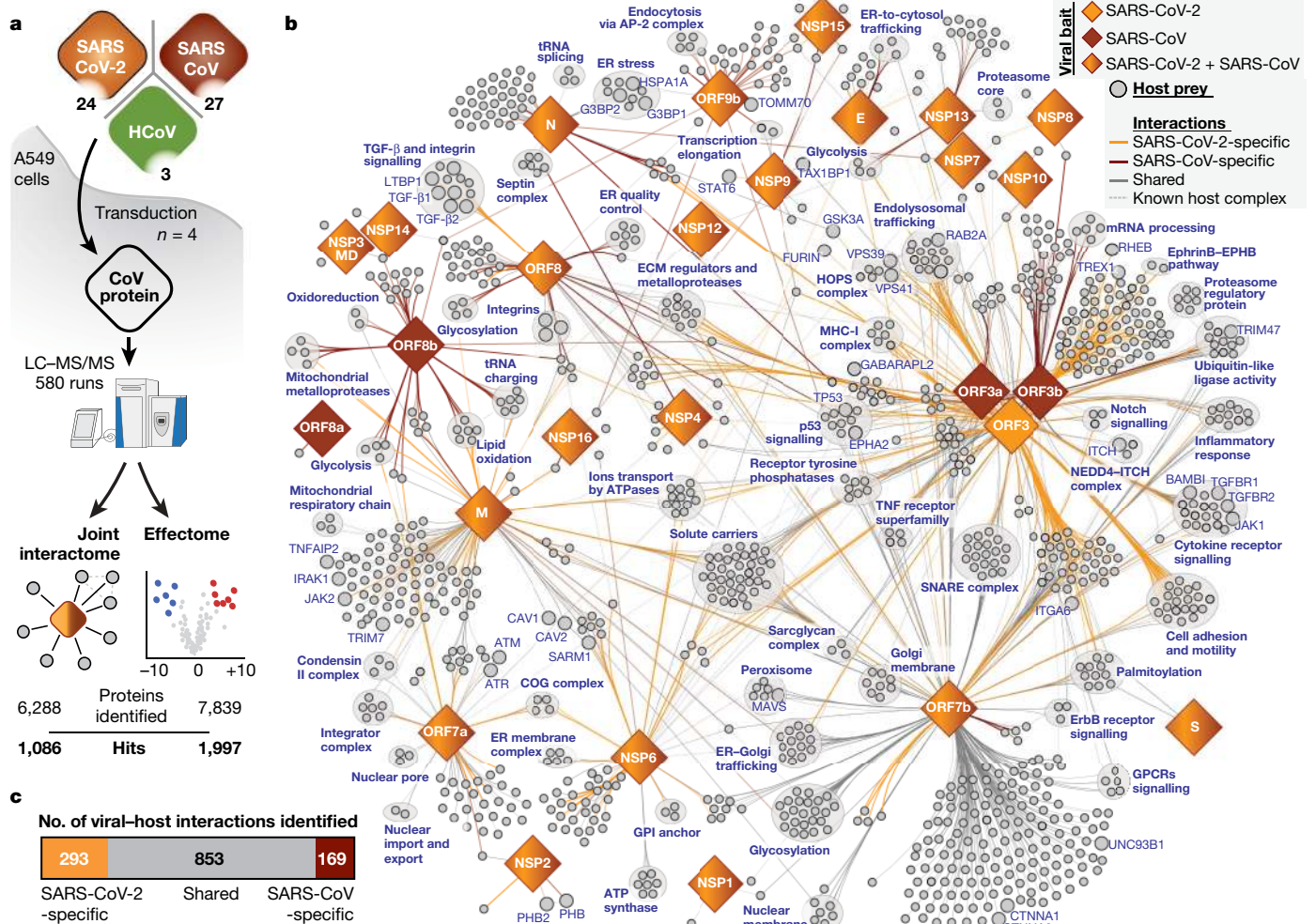# Multilevel proteomics reveals host perturbations by SARS-CoV-2 and SARS-CoV

Alexey Stukalov[1,15], Virginie Girault[1,15], Vincent Grass[1,15], Ozge Karayel[2,15], Valter Bergant[1,15], Christian Urban[1,15], Darya A. Haas[1,15], Yiqi Huang[1,15], Lila Oubraham[1], Anqi Wang[1], M. Sabri Hamad[1], Antonio Piras[1], Fynn M. Hansen[2], Maria C. Tanzer[2], Igor Paron[2], Luca Zinzula[3], Thomas Engleitner[4], Maria Reinecke[5,6,7], Teresa M. Lavacca[1], Rosina Ehmann[8,9], Roman Wölfel[8,9], Jörg Jores[10], Bernhard Kuster[5,6,7], Ulrike Protzer[1,9], Roland Rad[4], John Ziebuhr[11], Volker Thiel[12,13], Pietro Scaturro[1,14], Matthias Mann[2] & Andreas Pichlmair[1,9 ✉]

The emergence and global spread of SARS-CoV-2 has resulted in the urgent need for an in-depth understanding of molecular functions of viral proteins and their interactions with the host proteome. Several individual omics studies have extended our knowledge of COVID-19 pathophysiology[1–10]. Integration of such datasets to obtain a holistic view of virus–host interactions and to define the pathogenic properties of SARS-CoV-2 is limited by the heterogeneity of the experimental systems. Here we report a concurrent multi-omics study of SARS-CoV-2 and SARS-CoV. Using state-of-the-art proteomics, we profiled the interactomes of both viruses, as well as their influence on the transcriptome, proteome, ubiquitinome and phosphoproteome of a lung-derived human cell line. Projecting these data onto the global network of cellular interactions revealed crosstalk between the perturbations taking place upon infection with SARS-CoV-2 and SARS-CoV at different levels and enabled identification of distinct and common molecular mechanisms of these closely related coronaviruses. The TGF-β pathway, known for its involvement in tissue fibrosis, was specifically dysregulated by SARS-CoV-2 ORF8 and autophagy was specifically dysregulated by SARS-CoV-2 ORF3. The extensive dataset (available at https://covinet.innatelab.org) highlights many hotspots that could be targeted by existing drugs and may be used to guide rational design of virus- and host-directed therapies, which we exemplify by identifying inhibitors of kinases and matrix metalloproteases with potent antiviral effects against SARS-CoV-2.

To identify protein–protein interactions of SARS-CoV-2 and SARS-CoV and cellular proteins, we transduced A549 lung carcinoma cells with lentiviruses expressing individual haemagglutinin-tagged viral proteins (Fig. 1a, Extended Data Fig. 1a, Supplementary Table 1). Statistical modelling of quantitative data from affinity purification followed by mass spectrometry (AP–MS) analysis identified 1,801 interactions between 1,086 cellular proteins and 24 SARS-CoV-2 and 27 SARS-CoV bait proteins (Fig. 1b, Extended Data Fig. 1b, Supplementary Table 2), substantially increasing the number of reported interactions of SARS-CoV-2 and SARS-CoV[1,2,5,6,10,11] (Supplementary Table 10). The resulting virus–host interaction network revealed a wide range of cellular activities intercepted by SARS-CoV-2 and SARS-CoV (Fig. 1b, Extended Data Table 1, Supplementary Table 2). In particular, we observed that SARS-CoV-2 targets a number of key innate immunity regulators (ORF7b–MAVS and

ORF7b–UNC93B1), stress response components (N–HSPA1A) and DNA damage response mediators (ORF7a–ATM and ORF7a–ATR) (Fig. 1b, Extended Data Fig. 1c–e). Additionally, SARS-CoV-2 proteins interact with molecular complexes involved in intracellular trafficking (for example, endoplasmic reticulum–Golgi trafficking) and transport (for example, solute carriers and ion transport by ATPases) as well as cellular metabolism (for example, mitochondrial respiratory chain and glycolysis) (Fig. 1b, Extended Data Table 1, Supplementary Table 2). Comparing the AP–MS data of homologous SARS-CoV-2 and SARS-CoV proteins identified differences in the enrichment of individual host targets, highlighting potential virus-specific interactions (Fig. 1b (edge colour), c, Extended Data Figs. 1f, 2a, b, Supplementary Table 2). For instance, we recapitulated the known interactions between SARS-CoV NSP2 and prohibitins[12] (PHB and PHB2), but this interaction was not

Fig. 1 | Joint analysis of SARS-CoV-2 and SARS-CoV virus–host protein–protein interactomes. a, Systematic comparison of interactomes and host proteome changes (effectomes) of the 24 SARS-CoV-2 and 27 SARS-CoV viral proteins, using 3 homologues from human coronaviruses (HCoV-NL63 and HCoV-229E) as reference for pan-coronavirus specificity. **b**, Combined virus–host protein-interaction network of SARS-CoV-2 and SARS-CoV measured by AP–MS. Homologous viral proteins are displayed as a single node. Shared and virus-specific interactions are denoted by the edge colour. The edge intensity reflects the *P*-value for the interaction (with the smallest *P*-value represented by solid edges and the highest *P*-value (<0.001) represented by faded edges). ECM, extracellular matrix; ER, endoplasmic reticulum; GPCR, G-protein-coupled receptor; HOPS, homotypic fusion and protein-sorting; MHC, major histocompatibility complex; SNARE, soluble *N*-ethylmaleimide-sensitive factor attachment protein receptor; COG, conserved oligomeric Golgi. **c**, The numbers of unique and shared host interactions between the homologous proteins of SARS-CoV-2 and SARS-CoV.

conserved with SARS-CoV-2 NSP2, suggesting that the two viruses differ in their ability to modulate mitochondrial function and homeostasis through NSP2 (Extended Data Fig. 2a). The exclusive interaction of SARS-CoV-2 ORF8 with the TGF-β1–LTBP1 complex is another interaction that potentially explains the differences in pathogenicity of the two viruses (Extended Data Figs. 1f, 2b). Notably, disbalanced TGF-β signalling has been linked to lung fibrosis and oedema, a common complication of severe pulmonary diseases including COVID-19[13–16].

To map the virus–host interactions to the functions of viral proteins, we conducted a study of total proteomes of A549 cells expressing 54 individual viral proteins comprising the 'effectome' (Fig. 1a, Supplementary Table 3). This dataset provides clear links between changes in protein expression and virus–host interactions, as exemplified by ORF9b, which leads to a dysregulation of mitochondrial functions and binds to TOMM70, a known regulator of mitophagy[2,17] (Fig. 1b, Supplementary Tables 2, 3). Global pathway-enrichment analysis of the effectome dataset confirmed that ORF9b of both viruses led to mitochondrial dysregulation[2,18] (Extended Data Fig. 2c, Supplementary Table 3) and further highlighted virus-specific effects, as exemplified by the upregulation of proteins involved in cholesterol metabolism

(CYP51A1, DHCR7, IDI1 and SQLE) by SARS-CoV-2 NSP6 but not by SARS-CoV NSP6. Of note, cholesterol metabolism was recently shown to be implicated in SARS-CoV-2 replication and has been suggested as a promising target for drug development[19–21]. Besides perturbations at the pathway level, viral proteins also specifically modulated single host proteins, possibly explaining more specific molecular mechanisms involved in viral protein function. Focusing on the 180 most affected host proteins, we identified RCOR3, a putative transcriptional corepressor, as strongly upregulated by NSP4 of both viruses (Extended Data Figs. 2d, 3a). Notably, apolipoprotein B (APOB) was substantially regulated by ORF3 and NSP1 of SARS-CoV-2, suggesting that it has an important role in SARS-CoV-2 biology (Extended Data Fig. 3b).

Multi-omics profiling of virus infection

Although the interactome and the effectome provide in-depth information on the activity of individual viral proteins, we aimed to directly study their combined activities in the context of viral infection. To this end, we infected A549 cells expressing angiotensin-converting enzyme 2 (ACE2) (A549-ACE2 cells) (Extended Data Fig. 4a, b) with SARS-CoV-2
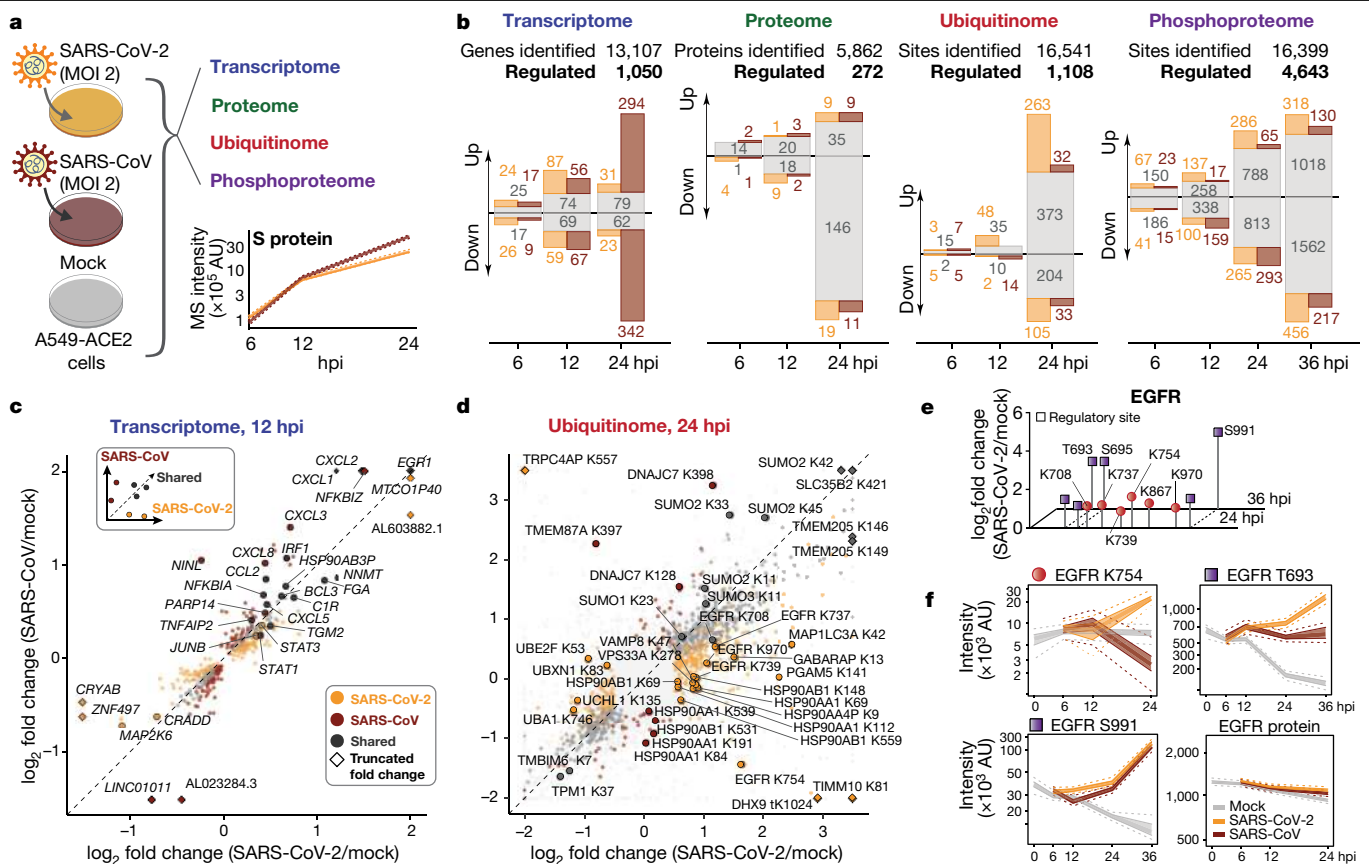
Fig. 2 | Multi-level profiling of SARS-CoV-2 and SARS-CoV infection.
a, Time-resolved profiling of parallel SARS-CoV-2 and SARS-CoV infection by multiple omics methods. The plot shows the mass spectroscopy (MS) intensity estimates for spike proteins of SARS-CoV-2 and SARS-CoV over time ($n = 4$ independent experiments). MOI, multiplicity of infection. **b**, The numbers of distinct transcripts, proteins, ubiquitination and phosphorylation sites that are significantly up- or downregulated at given time points after infection (relative to mock infection at the same time point). Transcripts, proteins or sites that are regulated similarly by SARS-CoV-2 and SARS-CoV infection are shown in grey, those regulated specifically by SARS-CoV-2 are in orange and those regulated by SARS-CoV are in brown. **c**, **d**, Comparison of host transcriptome 12 h (**c**) and ubiquitinome 24 h (**d**) after infection (hpi) with SARS-CoV-2 (x-axis) or SARS-CoV (y-axis) ($\log_2$ fold change in comparison to the mock infection samples at the same time point). Significantly regulated transcripts by moderated t-test with false discovery rate-corrected two-sided

P-value ≤0.05 (**c**) and significantly regulated sites by Bayesian linear model-based unadjusted two-sided P-value ≤10^{-3}, |$\log_2$ fold change| ≥0.5 (**d**) are coloured according to specificity as indicated. Diamonds indicate that the actual $\log_2$ fold change was truncated to fit into the plot. $n = 3$ independent experiments. **e**, Phosphorylation (purple squares) and ubiquitination (red circles) sites on EGFR that are regulated upon SARS-CoV-2 infection. The plot shows median $\log_2$ fold change of abundance compared with mock infection at 24 and 36 hpi. All identified phosphorylation sites have known regulatory function. **f**, Profile plots of time-resolved EGFR K754 ubiquitination, EGFR T693 and S991 phosphorylation, and total EGFR protein levels in A549-ACE2 cells infected with SARS-CoV-2 or SARS-CoV with indicated median (line), 50% (shaded region) and 95% (dotted line) confidence intervals. $n = 3$ (ubiquitination) or 4 (phosphorylation and total protein) independent experiments.

or SARS-CoV, and profiled the effects of viral infection on mRNA expression, protein abundance, ubiquitination and phosphorylation in a time-resolved manner (Fig. 2 a-b).

In line with previous reports[9,22], we found that both SARS-CoV-2 and SARS-CoV can downregulate the type I interferon response and activate a pro-inflammatory signature at transcriptome and proteome levels (Fig. 2a–c, Extended Data Fig. 4c–f, i, Supplementary Tables 4, 8, Supplementary Discussion 1). However, SARS-CoV elicited a more pronounced activation of the NF-κB pathway, correlating with its higher replication rate and potentially explaining the lower severity of pulmonary disease in cases of SARS-CoV-2 infection[23] (Supplementary Tables 4, 5). By contrast, SARS-CoV-2 infection led to higher expression of FN1 and SERPINE1, which may be linked to the specific recruitment of TGF-β factors (Fig. 1b), supporting regulation of TGF-β signalling by SARS-CoV-2.

To better understand the mechanisms underlying perturbation of cellular signalling, we performed comparative ubiquitination and phosphorylation profiling following infection with SARS-CoV-2 or SARS-CoV.

This analysis showed that 1,108 of 16,541 detected ubiquitination sites were differentially regulated by infection with SARS-CoV-2 or SARS-CoV (Fig. 2a, b, d, Extended Data Fig. 5a, Supplementary Table 6). More than half of the significant sites were regulated in a similar manner by both viruses. These included sites on SLC35 and SUMO family proteins, indicating possible regulation of sialic acid transport and the SUMO activity. SARS-CoV-2 specifically increased ubiquitination on autophagy-related factors (MAP1LC3A, GABARAP, VPS33A and VAMP8) as well as specific sites on EGFR (for example, K739, K754 and K970). In some cases, the two viruses targeted distinct sites on the same cellular protein, as exemplified by HSP90 family members (for example, K84, K191 and K539 on HSP90AA1) (Fig. 2d). Notably, a number of proteins (for example, ALCAM, ALDH3B1, CTNNA1, EDF1 and SLC12A2) exhibited concomitant ubiquitination and a decrease at the protein level after infection, pointing to ubiquitination-mediated protein degradation (Fig. 2d, Extended Data Figs. 4f, 5a, Supplementary Tables 5, 6). Among these downregulated proteins, EDF1 has a pivotal role in the maintenance of endothelial integrity and may be a link to endothelial dysfunctions
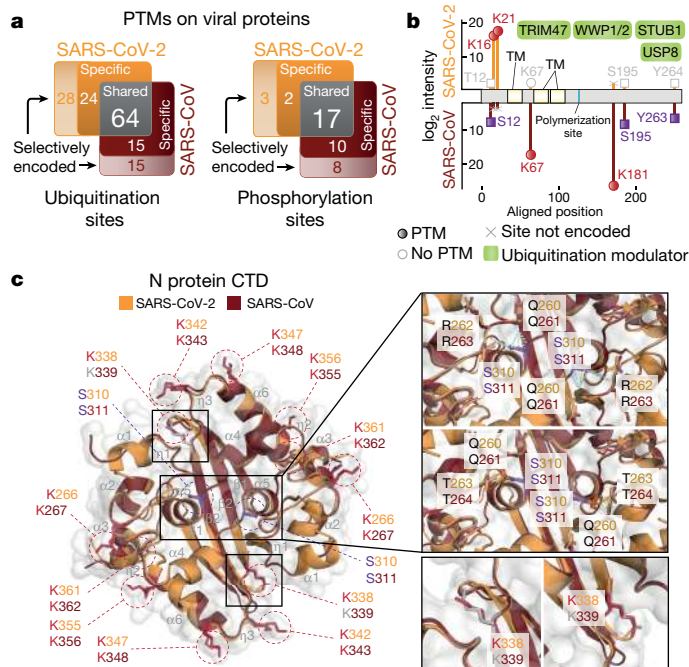
Fig. 3 | Phosphorylation and ubiquitination of SARS-CoV-2 and SARS-CoV viral proteins. a, Distribution of identified shared, differentially regulated and selectively encoded (sequence-specific) ubiquitination and phosphorylation sites on SARS-CoV-2 and SARS-CoV homologous proteins after infection of A549-ACE2 cells. PTM, post-translational modification. **b**, Mapping of the ubiquitination (red circles) and phosphorylation (purple squares) sites on an alignment of SARS-CoV-2 ORF3 and SARS-CoV ORF3a proteins, showing median $\log_2$ intensities in virus-infected A549-ACE2 cells at 24 hpi. Functional (blue) and topological (yellow) domains are mapped on each sequence. Ubiquitin-modifying enzymes binding to ORF3 and ORF3a as identified in our AP–MS experiments (Extended Data Fig. 1b) are indicated (green). TM, transmembrane domain. **c**, Surface and ribbon representation of superimposed SARS-CoV (Protein Data Bank (PDB) ID: 2CJR, brown) and SARS-CoV-2 (PDB ID: 6YUN, orange) N protein CTD dimers (r.m.s.d. values of 0.492 Å for 108 matching Cα atoms). Secondary structures are numbered in grey (prefixed with α for α-helix, β for β-strand and η for non-structured regions). Side chains are colour coded depending on whether they are in ubiquitinated (red), phosphorylated (purple) or unmodified (grey) sites. The K338 ubiquitination site unique to SARS-CoV-2 is shown as a close-up for both monomers (bottom right). Close-ups of inter-chain residue interactions established by non-phosphorylated (top right) and phosphorylated (middle right) SARS-CoV-2 S310 or SARS-CoV S311.

described for COVID-19[24,25]. Profound regulation of cellular signalling pathways was also observed at the phosphoproteomic level: among 16,399 total quantified phosphorylation sites, 4,643 showed significant changes after infection with SARS-CoV-2 or SARS-CoV (Extended Data Fig. 5b, c, Supplementary Table 7). Highly regulated sites were identified for the proteins of the MAPK pathways (for example, MAPKAPK2, MAP2K1, JUN and SRC), and proteins involved in autophagy signalling (for example, DEPTOR, RICTOR, OPTN, SQSTM1 and LAMTOR1) and viral entry (for example, ACE2 and RAB7A) (Extended Data Fig. 5b, d). Notably, RAB7A was recently shown to be an important host factor for SARS-CoV-2 infection that assists endosomal trafficking of ACE2 to the plasma membrane[26]. We observed higher phosphorylation at S72 of RAB7A in SARS-CoV-2 infection compared with SARS-CoV or mock infection; this site is implicated in RAB7A intracellular localization and molecular association[27]. The regulation of known phosphorylation sites suggests an involvement of central kinases (cyclin dependent kinases, AKT, MAPKs, ATM, and CHEK1) linked to cell survival, cell cycle progression, cell growth and motility, stress responses and the DNA

damage response; this was also supported by the analysis of enriched motifs (Extended Data Fig. 5e, f, Supplementary Tables 7, 8). Notably, SARS-CoV-2 infection, but not SARS-CoV infection, led to phosphorylation of the antiviral kinase EIF2AK2 (also known as PKR) at the critical regulatory residue S33[28]. This differential activation of EIF2AK2 could contribute to the difference in the growth kinetics of the two viruses (Supplementary Table 4, 5).
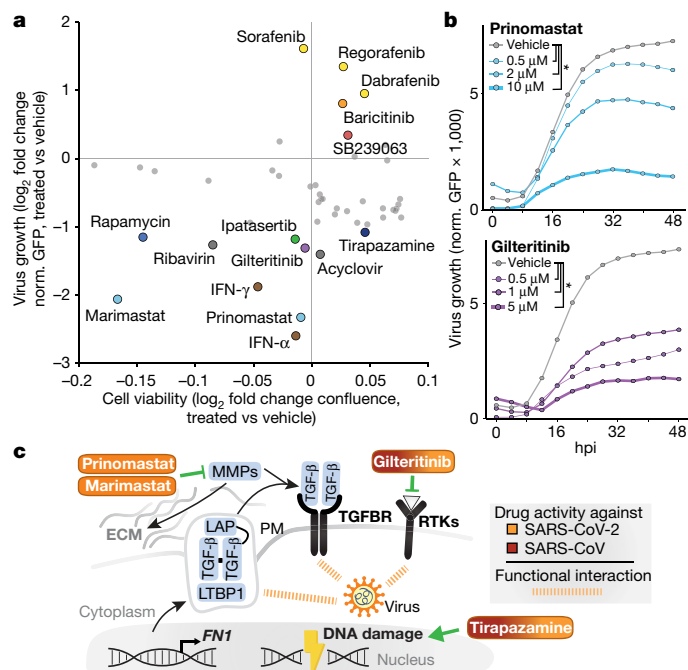
Our data clearly point to an interplay of phosphorylation and ubiquitination patterns on individual host proteins. For instance, EGFR showed increased ubiquitination on 6 lysine residues at 24 h post-infection (hpi) accompanied by increased phosphorylation of T693, S695 and S991 at 24 and 36 hpi (Fig. 2e, f). Ubiquitination of all six lysine residues on EGFR was more pronounced following infection with SARS-CoV-2 than with SARS-CoV. Moreover, vimentin, a central co-factor for coronavirus entry[29] and pathogenicity[30,31], displayed distinct phosphorylation and ubiquitination patterns on several sites early (for example, S420) or late (for example, S56, S72 and K334) in infection (Extended Data Fig. 6a, b). These findings underscore the value of testing different post-translational modifications simultaneously and suggest a concerted engagement of regulatory machineries to modify target protein functions and abundance.

Post-translational modification of viral proteins

The majority of viral proteins were also post-translationally modified. Of the 27 detected SARS coronavirus proteins, 21 were ubiquitinated. Nucleocapsid (N), spike (S), NSP2 and NSP3 were the most heavily modified proteins in both viruses (Extended Data Fig. 6c, Supplementary Table 6). Many ubiquitination sites were common to both viruses. Around half of the sites that were exclusively ubiquitinated in either virus were conserved between SARS-CoV and SARS-CoV-2. The remaining specifically regulated ubiquitination sites were unique to each virus, indicating that these acquired adaptations can be post-translationally modified and may recruit cellular proteins with distinct functions (Fig. 3a). Our interactome data identified several host E3 ligases (for example, we identified interactions between SARS-CoV-2 ORF3 and TRIM47, WWP1, WWP2 and STUB1; and between SARS-CoV-2 M and TRIM7) and deubiquitinating enzymes (for example, interactions between SARS-CoV-2 ORF3 and USP8; SARS-CoV-2 ORF7a and USP34; and SARS-CoV N and USP9X), suggesting crosstalk between ubiquitination and viral protein functions (Fig. 1b, Extended Data Fig. 6d, Supplementary Table 2). Of particular interest are extensive ubiquitination events on the S protein of both viruses (K97, K528, K825, K835, K921 and K947), which are distributed on functional domains (N-terminal domain, C-terminal domain (CTD), fusion peptide and heptad repeat 1 domain), potentially indicating critical regulatory functions that are conserved between the two viruses (Extended Data Fig. 6e). We observed phosphorylation of 5 SARS-CoV-2 proteins (M, N, S, NSP3 and ORF9b) and 8 SARS-CoV proteins (M, N, S, NSP1, NSP2, NSP3, ORF3 and ORF9b) (Extended Data Fig. 6f, Supplementary Table 7), on sites corresponding to known recognition motifs. In particular, CAMK4 and MAPKAPK2 potentially phosphorylate sites on S and N, respectively. Phosphorylation of cellular proteins suggested that the activities of these kinases were enriched in cells infected with SARS-CoV-2 or SARS-CoV (Extended Data Figs. 5e, f, 6e, g). Moreover, N proteins of both SARS-CoV-2 and SARS-CoV recruit GSK3, which could potentially be linked to phosphorylation events on these viral proteins (Fig. 1b, Extended Data Fig. 6g, Supplementary Table 7). Notably, we identified novel post-translationally modified sites located at functional domains of viral proteins; we detected ubiquitination at SARS-CoV-2 N K338 and phosphorylation on SARS-CoV-2 and SARS-CoV N S310 and S311 (Extended Data Fig. 6g). Mapping these sites to the atomic structure of the CTD[32,33] highlights critical positions for the function of the protein (Fig. 3c, Extended Data Fig. 6h, Supplementary Discussion 2). Collectively, while the identification of differentially regulated sites

Fig. 4 | A network diffusion approach enables identification of molecular pathways linking protein–protein interactions with downstream changes in the host proteome. a, The network diffusion approach to identify functional connections between the host targets of a viral protein and downstream proteome changes. The results of network diffusion are integrated with omics datasets of SARS coronavirus infection to streamline the identification of affected host pathways. **b**, Subnetworks of the network diffusion linking host targets of SARS-CoV-2 ORF3 to factors involved in autophagy. The thickness of directed edges is proportional to the random-walk transition probability. Black edges denote connections present in ReactomeFI. **c**, Overview of perturbations to host-cell autophagy induced by SARS-CoV-2. The pathway regulation is derived from the network diffusion model for SARS-CoV-2 ORF3 and NSP6 and is overlaid with the changes in protein levels, ubiquitination (Ubi) and phosphorylation (Pho) induced by SARS-CoV-2 infection. FC, fold change; PM, plasma membrane.

may indicate pathogen-specific functions, insights from conserved post-translational modifications may also provide useful knowledge for the development of targeted pan-antiviral therapies.

Viral perturbation of key cellular pathways

Our unified experimental design in a syngeneic system enabled direct time-resolved comparison of SARS-CoV-2 and SARS-CoV infection across different levels. Integrative pathway-enrichment analysis demonstrated that both viruses largely perturb the same cellular processes at multiple levels, albeit with distinct temporal patterns (Extended Data Fig. 7a). For instance, transcriptional downregulation of proteins involved in tau protein kinase activity and Fe ion sequestration at 6 hpi was followed by a decrease in protein abundance after 12 hpi

(Supplementary Table 8). RHO GTPase activation, mRNA processing and the role of ABL in ROBO–SLIT signalling appeared to be regulated mostly through phosphorylation (Extended Data Fig. 7a). By contrast, processes connected to cellular integrity such as the formation of senescence-associated heterochromatin foci, apoptosis-induced DNA fragmentation and amino acid transport across the plasma membrane were modulated through concomitant phosphorylation and ubiquitination events, suggesting molecular links between these post-translational modifications. Ion transporters, especially the SLC12 family of cation-coupled chloride cotransporters—previously identified as cellular factors in pulmonary inflammation[34]—were also regulated at multiple levels, evidenced by reduced protein abundance as well as differential post-translational modifications (Extended Data Fig. 7a).

The pathway-enrichment analysis provided a global and comprehensive picture of how SARS-CoV-2 and SARS-CoV affect the host. We next applied an automated approach to systematically explore the underlying molecular mechanisms contained in the viral interactome and effectome data. We mapped the measured interactions and effects of each viral protein onto the global network of cellular interactions[35] and applied a network diffusion approach[36] (Fig. 4a). This type of analysis uses known cellular protein–protein interactions, signalling and regulation events to identify connection points between cellular proteins that interact with viral proteins and the proteins affected by the expression of these viral proteins (Extended Data Figs. 1b, 2d, Supplementary Tables 2, 3). The connections inferred from the real data were significantly shorter than for randomized data, validating the relevance of the approach and the quality of the data (Extended Data Fig. 8a, b). The findings from this approach include the potential mechanisms by which ORF3 and NSP6 may regulate autophagy, the modulation of innate immunity by M, ORF3 and ORF7b, and the perturbation of integrin–TGF-β–EGFR–receptor tyrosine kinase signalling by ORF8 of SARS-CoV-2 (Fig. 4b, Extended Data Fig. 8c, d). Enriching these subnetworks with data on SARS-CoV-2 infection-dependent mRNA abundance, protein abundance, phosphorylation and ubiquitination (Fig. 4a) provided insights into the regulatory mechanisms activated by SARS-CoV-2. For instance, the analysis confirmed a role of NSP6 in both SARS-CoV-2- and SARS-CoV-induced autophagy[37] and revealed the SARS-CoV-2 specific inhibition of autophagic flux by ORF3 protein, leading to the accumulation of autophagy receptors (SQSTM1, GABARAPL2, NBR1, CALCOCO2, MAP1LC3A, MAP1LC3B and TAX1BP1), consistent with the accumulation of MAP1LC3B protein observed in cells infected with SARS-CoV-2 (Fig. 4c, Extended Data Fig. 8e, f). This inhibition may be a result of the interaction of the ORF3 protein with the HOPS complex (VPS11, VPS16, VPS18, VPS39 and VPS41), which is essential for autophagosome–lysosome fusion, as well as the differential phosphorylation of regulatory sites (for example, on TSC2, mTORC1 complex, ULK1, RPS6 and SQSTM1) and ubiquitination of key components (MAP1LC3A, GABARAPL2, VPS33A and VAMP8) (Fig. 4c, Extended Data Fig. 8g). This inhibition of autophagosome function may have direct consequences for protein degradation. The abundance of APOB, a protein that is degraded via autophagy[38], was selectively increased after SARS-CoV-2 infection or expression of SARS-CoV-2 ORF3 (Extended Data Fig. 3b, 8h). Accumulating APOB levels could increase the risk of arterial thrombosis[39], one of the main complications contributing to lung, heart and kidney failure in patients with COVID-19[40]. The inhibition of the interferon (IFN)-α and IFN-β response observed at transcriptional and proteome levels was similarly explained by the network diffusion analysis (Extended Data Fig. 8i), which implicated multiple proteins of SARS-CoV-2 in the disruption of antiviral immunity. Further experiments functionally corroborated the inhibition of IFN-α and IFN-β induction or signalling by ORF3, ORF6, ORF7a, ORF7b and ORF9b (Extended Data Fig. 8j). Upon virus infection, we observed the regulation of TGF-β and EGFR pathways modulating cell survival, motility and innate immune responses (Extended Data Fig. 9a–d). Specifically, our network diffusion analysis revealed a connection between

Fig. 5 | Identification of pathways targeted by SARS-CoV-2 using a multi-omics profiling approach enables systematic testing of candidate antiviral therapies. a, b, A549-ACE2 cells were treated with the indicated drugs 6 h before infection with SARS-CoV-2-GFP (MOI of 3). Changes in cell viability and virus growth (**a**) in drug-treated cells compared with untreated A549-ACE2 cells at 48 hpi. A confluence cut-off of −0.2 log₂ fold change was applied to remove cytotoxic compounds. **b**, Time courses of virus replication after pre-treatment of cells with prinomastat or gilteritinib. $n = 4$ independent experiments; $^*P \leq 0.01$ compared with control treatment, unadjusted two-sided Wilcoxon test. Norm., normalized. **c**, Drugs potentially targeting pathways identified in our study. Colour indicates antiviral activity against SARS-CoV-2 or SARS-CoV (brown–orange gradient) or SARS-CoV-2 specifically (orange), as inferred from in vitro experiments.

the binding of the ORF8 and ORF3 proteins to TGF-β-associated factors (TGF-β1, TGF-β2, LTBP1, TGFBR2, FURIN and BAMBI), the differential expression of extracellular matrix regulators (FERMT2 and CDH1) and the virus-induced upregulation of fibrinogens (FGA, FGB), fibronectin (FN1) and SERPINE1[41] (Extended Data Fig. 9a, b). The increased phosphorylation of proteins involved in MAPK signalling (for example, SHC1 (on S139), SOS1 (S1134/S1229), JUN (S63/S73), MAPKAPK2 (T334) and p38 (T180/Y182)) and receptor tyrosine kinase signalling (for example, phosphorylation of PI3K complex members PDPK1 (S241) and RPS6KA1 (S380)) as well as a higher expression of JUN, FOS and EGR1 are further indications of regulation of TGF-β and EGFR pathways (Extended Data Fig. 9a, c, d). In turn, TGF-β and EGFR signalling are known to be potentiated by integrin signalling and activation of YAP-dependent transcription[42], which we observed to be regulated in a time-dependent manner upon SARS-CoV-2 infection (Extended Data Fig. 9a). As well as promoting virus replication, activation of these pathways has been implicated in fibrosis[13–15], one of the hallmarks of COVID-19[16].

Data-guided drug identification and testing

Together, the viral protein–host protein interactions and regulation of pathways observed at multiple levels identify potential points for targeting SARS-CoV and SARS-CoV-2 using well-characterized selective antiviral drugs. To test antiviral efficacy, we used time-lapse fluorescent microscopy of infection with a GFP reporter SARS-CoV-2[43]. Inhibition of virus replication by treatment with IFN-α corroborated previous conclusions that efficient SARS-CoV-2 replication involves the inactivation

of this pathway at an early step[9,44] and confirmed the reliability of this screening approach (Extended Data Fig. 10a). We tested a panel of 48 drugs that modulate the pathways perturbed by the virus for their effects on SARS-CoV-2 replication (Fig. 5a, Supplementary Table 9). Of note, inhibitors of B-RAF (sorafenib, regorafenib and dabrafenib), JAK1/2 (baricitinib) and MAPK (SB239063), which are commonly used to treat cancer and autoimmune diseases[45–47], significantly increased virus growth in an in vitro model of infection (Fig. 5a, Extended Data Fig. 10b, Supplementary Table 9). By contrast, inducers of DNA damage (tirapazamine and rabusertib) or an mTOR inhibitor (rapamycin) suppressed virus growth. The highest antiviral activity was observed for gilteritinib (a designated inhibitor of FLT3 and AXL), ipatasertib (an AKT inhibitor), prinomastat and marimastat (matrix metalloprotease (MMP) inhibitors) (Fig. 5a, b, Extended Data Fig. 10c, Supplementary Table 9). These compounds profoundly inhibited replication of SARS-CoV-2 while having no effects or minor effects on cell growth (Extended Data Fig. 10b, Supplementary Table 9). Quantitative PCR analysis indicated antiviral activities for gilteritinib and tirapazamine against SARS-CoV-2 and SARS-CoV (Fig. 5c, Extended Data Fig. 10d, e). Notably, prinomastat and marimastat, specific inhibitors of MMP2 and MMP9, showed selective activity against SARS-CoV-2 but not against SARS-CoV (Fig. 5c, Extended Data Fig. 10f, g). Activities of MMPs have been linked to TGF-β activation and pleural effusions, alveolar damage and neuroinflammation (for example, Kawasaki disease), all of which are characteristic of COVID-19[23,48–51].

This drug screen demonstrates the value of our combined dataset, which profiles SARS-CoV-2 infection at multiple levels. We expect that further exploration of these rich data by the scientific community and additional studies of the interplay between different omics levels will substantially advance our molecular understanding of coronavirus biology, including the pathogenicity associated with specific human coronaviruses, such as SARS-CoV-2 and SARS-CoV. Moreover, this resource, together with complementary approaches[26,52–54], will streamline the search for antiviral compounds and serve as a base for rational design of combination therapies that target the virus from multiple synergistic angles, thus potentiating the effect of individual drugs while minimizing potential side effects on healthy tissues.

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/s41586-021-03493-4.

1. Gordon, D. E. et al. A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature* **583**, 459–468 (2020).
2. Gordon, D. E. et al. Comparative host–coronavirus protein interaction networks reveal pan-viral disease mechanisms. *Science* **370**, eabe9403 (2020).
3. Bojkova, D. et al. Proteomics of SARS-CoV-2-infected host cells reveals therapy targets. *Nature* **583**, 469–472 (2020).
4. Messner, C. B. et al. Ultra-high-throughput clinical proteomics reveals classifiers of COVID-19 infection. *Cell Syst.* **11**, 11–24.e4 (2020).
5. Samavarchi-Tehrani, P. et al. A SARS-CoV-2–host proximity interactome. Preprint at https://doi.org/10.1101/2020.09.03.282103 (2020).
6. Laurent, E. M. N. et al. Global BioID-based SARS-CoV-2 proteins proximal interactome unveils novel ties between viral polypeptides and host factors involved in multiple COVID19-associated mechanisms. Preprint at https://doi.org/10.1101/2020.08.28.272955 (2020).
7. Klann, K. et al. Growth factor receptor signaling inhibition prevents SARS-CoV-2 replication. *Mol. Cell* **80**, 164–174.e4 (2020).
8. Huang, J. et al. SARS-CoV-2 infection of pluripotent stem cell-derived human lung alveolar type 2 cells elicits a rapid epithelial-intrinsic inflammatory response. *Cell Stem Cell* **27**, 962–973.e7 (2020).
9. Blanco-Melo, D. et al. Imbalanced host response to SARS-CoV-2 drives development of COVID-19. *Cell* **181**, 1036–1045.e9 (2020).
10. Li, J. et al. Virus–host interactome and proteomic survey reveal potential virulence factors influencing SARS-CoV-2 pathogenesis. *Med* **2**, 99–112.e7 (2021).
11. von Brunn, A. et al. Analysis of intraviral protein–protein interactions of the SARS coronavirus ORFeome. *PLoS ONE* **2**, e459 (2007).

Article

12. Cornillez-Ty, C. T., Liao, L., Yates, J. R., III, Kuhn, P. & Buchmeier, M. J. Severe acute respiratory syndrome coronavirus nonstructural protein 2 interacts with a host protein complex involved in mitochondrial biogenesis and intracellular signaling. *J. Virol.* **83**, 10314–10318 (2009).

13. Andrianifahanana, M. et al. ERBB receptor activation is required for profibrotic responses to transforming growth factor β. *Cancer Res.* **70**, 7421–7430 (2010).

14. Pittet, J.-F. et al. TGF-β is a critical mediator of acute lung injury. *J. Clin. Invest.* **107**, 1537–1544 (2001).

15. George, P. M., Wells, A. U. & Jenkins, R. G. Pulmonary fibrosis and COVID-19: the potential role for antifibrotic therapy. *Lancet Respir. Med.* **8**, 807–815 (2020).

16. Mo, X. et al. Abnormal pulmonary function in COVID-19 patients at time of hospital discharge. *Eur. Respir. J.* **55**, 2001217 (2020).

17. Heo, J.-M. et al. Integrated proteogenetic analysis reveals the landscape of a mitochondrial-autophagosome synapse during PARK2-dependent mitophagy. *Sci. Adv.* **5**, eaay4624 (2019).

18. Shi, C.-S. et al. SARS-coronavirus open reading frame-9b suppresses innate immunity by targeting mitochondria and the MAVS/TRAF3/TRAF6 signalosome. *J. Immunol.* **193**, 3080–3089 (2014).

19. Hoagland, D. A. et al. Modulating the transcriptional landscape of SARS-CoV-2 as an effective method for developing antiviral compounds. Preprint at https://doi.org/10.1101/2020.07.12.199687 (2020).

20. Castiglione, V., Chiriacò, M., Emdin, M., Taddei, S. & Vergaro, G. Statin therapy in COVID-19 infection. *Eur. Heart J. Cardiovasc. Pharmacother.* **6**, 258–259 (2020).

21. Radenkovic, D., Chawla, S., Pirro, M., Sahebkar, A. & Banach, M. Cholesterol in relation to COVID-19: should we care about it? *J. Clin. Med.* **9**, 1909 (2020).

22. Chu, H. et al. Comparative replication and immune activation profiles of SARS-CoV-2 and SARS-CoV in human lungs: an ex vivo study with implications for the pathogenesis of COVID-19. *Clin. Infect. Dis.* **71**, 1400–1409 (2020).

23. Zhu, Z. et al. From SARS and MERS to COVID-19: a brief summary and comparison of severe acute respiratory infections caused by three highly pathogenic human coronaviruses. *Respir. Res.* **21**, 224 (2020).

24. Cazzaniga, A., Locatelli, L., Castiglioni, S. & Maier, J. The contribution of EDF1 to PPARγ transcriptional activation in VEGF-treated human endothelial cells. *Int. J. Mol. Sci.* **19**, 1830 (2018).

25. Gavriilaki, E. et al. Endothelial dysfunction in COVID-19: lessons learned from coronaviruses. *Curr. Hypertens. Rep.* **22**, 63 (2020).

26. Daniloski, Z. et al. Identification of required host factors for SARS-CoV-2 infection in human cells. *Cell* **184**, 92–105 (2021).

27. Shinde, S. R. & Maddika, S. PTEN modulates EGFR late endocytic trafficking and degradation by dephosphorylating Rab7. *Nat. Commun.* **7**, 10689 (2016).

28. Wang, D. et al. Auto-phosphorylation represses protein kinase R activity. *Sci. Rep.* **7**, 44340 (2017).

29. Yu, Y. T.-C. et al. Surface vimentin is critical for the cell entry of SARS-CoV. *J. Biomed. Sci.* **23**, 14 (2016).

30. dos Santos, G. et al. Vimentin regulates activation of the NLRP3 inflammasome. *Nat. Commun.* **6**, 6574 (2015).

31. Ramos, I., Stamatakis, K., Oeste, C. L. & Pérez-Sala, D. Vimentin as a multifaceted player and potential therapeutic target in viral infections. *Int. J. Mol. Sci.* **21**, 4675 (2020).

32. Zinzula, L. et al. High-resolution structure and biophysical characterization of the nucleocapsid phosphoprotein dimerization domain from the Covid-19 severe acute respiratory syndrome coronavirus 2. *Biochem. Biophys. Res. Commun.* **538**, 54–62 (2021).

33. Chen, C.-Y. et al. Structure of the SARS coronavirus nucleocapsid protein RNA-binding dimerization domain suggests a mechanism for helical packaging of viral RNA. *J. Mol. Biol.* **368**, 1075–1086 (2007).

34. Perry, J. S. A. et al. Interpreting an apoptotic corpse as anti-inflammatory involves a chloride sensing pathway. *Nat. Cell Biol.* **21**, 1532–1543 (2019).

35. Wu, G., Dawson, E., Duong, A., Haw, R. & Stein, L. ReactomeFIViz: a Cytoscape app for pathway and network-based data analysis. *F1000Res.* **3**, 146 (2014).

36. Reyna, M. A., Leiserson, M. D. M. & Raphael, B. J. Hierarchical HotNet: identifying hierarchies of altered subnetworks. *Bioinformatics* **34**, i972–i980 (2018).

37. Cottam, E. M., Whelband, M. C. & Wileman, T. Coronavirus NSP6 restricts autophagosome expansion. *Autophagy* **10**, 1426–1441 (2014).

38. Ohsaki, Y., Cheng, J., Fujita, A., Tokumoto, T. & Fujimoto, T. Cytoplasmic lipid droplets are sites of convergence of proteasomal and autophagic degradation of apolipoprotein B. *Mol. Biol. Cell* **17**, 2674–2683 (2006).

39. Khalil, M. F., Wagner, W. D. & Goldberg, I. J. Molecular interactions leading to lipoprotein retention and the initiation of atherosclerosis. *Arterioscler. Thromb. Vasc. Biol.* **24**, 2211–2218 (2004).

40. Nicolai, L. et al. Immunothrombotic dysregulation in COVID-19 pneumonia is associated with respiratory failure and coagulopathy. *Circulation* **142**, 1176–1189 (2020).

41. Zavadil, J. et al. Genetic programs of epithelial cell plasticity directed by transforming growth factor-β. *Proc. Natl Acad. Sci. USA* **98**, 6686–6691 (2001).

42. Qin, Z., Xia, W., Fisher, G. J., Voorhees, J. J. & Quan, T. YAP/TAZ regulates TGF-β/Smad3 signaling by induction of Smad7 via AP-1 in human skin dermal fibroblasts. *Cell Commun. Signal.* **16**, 18 (2018).

43. Thi Nhu Thao, T. et al. Rapid reconstruction of SARS-CoV-2 using a synthetic genomics platform. *Nature* **582**, 561–565 (2020).

44. Mantlo, E., Bukreyeva, N., Maruyama, J., Paessler, S. & Huang, C. Antiviral activities of type I interferons to SARS-CoV-2 infection. *Antiviral Res.* **179**, 104811 (2020).

45. Seebacher, N. A., Stacy, A. E., Porter, G. M. & Merlot, A. M. Clinical development of targeted and immune based anti-cancer therapies. *J. Exp. Clin. Cancer Res.* **38**, 156 (2019).

46. O'Shea, J. J., Kontzias, A., Yamaoka, K., Tanaka, Y. & Laurence, A. Janus kinase inhibitors in autoimmune diseases. *Ann. Rheum. Dis.* **72** (Suppl 2), ii111–ii115 (2013).

47. Yong, H.-Y., Koh, M.-S. & Moon, A. The p38 MAPK inhibitors for the treatment of inflammatory diseases and cancer. *Expert Opin. Investig. Drugs* **18**, 1893–1905 (2009).

48. Hsieh, W.-Y. et al. ACE/ACE2 ratio and MMP-9 activity as potential biomarkers in tuberculous pleural effusions. *Int. J. Biol. Sci.* **8**, 1197–1205 (2012).

49. Ueland, T. et al. Distinct and early increase in circulating MMP-9 in COVID-19 patients with respiratory failure. *J. Infect.* **81**, e41–e43 (2020).

50. Villalta, P. C., Rocic, P. & Townsley, M. I. Role of MMP2 and MMP9 in TRPV4-induced lung injury. *Am. J. Physiol. Lung Cell. Mol. Physiol.* **307**, L652–L659 (2014).

51. Marten, N. W. & Zhou, J. in *Experimental Models of Multiple Sclerosis* (eds Lavi, E. & Constantinescu, C. S.) 839–848 (Springer, 2005).

52. Zhang, J.-Y. et al. Single-cell landscape of immunological responses in patients with COVID-19. *Nat. Immunol.* **21**, 1107–1118 (2020).

53. Ellinghaus, D. et al. Genomewide association study of severe Covid-19 with respiratory failure. *N. Engl. J. Med.* **383**, 1522–1534 (2020).

54. Ali, A. & Vijayan, R. Dynamics of the ACE2–SARS-CoV-2/SARS-CoV spike protein interface reveal unique mechanisms. *Sci. Rep.* **10**, 14214 (2020).

Methods

No statistical methods were used to predetermine sample size. The experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessment.

Cell lines and reagents

HEK293T, A549, Vero E6 and HEK293-R1 cells were authentified by PCR-single-locus-technology and their respective culturing conditions were described previously[55]. All cell lines were tested to be mycoplasma free. Expression constructs for C-terminal HA-tagged viral open reading frames were synthesized (Twist Bioscience and BioCat) and cloned into pWPI vector as described previously[56] with the following modifications: a starting ATG codon was added, internal canonical splicing sites were replaced with synonymous mutations and a C-terminal HA-tag, followed by an amber stop codon, was added to individual viral open reading frames. A C-terminally HA-tagged ACE2 sequence was amplified from an ACE2 expression vector (provided by S. Pöhlmann)[57] into the lentiviral vector pWPI-puro. A549 cells were transduced twice, and A549-ACE2 cells were selected with puromycin. Lentivirus production, transduction of cells and antibiotic selection were performed as described previously[52]. RNA isolation (Macherey–Nagel NucleoSpin RNA plus), reverse transcription (TaKaRa Bio PrimeScript RT with gDNA eraser) and quantitative PCR with reverse transcription (RT–qPCR) (Thermo-Fisher Scientific PowerUp SYBR green) were performed as described previously[54]. RNA isolation for next generation sequencing applications was performed according to the manufacturer's protocol (Qiagen RNeasy mini kit, RNase free DNase set). For detection of protein abundance by western blotting, HA–horseradish peroxidase (HRP) (Sigma-Aldrich; H6533; 1:2,500 dilution), ACTB–HRP (Santa Cruz; sc-47778; 1:5,000 dilution), MAP1LC3B (Cell Signaling; 3868; 1:1,000 dilution), MAVS (Cell Signaling; 3993; 1:1,000 dilution), HSPA1A (Cell Signaling; 4873; 1:1,000 dilution), TGF-β (Cell Signaling; 3711; 1:1,000 dilution), phospho-p38 (T180/Y182) (Cell Signaling; 4511; 1:1,000 dilution), p38 (Cell Signaling; 8690; 1:1,000 dilution) and SARS-CoV-2 or SARS-CoV N protein (Sino Biological; 40143-MM05; 1:1,000 dilution) antibodies were used. Secondary antibodies detecting mouse (Cell Signaling; 7076; 1:5,000 dilution; Jackson ImmunoResearch; 115-035-003; 1:5,000 dilution), rat (Invitrogen; 31470; 1:5,000 dilution), and rabbit IgG (Cell Signaling; 7074; 1:5,000 dilution) were coupled to HRP. For AP–MS and affinity purification–western blotting applications, HA beads (Sigma-Aldrich and Thermo Fisher Scientific) and Streptactin II beads (IBA Lifesciences) were used. Imaging of western blots was performed as described[58]. Recombinant human IFN-α used for stimulation of cells in the reporter assay was a gift from P. Stäheli (Institute of Virology, University of Freiburg), recombinant human IFN-γ was purchased from PeproTech, and IVT4 was produced as described before[59]. All compounds tested in the viral inhibitor assay are listed in Supplementary Table 9.

Virus strains, stock preparation, plaque assay and in vitro infection

SARS-CoV-Frankfurt-1, SARS-CoV-2-MUC-IMB-1 and SARS-CoV-2-GFP strains[43] were produced by infecting Vero E6 cells cultured in DMEM medium (10% FCS, 100 μg ml⁻¹ Streptomycin, 100 IU ml⁻¹ penicillin) for 2 days (MOI of 0.01). Viral stock was collected and spun twice (1,000g for 10 min) before storage at −80 °C. Titre of viral stock was determined by plaque assay. Confluent monolayers of Vero E6 cells were infected with serial fivefold dilutions of virus supernatants for 1 h at 37 °C. The inoculum was removed and replaced with serum-free MEM (Gibco, Life Technologies) containing 0.5% carboxymethylcellulose (Sigma-Aldrich). Two days after infection, cells were fixed for 20 min at room temperature with formaldehyde added directly to the medium to a final concentration of 5%. Fixed cells were washed extensively with PBS before staining with water containing 1% crystal violet and 10% ethanol for 20 min. After rinsing with PBS, the number of plaques was counted and the virus titre was calculated.

A549-ACE2 cells were infected with either SARS-CoV-Frankfurt-1 or SARS-CoV-2-MUC-IMB-1 strains (MOI of 2) for the subsequent experiments. At each time point, the samples were washed once with 1× TBS buffer and collected in sodium deoxycholate (SDC) lysis buffer (100 mM Tris HCl pH 8.5; 4% SDC) for proteome-phosphoproteome-ubiquitinome analysis, sodium dodecyl sulfate (SDS) lysis buffer (62.5 mM Tris HCl pH 6.8; 2% SDS; 10% glycerol; 50 mM DTT; 0.01% bromophenol blue) for western blot, or buffer RLT (Qiagen) for transcriptome analysis. The samples were heat-inactivated and frozen at −80 °C until further processing.

Affinity purification and mass spectrometric analyses of SARS-CoV-2, SARS-CoV, HCoV-229E and HCoV-NL63 proteins expressed in A549 cells

To determine the interactomes of SARS-CoV-2 and SARS-CoV and the interactomes of an accessory protein (encoded by ORF4 or ORF4a of HCoV-229E or ORF3 of HCoV-NL63) that presumably represents a homologue of the ORF3 and ORF3a proteins of SARS-CoV-2 and SARS-CoV, respectively, four replicate affinity purifications were performed for each HA-tagged viral protein. A549 cells (6 × 10⁶ cells per 15-cm dish) were transduced with lentiviral vectors encoding HA-tagged SARS-CoV-2, SARS-CoV or HCoV-229E/NL63 proteins and protein lysates were prepared from cells collected 3 days after transduction. Cell pellets from two 15-cm dishes were lysed in lysis buffer (50 mM Tris-HCl pH 7.5, 100 mM NaCl, 1.5 mM MgCl₂, 0.2% (v/v) NP-40, 5% (v/v) glycerol, cOmplete protease inhibitor cocktail (Roche), 0.5% (v/v) 750 U/μl SmDNase) and sonicated (5 min, 4 °C, 30 s on, 30 s off, low settings; Bioruptor, Diagenode). Following normalization of protein concentrations of cleared lysates, virus protein-bound host proteins were enriched by adding 50 μl anti-HA-agarose slurry (Sigma-Aldrich, A2095) with constant agitation for 3 h at 4 °C. Non-specifically bound proteins were removed by four subsequent washes with lysis buffer followed by three detergent-removal steps with washing buffer (50 mM Tris-HCl pH 7.5, 100 mM NaCl, 1.5 mM MgCl₂, 5% (v/v) glycerol). Enriched proteins were denatured, reduced, alkylated and digested by addition of 200 μl digestion buffer (0.6 M guanidinium chloride, 1 mM tris(2-carboxyethyl)phosphine (TCEP), 4 mM chloroacetamide (CAA), 100 mM Tris-HCl pH 8, 0.5 μg LysC (WAKO Chemicals) and 0.5 μg trypsin (Promega) at 30 °C overnight. Peptide purification on StageTips with three layers of C18 Empore filter discs (3M) and subsequent mass spectrometry analysis was performed as described previously[55,56]. In brief, purified peptides were loaded onto a 20-cm reverse-phase analytical column (75 μm diameter; ReproSil-Pur C18-AQ 1.9 μm resin; Dr Maisch) and separated using an EASY-nLC 1200 system (Thermo Fisher Scientific). A binary buffer system consisting of buffer A (0.1% formic acid (FA) in H₂O) and buffer B (80% acetonitrile (ACN), 0.1% FA in H₂O) with a 90-min gradient (5–30% buffer B (65 min), 30–95% buffer B (10 min), wash out at 95% buffer B (5 min), decreased to 5% buffer B (5 min), and 5% buffer B (5 min)) was used at a flow rate of 300 nl per min. Eluting peptides were directly analysed on a Q-Exactive HF mass spectrometer (Thermo Fisher Scientific). Data-dependent acquisition included repeating cycles of one MS1 full scan (300–1650 m/z, $R = 60,000$ at 200 m/z) at an ion target of 3 × 10⁶, followed by 15 MS2 scans of the highest abundant isolated and higher-energy collisional dissociation (HCD) fragmented peptide precursors ($R = 15,000$ at 200 m/z). For MS2 scans, collection of isolated peptide precursors was limited by an ion target of 1 × 10⁵ and a maximum injection time of 25 ms. Isolation and fragmentation of the same peptide precursor was eliminated by dynamic exclusion for 20 s. The isolation window of the quadrupole was set to 1.4 m/z and HCD was set to a normalized collision energy of 27%.

Proteome analyses of cells expressing SARS-CoV-2, SARS-CoV, HCoV-229E or HCoV-NL63 proteins

For the determination of proteome changes in A549 cells expressing SARS-CoV-2, SARS-CoV, HCoV-229E or HCoV-NL63 proteins, a fraction of

Article

1 × 10^6 lentivirus-transduced cells from the affinity purification samples were lysed in guanidinium chloride buffer (6 M guanidinium chloride, 10 mM TCEP, 40 mM CAA, 100 mM Tris-HCl pH 8), boiled at 95 °C for 8 min and sonicated (10 min, 4 °C, 30 s on, 30 s off, high settings). Protein concentrations of cleared lysates were normalized to 50 µg, and proteins were pre-digested with 1 µg LysC at 37 °C for 1 h followed by a 1:10 dilution (100 mM Tris-HCl pH 8) and overnight digestion with 1 µg trypsin at 30 °C. Peptide purification on StageTips with three layers of C18 Empore filter discs (3M) and subsequent mass spectrometry analysis was performed as described previously[55,56]. In brief, 300 ng of purified peptides were loaded onto a 50-cm reversed-phase column (75 µm inner diameter, packed in house with ReproSil-Pur C18-AQ 1.9 µm resin (Dr Maisch)). The column temperature was maintained at 60 °C using a homemade column oven. A binary buffer system, consisting of buffer A (0.1% FA) and buffer B (80% ACN, 0.1% FA), was used for peptide separation, at a flow rate of 300 nl min⁻¹. An EASY-nLC 1200 system (Thermo Fisher Scientific), directly coupled online with the mass spectrometer (Q Exactive HF-X, Thermo Fisher Scientific) via a nano-electrospray source, was employed for nano-flow liquid chromatography. Peptides were eluted by a linear 80 min gradient from 5% to 30% buffer B (0.1% v/v FA, 80% v/v ACN), followed by a 4 min increase to 60% B, a further 4 min increase to 95% B, a 4 min plateau phase at 95% B, a 4 min decrease to 5% B and a 4 min wash phase of 5% B. To acquire MS data, the data-independent acquisition (DIA) scan mode operated by the XCalibur software (Thermo Fisher) was used. DIA was performed with one full MS event followed by 33 MS/MS windows in one cycle resulting in a cycle time of 2.7 s. The full MS settings included an ion target value of 3×10^6 charges in the 300–1,650 m/z range with a maximum injection time of 60 ms and a resolution of 120,000 at m/z 200. DIA precursor windows ranged from 300.5 m/z (lower boundary of first window) to 1,649.5 m/z (upper boundary of 33rd window). MS/MS settings included an ion target value of 3×10^6 charges for the precursor window with an Xcalibur-automated maximum injection time and a resolution of 30,000 at m/z 200.

To generate the proteome library for DIA measurements purified peptides from the first and the fourth replicates of all samples were pooled separately and 25 µg of peptides from each pool were fractionated into 24 fractions by high pH reversed-phase chromatography as described earlier[60]. During each separation, fractions were concatenated automatically by shifting the collection tube every 120 s. In total 48 fractions were dried in a vacuum centrifuge, resuspended in buffer A* (0.2% trifluoroacetic acid (TFA), 2% ACN) and subsequently analysed by a top-12 data-dependent acquisition (DDA) scan mode using the same LC gradient and settings. The mass spectrometer was operated by the XCalibur software (Thermo Fisher). DDA scan settings on full MS level included an ion target value of 3×10^6 charges in the 300–1,650 m/z range with a maximum injection time of 20 ms and a resolution of 60,000 at m/z 200. At the MS/MS level the target value was 10^5 charges with a maximum injection time of 60 ms and a resolution of 15,000 at m/z 200. For MS/MS events only, precursor ions with 2–5 charges that were not on the 20-s dynamic exclusion list were isolated in a 1.4 m/z window. Fragmentation was performed by higher-energy C-trap dissociation with a normalized collision energy of 27 eV.

Infected time-course proteome–phosphoproteome–diGly proteome sample preparation

Frozen lysates of infected A549-ACE2 cells collected at 6, 12 and 24 hpi (and 36 hpi for the phosphoproteomics study) were thawed on ice, boiled for 5 min at 95 °C and sonicated for 15 min (Branson Sonifierer). Protein concentrations were estimated by tryptophan assay[61]. To reduce and alkylate proteins, samples were incubated for 5 min at 45 °C with TCEP (10 mM) and CAA (40 mM). Samples were digested overnight at 37 °C using trypsin (1:100 w/w, enzyme/protein, Sigma-Aldrich) and LysC (1:100 w/w, enzyme/protein, Wako).

For proteome analysis, 10 µg of peptide material were desalted using SDB-RPS StageTips (Empore)[61]. In brief, samples were diluted with 1% TFA in isopropanol to a final volume of 200 µl and loaded onto StageTips, subsequently washed with 200 µl of 1% TFA in isopropanol and 200 µl 0.2% TFA/ 2% ACN. Peptides were eluted with 75 µl of 1.25% ammonium hydroxide (NH₄OH) in 80% ACN and dried using a SpeedVac centrifuge (Eppendorf, Concentrator Plus). They were resuspended in buffer A* (0.2% TFA, 2% ACN) before LC–MS/MS analysis. Peptide concentrations were measured optically at 280 nm (Nanodrop 2000, Thermo Scientific) and subsequently equalized using buffer A*. One microgram of peptide was analysed by LC–MS/MS.

The rest of the samples were diluted fourfold with 1% TFA in isopropanol and loaded onto SDB-RPS cartridges (Strata-X-C, 30 mg per 3 ml, Phenomenex), pre-equilibrated with 4 ml 30% MeOH/1% TFA and washed with 4 ml 0.2% TFA. Samples were washed twice with 4 ml 1% TFA in isopropanol, once with 0.2% TFA/2% ACN and eluted twice with 2 ml 1.25% NH₄OH/80% ACN. Eluted peptides were diluted with ddH₂O to a final ACN concentration of 35%, snap frozen and lyophilized.

For phosphopeptide enrichment, lyophilized peptides were resuspended in 105 µl of equilibration buffer (1% TFA/80% ACN) and the peptide concentration was measured optically at 280 nm (Nanodrop 2000, Thermo Scientific) and subsequently equalized using equilibration buffer. The AssayMAP Bravo robot (Agilent) performed the enrichment for phosphopeptides (150 µg) by priming AssayMAP cartridges (packed with 5 µl Fe³⁺-NTA) with 0.1% TFA in 99% ACN followed by equilibration in equilibration buffer and loading of peptides. Enriched phosphopeptides were eluted with 1% ammonium hydroxide, which was evaporated using a Speedvac for 20 min. Dried peptides were resuspended in 6 µl buffer A* and 5 µl was analysed by LC–MS/MS.

For diGly peptide enrichment, lyophilized peptides were reconstituted in IAP buffer (50 mM MOPS, pH 7.2, 10 mM Na₂HPO₄, 50 mM NaCl) and the peptide concentration was estimated by tryptophan assay. K-ε-GG remnant containing peptides were enriched using the PTMScan Ubiquitin Remnant Motif (K-ε-GG) Kit (Cell Signaling Technology). Cross-linking of antibodies to beads and subsequent immunopurification was performed with slight modifications as previously described[62]. In brief, two vials of cross-linked beads were combined and equally split into 16 tubes (~31 µg of antibody per tube). Equal peptide amounts (600 µg) were added to cross-linked beads, and the volume was adjusted with IAP buffer to 1 ml. After 1 h of incubation at 4 °C and gentle agitation, beads were washed twice with cold IAP and 5 times with cold ddH₂O. Thereafter, peptides were eluted twice with 50 µl 0.15% TFA. Eluted peptides were desalted and dried as described for proteome analysis with the difference that 0.2% TFA instead of 1% TFA in isopropanol was used for the first wash. Eluted peptides were resuspended in 9 µl buffer A* and 4 µl was subjected to LC–MS/MS analysis.

DIA measurements

Samples were loaded onto a 50-cm reversed-phase column (75 µm inner diameter, packed in house with ReproSil-Pur C18-AQ 1.9 µm resin (Dr Maisch)). The column temperature was maintained at 60 °C using a homemade column oven. A binary buffer system, consisting of buffer A (0.1% FA) and buffer B (80% ACN plus 0.1% FA) was used for peptide separation, at a flow rate of 300 nl min⁻¹. An EASY-nLC 1200 system (Thermo Fisher Scientific), directly coupled online with the mass spectrometer (Orbitrap Exploris 480, Thermo Fisher Scientific) via a nano-electrospray source, was employed for nano-flow liquid chromatography. The FAIMS device was placed between the nanoelectrospray source and the mass spectrometer and was used for measurements of the proteome and the PTM-library samples. Spray voltage was set to 2,650 V, RF level to 40 and heated capillary temperature to 275 °C.

For proteome measurements we used a 100 min gradient starting at 5% buffer B followed by a stepwise increase to 30% in 80 min, 60% in 4 min and 95% in 4 min. The buffer B concentration stayed at 95% for 4 min, decreased to 5% in 4 min and stayed there for 4 min. The mass spectrometer was operated in data-independent mode (DIA) with a full scan range of 350–1,650 m/z at 120,000 resolution at 200 m/z,

normalized automatic gain control (AGC) target of 300% and a maximum fill time of 28 ms. One full scan was followed by 22 windows with a resolution of 15,000, normalized AGC target of 1,000% and a maximum fill time of 25 ms in profile mode using positive polarity. Precursor ions were fragmented by HCD (NCE 30%). Each of the selected compensation voltage (CV) (−40, −55 and −70 V) was applied to sequential survey scans and MS/MS scans; the MS/MS CV was always paired with the appropriate CV from the corresponding survey scan.

For phosphopeptide samples, 5 µl were loaded and eluted with a 70-min gradient starting at 3% buffer B followed by a stepwise increase to 19% in 40 min, 41% in 20 min, 90% in 5 min and 95% in 5 min. The mass spectrometer was operated in DIA mode with a full scan range of 300–1,400 m/z at 120,000 resolution at 200 m/z and a maximum fill time of 60 ms. One full scan was followed by 32 windows with a resolution of 30,000. Normalized AGC target and maximum fill time were set to 1,000% and 54 ms, respectively, in profile mode using positive polarity. Precursor ions were fragmented by HCD (NCE stepped 25–27.5–30%). For the library generation, we enriched A549 cell lysates for phosphopeptides and measured them with 7 different CV settings (−30, −40, −50, −60, −70, −80 or −90 V) using the same DIA method. The noted CVs were applied to the FAIMS electrodes throughout the analysis.

For the analysis of K-ε-GG peptide samples, half of the samples were loaded. We used a 120-min gradient starting at 3% buffer B followed by a stepwise increase to 7% in 6 min, 20% in 49 min, 36% in 39 min, 45% in 10 min and 95% in 4 min. The buffer B concentration stayed at 95% for 4 min, decreased to 5% in 4 min and stayed there for 4 min. The mass spectrometer was operated in DIA mode with a full scan range of 300–1,350 m/z at 120,000 resolution at m/z 200, normalized AGC target of 300% and a maximum fill time of 20 ms. One full scan was followed by 46 windows with a resolution of 30,000. Normalized AGC target and maximum fill time were set to 1,000% and 54 ms, respectively, in profile mode using positive polarity. Precursor ions were fragmented by HCD (NCE 28%). For K-ε-GG peptide library, we mixed the first replicate of each sample and measured them with eight different CV setting (−35, −40, −45, −50, −55, −60, −70 or −80 V) using the same DIA method.

Processing of raw MS data
AP–MS data. Raw MS data files of AP–MS experiments conducted in DDA mode were processed with MaxQuant (version 1.6.14) using the standard settings and label-free quantification (LFQ) enabled (LFQ min ratio count 1, normalization type none, stabilize large LFQ ratios disabled). Spectra were searched against forward and reverse sequences of the reviewed human proteome including isoforms (UniprotKB, release 2019.10) and C-terminally HA-tagged SARS-CoV-2, SARS-CoV and HCoV proteins by the built-in Andromeda search engine[63].

In-house Julia scripts (https://doi.org/10.5281/zenodo.4541090) were used to define alternative protein groups: only the peptides identified in AP–MS samples were considered for being protein group-specific, protein groups that differed by the single specific peptide or had less than 25% different specific peptides were merged to extend the set of peptides used for protein group quantitation and reduce the number of protein isoform-specific interactions.

Viral protein overexpression and DIA MS data. Spectronaut version 13 (Biognosys) with the default settings was used to generate the proteome libraries from DDA runs by combining files of respective fractionations using the human fasta file (Uniprot, 2019.10, 42 431 entries) and viral bait sequences. Proteome DIA files were analysed using the proteome library with the default settings and disabled cross run normalization.

SARS-CoV-2/SARS-CoV-infected proteome/PTM DIA MS data. Spectronaut version 14 (Biognosys)[64] was used to generate the libraries and analyse all DIA files using the human fasta file (UniprotKB, release 2019.10) and sequences of SARS-CoV-2/SARS-CoV proteins (UniProt, release 2020.08). Orf1a polyprotein sequences were split into separate protein chains according to the cleavage positions specified in the UniProt. For the generation of the PTM-specific libraries, the DIA single CV runs were combined with the actual DIA runs and either phosphorylation at serine, threonine or tyrosine, or GlyGly at lysine, was added as variable modification to default settings. The maximum number of fragment ions per peptide was increased to 25. The proteome DIA files were analysed using direct DIA approach with default settings and disabled cross run normalization. All post-translational modification DIA files were analysed using their respective hybrid library and either phosphorylation at Serine/Threonine/Tyrosine or GlyGly at Lysine was added as an additional variable modification to default settings with LOESS normalization and disabled PTM localization filter.

A collection of in-house Julia scripts (https://doi.org/10.5281/zenodo.4541090) were used to process the elution group (EG) -level Spectronaut reports, identify PTMs and assign EG-level measurements to PTMs. The PTM was considered if at least once it was detected with ≥0.75 localization probability in EG with q-value ≤10^{-3}. For further analysis of given PTM, only the measurements with ≥0.5 localization probability and EG q-value ≤10^{-2} were used.

Bioinformatic analysis
Unless otherwise specified, the bioinformatic analysis was done in R (version 3.6), Julia (version 1.5) and Python (version 3.8) using a collection of in-house scripts (https://doi.org/10.5281/zenodo.4541090 and https://doi.org/10.5281/zenodo.4541082).

Datasets. The following public datasets were used in the study: Gene Ontology and Reactome annotations (http://download.baderlab.org/EM_Genesets/April_01_2019/Human/UniProt/Human_GO_AllPathways_with_GO_iea_April_01_2019_UniProt.gmt); IntAct Protein Interactions (https://www.ebi.ac.uk/intact/, v2019.12); IntAct Protein Complexes (https://www.ebi.ac.uk/complexportal/home, v2019.12); CORUM Protein Complexes (http://mips.helmholtz-muenchen.de/corum/download/allComplexes.xml.zip, v2018.3); Reactome Functional Interactions (https://reactome.org/download/tools/ReatomeFIs/FIsInGene_020720_with_annotations.txt.zip); Human (v2019.10), Human-CoV, SARS-CoV-2 and SARS-CoV (v2020.08) protein sequences: https://uniprot.org.

Statistical analysis of MS data. MaxQuant and Spectronaut output files were imported into R using in-house maxquantUtils R package (https://doi.org/10.5281/zenodo.4536603). For all MS datasets, the Bayesian linear random effects models were used to define how the abundances of proteins change between the conditions. To specify and fit the models we used the msglm R package (https://doi.org/10.5281/zenodo.4536605), which uses the rstan package (version 2.19)[65] for inferring the posterior distribution of the model parameters. In all the models, the effects corresponding to the experimental conditions have regularized horseshoe+ priors[66], whereas the batch effects have normally distributed priors. Laplacian distribution was used to model the instrumental error of MS intensities. For each MS instrument used, the heteroscedastic intensities noise model was calibrated with the technical replicate MS data of the instrument. These data were also used to calibrate the logit-based model of missing MS data (the probability that the MS instrument will fail to identify the protein given its expected abundance in the sample). The model was fit using unnormalized MS intensities data. Instead of transforming the data by normalization, the inferred protein abundances were scaled by the normalization multiplier of each individual MS sample to match the expected MS intensity of that sample. This allows taking the signal-to-noise variation between the samples into account when fitting the model. Due to high computational intensity, the model was applied to each protein group separately. For all the models, 4,000 iterations (2,000 warmup + 2,000 sampling) of the no-U-turn Markov Chain Monte Carlo were performed in 7 or 8 independent chains, every 4th sample was collected

Article

for posterior distribution of the model parameters. For estimating the statistical significance of protein abundance changes between the two experimental conditions, the P-value was defined as the probability that a random sample from the posterior distribution of the first condition would be smaller (or larger) than a random sample drawn from the second condition. No-multiple hypothesis testing corrections were applied, since this is handled by the choice of the model priors.

Statistical analysis of AP–MS data and filtering for specific interactions. The statistical model was applied directly to the MS1 intensities of protein group-specific LC peaks (evidence.txt table of MaxQuant output). In R GLM formula language, the model could be specified as

$$\log(\text{Intensity}) \approx 1 + \text{APMS} + \text{Bait} + \text{Bait}:\text{Virus} + \text{MS1peak} + \text{MSbatch},$$

where the APMS effect models the average shift of intensities in AP–MS data in comparison to full proteome samples, Bait is the average enrichment of a protein in AP–MS experiments of homologous proteins of both SARS-CoV and SARS-CoV-2, and Bait:Virus corresponds to the virus-specific changes in protein enrichment. MS1peak is the log ratio between the intensity of a given peak and the total protein abundance (the peak is defined by its peptide sequence, PTMs and the charge; it is assumed that the peak ratios do not depend on experimental conditions[67]), and MSbatch accounts for batch-specific variations of protein intensity. APMS, Bait and Bait:Virus effects were used to reconstruct the batch effect-free abundance of the protein in AP–MS samples.

The modelling provided the enrichment estimates for each protein in each AP experiment. Specific AP–MS interactions had to pass the two tests. In the first test, the enrichment of the candidate protein in a given bait AP was compared against the background, which was dynamically defined for each interaction to contain the data from all other baits, where the abundance of the candidate was within 50–90% percentile range (excluding top 10% baits from the background allowed the protein to be shared by a few baits in the resulting AP–MS network). The non-targeting control and Gaussian luciferase baits were always preserved in the background. Similarly, to filter out any potential side-effects of very high bait protein expression, the ORF3 homologues were always present in the background of M interactors and vice versa. To rule out the influence of the batch effects, the second test was applied. It was defined similarly to the first one, but the background was constrained to the baits of the same batch, and 40–80% percentile range was used. In both tests, the protein has to be fourfold enriched over the background (16 fold for highly expressed baits: ORF3, M, NSP13, NSP5, NSP6, ORF3a, ORF7b, ORF8b and HCoV-229E ORF4a) with P-value $\leq 10^{-3}$.

Additionally, we excluded the proteins that, in the viral protein expression data, have shown upregulation, and their enrichment in AP–MS data was less than 16 times stronger than observed upregulation effects. Finally, to exclude the carryover of material between the samples sequentially analysed by MS, we removed the putative interactors, which were also enriched at higher levels in the samples of the preceding bait, or the one before it.

For the analysis of interaction specificity between the homologous viral proteins, we estimated the significance of interaction enrichment difference (corrected by the average difference between the enrichment of the shared interactors to adjust for the bait expression variation). Specific interactions have to be fourfold enriched in comparison to the homologue with P-value $\leq 10^{-3}$.

Statistical analysis of DIA proteome effects upon viral protein overexpression. The statistical model of the viral protein overexpression dataset was similar to AP–MS data, except that protein-level intensities provided by Spectronaut were used. The PCA analysis of the protein intensities has identified that the second principal component is associated with the batch-dependent variations between the samples.

To exclude their influence, this principal component was added to the experimental design matrix as an additional batch effect.

As with AP–MS data, the two statistical tests were used to identify the significantly regulated proteins (column 'is_change' in Supplementary Table 3). First, the absolute value of median $\log_2$-fold change of the protein abundance upon overexpression of a given viral protein in comparison to the background had to be above 1.0 with P-value $\leq 10^{-3}$. The background was individually defined for each analysed protein. It was composed of experiments, where the abundance of given protein was within the 20–80% percentile range of all measured samples. Second, the protein had to be significantly regulated (same median $\log_2$-fold change and P-value thresholds applied) against the batch-specific background (defined similarly to the global background, but using only the samples of the same batch).

An additional stringent criterion was applied to select the most significant changes (column 'is_top_change' in Supplementary Table 3; Extended Data Fig. 1i).

For each protein we classified bait-induced changes as: 'high' when $|\text{median} \log_2 \text{fold-change}| \geq 1$ and P-value $\leq 10^{-10}$ both in background and batch comparisons; 'medium' if $10^{-10} < P$-value $\leq 10^{-4}$ with same fold-change requirement; and 'low' if $10^{-4} < P$-value $\leq 10^{-2}$ with the same fold-change requirement. All other changes were considered non-significant.

We then required that 'shared' top-regulated proteins should have exactly one pair of SARS-CoV-2 and SARS-CoV high- or medium-significant homologous baits among the baits with either up- or downregulated changes and no other baits with significant changes of the same type.

We further defined 'SARS-CoV-2-specific' or 'SARS-CoV-specific' top-regulated proteins to be the ones with exactly one high-significant change, and no other significant changes of the same sign. For 'specific' hits we additionally required that in the comparison of high-significant bait to its homologue $|\text{median} \log_2 \text{fold-change}| \geq 1$ and P-value $\leq 10^{-3}$. When the homologous bait was missing (SARS-CoV-2 NSP1, SARS-CoV ORF8a and SARS-CoV ORF8b), we instead required that in the comparison of the high-significant change to the background $|\text{median} \log_2 \text{fold-change}| \geq 1.5$.

The resulting network of most affected proteins was imported and prepared for publication in Cytoscape v.3.8.1[68].

Statistical analysis of DIA proteomic data of SARS-CoV-2 and SARS-CoV-infected A549-ACE2 cells. Similarly to the AP–MS DDA data, the linear Bayesian model was applied to the EG-level intensities. To model the protein intensity, the following linear model (in R notation) was used:

$$\log(\text{Intensity}(t)) \sim 1 + \sum_{t_i \leq t} (\text{after}(t_i) + (\text{infection} \pm \text{CoV2}):\text{after}(t_i))$$
$$+ \text{EG},$$

where the after(t_i) effect corresponds to the protein abundance changes in mock-infected samples that happened between t_{i-1} and t_i after infection and it is applied to the modelled intensity at all time points starting from t_i; infection:after(t_i) ($t_i = 6, 12, 24$) is the common effect of SARS-CoV-2 and SARS-CoV infections occurring between t_{i-1} and t_i; CoV2:after(t_i) is the virus-specific effect within t_{i-1} and t_i hpi that is added to the log intensity for SARS-CoV-2-infected samples and subtracted from the intensity for SARS-CoV ones; EG is the elution group-specific shift in the measured log-intensities.

The absolute value of median $\log_2$ fold change between the conditions above 0.25 and the corresponding unadjusted P-value $\leq 10^{-3}$ were used to define the significant changes at a given time point in comparison to mock infection. We also required that the protein group is quantified in at least two replicates of at least one of the compared conditions. Additionally, if for one of the viruses (for example, SARS-CoV-2)

only the less stringent condition ($|\text{median}\log_2\text{ fold-change}| \geq 0.125$, P-value $\leq 10^{-2}$) was fulfilled, but the change was significant in the infection of the other virus (SARS-CoV), and the difference between the viruses was not significant, the observed changes were considered significant for both viruses.

Statistical analysis of DIA phosphoproteome and ubiquitinome data of SARS-CoV-2 and SARS-CoV infections. The data from single-double- and triple-modified peptides were analysed separately and, for a given PTM, the most significant result was reported.

The data were analysed with the same Bayesian linear model as proteome SARS-CoV and SARS-CoV-2 infection data. In addition to the intensities normalization, for each replicate sample the scale of the effects in the experimental design matrix was adjusted, so that on average the correlation between log fold changes of the replicates was 1:1. The same logic as for the proteome analysis, was applied to identify significant changes, but the median $\log_2$ fold change had to be larger than 0.5, or 0.25 for the less stringent test. We additionally required that the PTM peptides are quantified in at least two replicates of at least one of the compared conditions. To ignore the changes in PTM site intensities that are due to proteome-level regulation, we excluded PTM sites on significantly regulated proteins if the directions of protein and PTM site changes were the same and the difference between their median $\log_2$ fold changes was less than two. Phosphoproteomics data were further analysed with Ingenuity Pathway Analysis software (Qiagen; https://www.qiagenbioinformatics.com/products/ingenuity-pathway-analysis)

Transcriptomic analysis of SARS-CoV-2 and SARS-CoV infected A549-ACE2 cells. For the analysis of the transcriptome data, Gencode gene annotations v28 and the human reference genome GRCh38 were derived from the Gencode homepage (EMBL-EBI). Viral genomes were derived from GenBank (SARS-CoV-2 - LR824570.1, and SARS-CoV - AY291315.1). Dropseq tool v1.12 was used for mapping raw sequencing data to the reference genome. The resulting UMI filtered count matrix was imported into R v3.4.4. CPM (counts per million) values were calculated for the raw data and genes having a mean cpm value less than 1 were removed from the dataset. A dummy variable combining the covariates infection status (mock, SARS-CoV, SARS-CoV-2) and time point was used for modelling the data within Limma (v3.46.0)[69].

Data were transformed with the Voom method[69] followed by quantile normalization. Differential testing was performed between infection states at individual time points by calculating moderated t-statistics and P-values for each host gene. A gene was considered to be significantly regulated if the false discovery rate-adjusted P-value was below 0.05.

Gene set enrichment analysis. We used Gene Ontology, Reactome and other EnrichmentMap gene sets of human proteins (version 2020.10)[70] as well as protein complexes annotations from IntAct Complex Portal (version 2019.11)[71] and CORUM (version 2019)[72]. PhosphoSitePlus (version 2020.08) was used for known kinase-substrate and regulatory sites annotations, Perseus (version 1.6.14.0)[73] was used for annotation of known kinase motifs. For transcription factor enrichment analysis (Extended Data Fig. 2e) the significantly regulated transcripts were submitted to ChEA3 web-based application[74] and ENCODE data on transcription factor–target gene associations were used[75].

To find the non-redundant collection of annotations describing the unique and shared features of multiple experiments in a dataset (Fig. 1d, Extended Data Fig. 2l, m), we used in-house Julia package OptEnrichedSetCover.jl (https://doi.org/10.5281/zenodo.4536596), which employs evolutionary multi-objective optimization technique to find a collection of annotation terms that have both significant enrichments in the individual experiments and minimal pairwise overlaps.

The resulting set of terms was further filtered by requiring that the annotation term has to be significant with the specified unadjusted Fisher's exact test P-value cut-off in at least one of the experiments or comparisons (the specific cut-off value is indicated in the figure legend of the corresponding enrichment analysis).

The generation of diagonally-split heat maps was done with the VegaLite.jl package (https://github.com/queryverse/VegaLite.jl).

Viral PTMs alignment. For matching the PTMs of SARS-CoV-2 and SARS-CoV the protein sequences were aligned using the BioAlignments. jl Julia package (v.2.0; https://github.com/BioJulia/BioAlignments.jl) with the Needleman–Wunsch algorithm using BLOSUM80 substitution matrix, and applying −5 and −3 penalties for the gap and extension, respectively.

For the cellular proteins, we required that the viral phosphorylation or ubiquitination site is observed with q-value $\leq 10^{-3}$ and localization probability ≥ 0.75. For the PTMs with lower confidence (q-value $\leq 10^{-2}$ and localization probability ≥ 0.5) we required that the same site is observed with high confidence at the matching position of the orthologous protein of the other virus.

Network diffusion analysis. To systematically detect functional interactions, which may connect the cellular targets of each viral protein (interactome dataset) with the downstream changes it induces on proteome level (effectome dataset), we have used the network diffusion-based HierarchicalHotNet method[36] as implemented in Julia package HierarchicalHotNet.jl (https://doi.org/10.5281/zenodo.4536590). Specifically, for network diffusion with restart, we used the ReactomeFI network (version 2019)[35] of cellular functional interactions, reversing the direction of functional interaction (for example, replacing kinase→substrate interaction with substrate→kinase). The proteins with significant abundance changes upon bait overexpression ($|\text{median}(\log_2\text{ fold change})| \geq 0.25$, $P \leq 10^{-2}$ both in the comparison against the controls and against the baits of the same batch) were used as the sources of signal diffusion with weights set to $w_i = \sqrt{|\text{median}\log_2(\text{fold change})| \cdot |\log_{10}P\text{-value}|}$, otherwise the node weight was set to zero. The weight of the edge $g_i{\rightarrow}g_j$ was set to $w_{i,j} = 1 + w_j$. The restart probability was set to 0.4, as suggested in the original publication, so that the probability of the random walk to stay in the direct neighbourhood of the node is the same as the probability to visit more distant nodes. To find the optimal cutting threshold of the resulting hierarchical tree of strongly connected components (SCCs) of the weighted graph corresponding to the stationary distribution of signal diffusion and to confirm the relevance of predicted functional connections, the same procedure was applied to 1,000 random permutations of vertex weights as described in Reyna et al.[36] (vertex weights are randomly shuffled between the vertices with similar in and out degrees). Since cutting the tree of SCCs at any threshold t (keeping only the edges with weights above t) and collapsing each resulting SCC into a single node produces the directed acyclic graph of connections between SCCs, it allowed efficient enumeration of the paths from the 'source' nodes (proteins strongly perturbed by viral protein expression with vertex weight w, $w \geq 1.5$) to the 'sink' nodes (interactors of the viral protein). At each threshold t, the average inverse of the path length from source to sink nodes was calculated as:

$$L_{\text{avg}}^{-1}(t) = \frac{1}{N_{\text{source}} \cdot N_{\text{sink}}} \sum_p L_{\text{SCC}}^{-1}(p),$$

where N_{source} is the number of sources, N_{sink} is the number of sinks, $L_{\text{SCC}}(p)$ is the number of SCCs that the given path p from source to sink goes through, and the sum is for all paths from sources to sinks. The metric changes from 1 (all sources and sinks in the same SCC) to 0 (no or infinitely long paths between sources and sinks). For the generation of the diffusion networks we were using the t_{opt} threshold that maximized the difference between $L_{\text{avg}}^{-1}(t)$ for the real data and the third quartile of $L_{\text{avg}}^{-1}(t)$ for randomly shuffled data.

Article

In the generated SCC networks, the direction of the edges was reverted back, and the results were exported as GraphML files using in-house Julia scripts (https://doi.org/10.5281/zenodo.4541090). The catalogue of the networks for each viral bait is available as Supplementary Data 1.

To assess the significance of edges in the resulting network, we calculated the P-value of the edge $g_i \rightarrow g_j$ as the probability that the transition probability between the given pair of genes based on permuted data is higher than the transition probability based on the real data:

$$P(w_{\text{real}}(g_i, g_j) \leq w_{\text{perm}}(g_i, g_j)).$$

This P-value was stored as the 'prob_perm_walkweight_greater' edge attribute of GraphML output. The specific subnetworks predicted by the network diffusion (Fig. 4b–d) were filtered for edges with $P \leq 0.05$.

When the $g_i \rightarrow g_j$ connection was not present in the ReactomeFI network, to recover the potential short pathways connecting g_i and g_j, ReactomeFI was searched for intermediate g_k nodes, such that the edges $g_i \rightarrow g_k$ and $g_k \rightarrow g_j$ are present in ReactomeFI. The list of these short pathways is provided as the 'flowpaths' edge attribute in GraphML output.

The GraphML output of network diffusion was prepared for publication using yEd (v.3.20; https://www.yworks.com).

Intersection with other SARS coronavirus datasets. The intersection between the data generated by this study and other publicly available datasets was done using the information from respective supplementary tables. When multiple viruses were used in a study, only the comparisons with SARS-CoV and SARS-CoV-2 were included. For time-resolved data, all time points up to 24 hpi were considered. The dataset coverage was defined as the number of reported distinct protein groups for proteomic studies and genes for transcriptomic studies. Confident interactions or significant regulations were filtered according to the criteria specified in the original study. A hit was considered as 'confirmed' when it was significant both in the present study and the external data and showed the same trend.

qRT–PCR analysis

RNA isolation from SARS-CoV and SARS-CoV-2 infected A549-ACE2 cells was performed as described above (Qiagen). Five hundred nanograms total RNA was used for reverse transcription with PrimeScript RT with gDNA eraser (Takara). For relative transcript quantification PowerUp SYBR Green (Applied Biosystems) was used. Primer sequences can be provided upon request.

Co-immunoprecipitation and western blot analysis

HEK293T cells were transfected with pWPI plasmid encoding single HA-tagged viral proteins, alone or together with pTO-SII-HA expressing host factor of interest. 48 h after transfection, cells were washed in PBS, flash frozen in liquid nitrogen and kept at −80 °C until further processing. Co-immunoprecipitation experiments were performed as described previously[55,56]. In brief, cells were lysed in lysis buffer (50 mM Tris-HCl pH 7.5, 100 mM NaCl, 1.5 mM MgCl$_2$, 0.2% (v/v) NP-40, 5% (v/v) glycerol, cOmplete protease inhibitor cocktail (Roche), 0.5% (v/v) 750 U µl^{-1} Sm DNase) and sonicated (5 min, 4 °C, 30 s on, 30 s off, low settings; Bioruptor, Diagenode SA). HA or Streptactin beads were added to cleared lysates and samples were incubated for 3 h at 4 °C under constant rotation. Beads were washed six times in the lysis buffer and resuspended in 1× SDS sample buffer 62.5 mM Tris-HCl pH 6.8, 2% SDS, 10% glycerol, 50 mM DTT, 0.01% bromophenol blue). After boiling for 5 min at 95 °C, a fraction of the input lysate and elution were loaded on NuPAGE Novex 4–12% Bis-Tris (Invitrogen), and further submitted to western blotting using Amersham Protran nitrocellulose membranes. Imaging was performed by HRP luminescence (ECL, Perkin Elmer).

SARS-CoV-2 infected A549-ACE2 cell lysates were sonicated (10 min, 4 °C, 30 s on, 30 s off, low settings; Bioruptor, Diagenode). Protein concentration was adjusted based on Pierce660 assay supplemented with ionic detergent compatibility reagent. After boiling for 5 min at 95 °C and brief centrifugation at maximum speed, the samples were loaded on NuPAGE Novex 4–12% Bis-Tris (Invitrogen), and blotted onto 0.22 µm Amersham Protran nitrocellulose membranes (Merck). Primary and secondary antibody stainings were performed according to the manufacturer's recommendations. Imaging was performed by HRP luminescence using Femto kit (ThermoFischer Scientific) or Western Lightning PlusECL kit (Perkin Elmer).

Mapping of post-translational modification sites on the N CTD structure

N CTD dimers of SARS-CoV-2 (PDB: 6YUN) and SARS-CoV (PDB: 2CJR) were superimposed by aligning the α-carbons backbone over 111 residues (from position 253/254 to position 364/365 following SARS-CoV-2/SARS-CoV numbering) by using the tool MatchMaker[76] as implemented in the Chimera software[77]. Ubiquitination sites were visually inspected and mapped by using the PyMOL software (https://pymol.org). Phosphorylation on Ser310/311 was simulated in silico by using the PyTMs plugin as implemented in PyMOL[78]. Inter-chain residue contacts, dimer interface area, free energy and complex stability were comparatively analysed between non-phosphorylated and phosphorylated SARS-CoV-2 and SARS-CoV N CTD by using the PDBePISA server[79]. Poisson–Boltzmann electrostatic surface potential of native and post-translationally modified N CTD was calculated by using the PBEQ Solver tool on the CHARMM-GUI server by preserving existing hydrogen bonds[80]. Molecular graphics depictions were produced with the PyMOL software.

Reporter assay and IFN bioassay

The following reporter constructs were used in this study: pISRE-luc was purchased from Stratagene, EF1-α-ren was obtained from E. Gürlevik (Department of Gastroenterology, Hepatology and Endocrinology, Hannover Medical School, Germany), pCAGGS-Flag-RIG-I was obtained from C. Basler (Department of Microbiology, Mount Sinai School of Medicine, USA), pIRF1-GAS-ff-luc, pWPI-SMN1-flag and pWPI-NS5 (ZIKV)-HA was described previously[56,81].

For the reporter assay, HEK293-R1 cells were plated in 24-well plates 24 h before transfection. Firefly reporter and *Renilla* transfection control were transfected together with plasmids expressing viral proteins using polyethylenimine (PEI, Polysciences) for untreated and treated conditions. In 18 h cells were stimulated for 8 h with a corresponding inducer and collected in the passive lysis buffer (Promega). Luminescence of Firefly and *Renilla* luciferases was measured using dual-luciferase-reporter assay (Promega) according to the manufacturer's instructions in a microplate reader (Tecan).

Total amounts of IFN-α and IFN-β in cell supernatants were measured by using 293T cells stably expressing the firefly luciferase gene under the control of the mouse Mx1 promoter (Mx1-luc reporter cells)[82]. In brief, HEK293-R1 cells were seeded, transfected with pCAGGS-flag-RIG-I plus viral protein constructs and stimulated as described above. Cell supernatants were collected in 8 h. Mx1-luc reporter cells were seeded into 96-well plates in triplicates and were treated 24 h later with supernatants. At 16 h after incubation, cells were lysed in the passive lysis buffer (Promega), and luminescence was measured with a microplate reader (Tecan). The assay sensitivity was determined by a standard curve.

Viral inhibitor assay

A549-ACE2 cells were seeded into 96-well plates in DMEM medium (10% FCS, 100 µg ml^{-1} streptomycin, 100 IU ml^{-1} penicillin) one day before infection. Six hours before infection, or at the time of infection, the medium was replaced with 100 µl of DMEM medium containing either the compounds of interest or DMSO as a control. Infection was performed by adding 10 µl of SARS-CoV-2-GFP (MOI of 3) per well and plates were placed in the IncuCyte S3 Live-Cell Analysis System (Essen

Bioscience), where whole well real-time images of mock (phase channel) and infected (GFP and phase channel) cells were captured every 4 h for 48 h. Cell viability (mock) and virus growth (mock and infected) were assessed as the cell confluence per well (phase area) and GFP area normalized by cell confluence per well (GFP area/phase area) respectively using IncuCyte S3 Software (Essen Bioscience; version 2019B rev2).

For comparative analysis of antiviral treatment activity against SARS-CoV and SARS-CoV-2, A549-ACE2 cells were seeded in 24-well plates, as previously described. Treatment was performed for 6 h with 0.5 ml of DMEM medium containing either the compounds of interest or DMSO as a control, and infected with SARS-CoV-Frankfurt-1 or SARS-CoV-2-MUC-IMB-1 (MOI of 1) for 24 h. Total cellular RNA was collected and analysed by RT–qPCR, as previously described.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

The raw sequencing data for this study have been deposited with the ENA at EMBL-EBI under accession number PRJEB38744. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE[83] partner repository with the dataset identifiers PXD022282, PXD020461 and PXD020222. Protein interactions identified in this study have been submitted to the IMEx (https://www.imexconsortium.org) consortium through IntAct[84] with the identifier IM-28109. The data and analysis results are accessible online via the interactive web interface at https://covinet.innatelab.org.

Code availability

In-house R and Julia packages and scripts used for the bioinformatics analysis of the data have been deposited to public GitHub repositories: https://doi.org/10.5281/zenodo.4536605, https://doi.org/10.5281/zenodo.4536603, https://doi.org/10.5281/zenodo.4536590, https://doi.org/10.5281/zenodo.4536596, https://doi.org/10.5281/zenodo.4541090 and https://doi.org/10.5281/zenodo.4541082.

55. Hubel, P. et al. A protein-interaction network of interferon-stimulated genes extends the innate immune system landscape. *Nat. Immunol.* **20**, 493–502 (2019).
56. Scaturro, P. et al. An orthogonal proteomic survey uncovers novel Zika virus host factors. *Nature* **561**, 253–257 (2018).
57. Hoffmann, M. et al. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* **181**, 271–280.e8 (2020).
58. Gebhardt, A. et al. The alternative cap-binding complex is required for antiviral defense in vivo. *PLoS Pathog.* **15**, e1008155 (2019).
59. Goldeck, M., Schlee, M., Hartmann, G. & Hornung, V. Enzymatic synthesis and purification of a defined RIG-I ligand. *Methods Mol. Biol.* **1169**, 15–25 (2014).
60. Kulak, N. A., Geyer, P. E. & Mann, M. Loss-less nano-fractionator for high sensitivity, high coverage proteomics. *Mol. Cell. Proteomics* **16**, 694–705 (2017).
61. Kulak, N. A., Pichler, G., Paron, I., Nagaraj, N. & Mann, M. Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells. *Nat. Methods* **11**, 319–324 (2014).
62. Udeshi, N. D., Mertins, P., Svinkina, T. & Carr, S. A. Large-scale identification of ubiquitination sites by mass spectrometry. *Nat. Protoc.* **8**, 1950–1960 (2013).
63. Tyanova, S., Temu, T. & Cox, J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat. Protoc.* **11**, 2301–2319 (2016).
64. Bober, M. & Miladinovic, S. General guidelines for validation of decoy models for HRM/DIA/SWATH as exemplified using Spectronaut. *F1000posters* https://f1000research.com/posters/1097512 (2015).
65. Carpenter, B. et al. Stan: a probabilistic programming language. *J. Stat. Softw.* **76**, 1–32 (2017).
66. Bhadra, A., Datta, J., Polson, N. G. & Willard, B. The Horseshoe+ estimator of ultra-sparse signals. *Bayesian Anal.* **12**, 1105–1131 (2017).
67. Goeminne, L. J. E., Gevaert, K. & Clement, L. Peptide-level robust ridge regression improves estimation, sensitivity, and specificity in data-dependent quantitative label-free shotgun proteomics. *Mol. Cell. Proteomics* **15**, 657–668 (2016).
68. Shannon, P. et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).
69. Ritchie, M. E. et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).
70. Merico, D., Isserlin, R., Stueker, O., Emili, A. & Bader, G. D. Enrichment map: a network-based method for gene-set enrichment visualization and interpretation. *PLoS ONE* **5**, e13984 (2010).
71. Meldal, B. H. M. et al. Complex Portal 2018: extended content and enhanced visualization tools for macromolecular complexes. *Nucleic Acids Res.* **47** (D1), D550–D558 (2019).
72. Giurgiu, M. et al. CORUM: the comprehensive resource of mammalian protein complexes-2019. *Nucleic Acids Res.* **47** (D1), D559–D563 (2019).
73. Tyanova, S. et al. The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods* **13**, 731–740 (2016).
74. Keenan, A. B. et al. ChEA3: transcription factor enrichment analysis by orthogonal omics integration. *Nucleic Acids Res.* **47** (W1), W212–W224 (2019).
75. Landt, S. G. et al. ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. *Genome Res.* **22**, 1813–1831 (2012).
76. Meng, E. C., Pettersen, E. F., Couch, G. S., Huang, C. C. & Ferrin, T. E. Tools for integrated sequence-structure analysis with UCSF Chimera. *BMC Bioinformatics* **7**, 339 (2006).
77. Pettersen, E. F. et al. UCSF Chimera–a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
78. Warnecke, A., Sandalova, T., Achour, A. & Harris, R. A. PyTMs: a useful PyMOL plugin for modeling common post-translational modifications. *BMC Bioinformatics* **15**, 370 (2014).
79. Paxman, J. J. & Heras, B. Bioinformatics tools and resources for analyzing protein structures. *Methods Mol. Biol.* **1549**, 209–220 (2017).
80. Jo, S., Vargyas, M., Vasko-Szedlar, J., Roux, B. & Im, W. PBEQ-Solver for online visualization of electrostatic potential of biomolecules. *Nucleic Acids Res.* **36**, W270–W275 (2008).
81. Vogt, C. et al. The interferon antagonist ML protein of thogoto virus targets general transcription factor IIB. *J. Virol.* **82**, 11446–11453 (2008).
82. Jorns, C. et al. Rapid and simple detection of IFN-neutralizing antibodies in chronic hepatitis C non-responsive to IFN-alpha. *J. Med. Virol.* **78**, 74–82 (2006).
83. Perez-Riverol, Y. et al. The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res.* **47** (D1), D442–D450 (2019).
84. Orchard, S. et al. The MIntAct project–IntAct as a common curation platform for 11 molecular interaction databases. *Nucleic Acids Res.* **42**, D358–D363 (2014).

Extended Data Fig. 1 | See next page for caption.

Extended Data Fig. 1 | SARS-CoV-2 and SARS-CoV proteins expressed in A549 cells target host proteins. a, Expression of HA-tagged viral proteins in stably transduced A549 cells, used in AP–MS and proteome expression measurements. When several bands are present in a single lane, * or ► mark the band with the expected molecular weight ($n = 4$ independent experiments). For gel source data, see Supplementary Fig. 1. **b**, Extended version of the virus-host protein–protein interaction network with 24 SARS-CoV-2 and 27 SARS-CoV proteins, as well as ORF3 of HCoV-NL63 and ORF4 and ORF4a of HCoV-229E, used as baits. Host targets regulated upon viral protein overexpression are highlighted (see the in-plot legend). **c**–**f**, Co-precipitation experiments in HEK293T cells showing a specific enrichment of endogenous MAVS co-precipitated with C-terminal HA-tagged ORF7b of SARS-CoV-2 and SARS-CoV (negative controls: SARS-CoV-2 ORF6-HA, ORF7a-HA) (**c**), ORF7b-HA of SARS-CoV-2 and SARS-CoV co-precipitated with SII-HA-UNC93B1 (control precipitation: SII-HA-RSAD2) (**d**), endogenous HSPA1A co-precipitated with N-HA of SARS-CoV-2 and SARS-CoV (control: SARS-CoV-2 ORF6-HA) (**e**) and endogenous TGF-β with ORF8-HA of SARS-CoV-2 vs ORF8-HA, ORF8a-HA, ORF8b-HA of SARS-CoV or ORF9b-HA of SARS-CoV-2 (**f**), ($n = 2$ independent experiments). For gel source data, see Supplementary Fig. 1. AP–MS: affinity-purification coupled to mass spectrometry; MD: Macro domain; NSP: Non-structural protein.
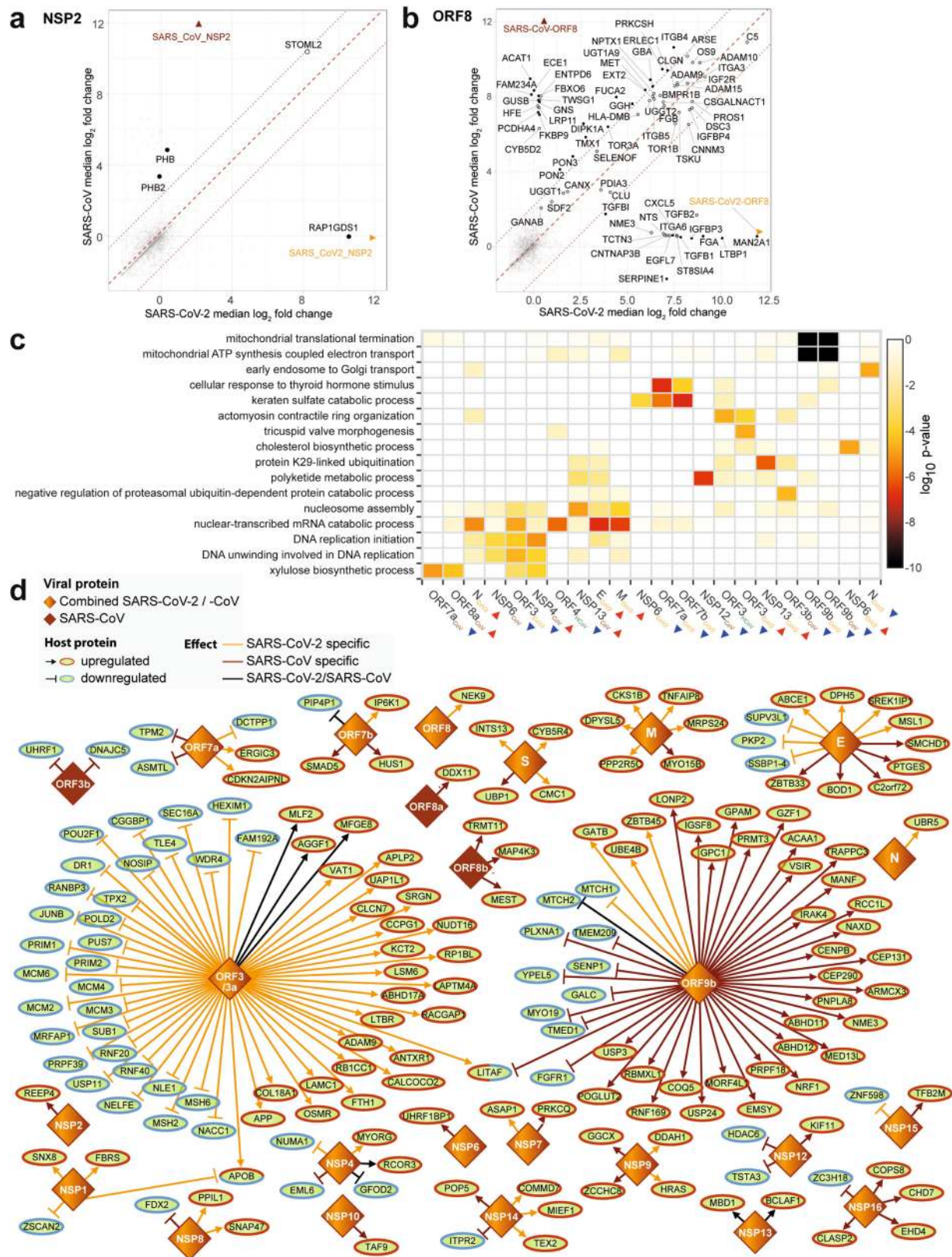
Extended Data Fig. 2 | SARS-CoV-2 and SARS-CoV proteins trigger shared and specific interactions with host factors, and induce changes to the host proteome. a, **b**, Differential enrichment of proteins in NSP2 (**a**) and ORF8 (**b**) of SARS-CoV-2 (*x*-axis) vs SARS-CoV (*y*-axis) AP–MS experiments (*n* = 4 independent experiments). **c**, Gene Ontology Biological Processes enriched among the cellular proteins that are up- (red arrow) or down- (blue arrow)

regulated upon overexpression of individual viral proteins. **d**, The most affected proteins from the effectome data of protein changes upon viral bait overexpression in A549 cells (see materials and methods for the exact protein selection criteria). Homologous viral proteins are displayed as a single node. Shared and virus-specific effects are denoted by the edge colour. NSP: Non-structural protein.
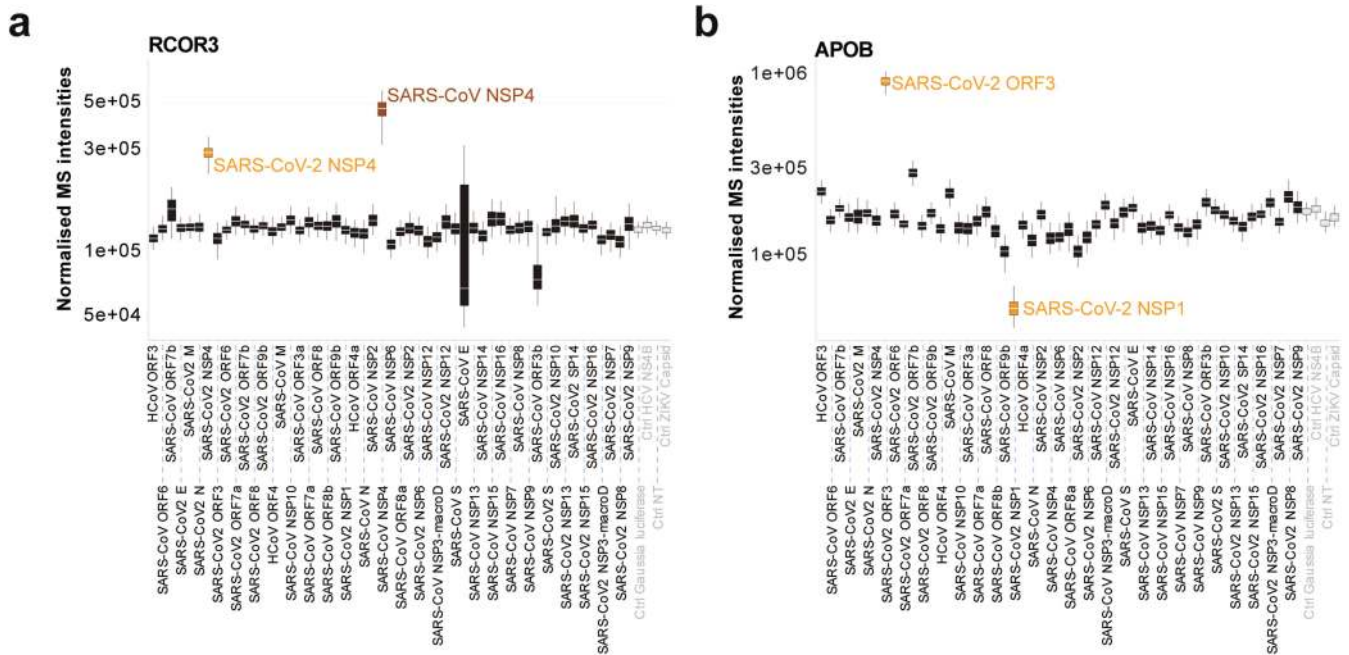
Extended Data Fig. 3 | RCOR3 and APOB regulation upon SARS-CoV-2 and SARS-CoV protein overexpression. a, b, Normalized intensities of selected candidates specifically perturbed by individual viral proteins: RCOR3 was upregulated both by SARS-CoV-2 and SARS-CoV NSP4 proteins (**a**), APOB was upregulated by ORF3 and downregulated by NSP1 specifically to SARS-CoV-2 (**b**). The box and the whiskers represent 50% and 95% confidence intervals, and the white line corresponds to the median of the $\log_2$ fold change upon viral protein overexpression (n = 4 independent experiments).
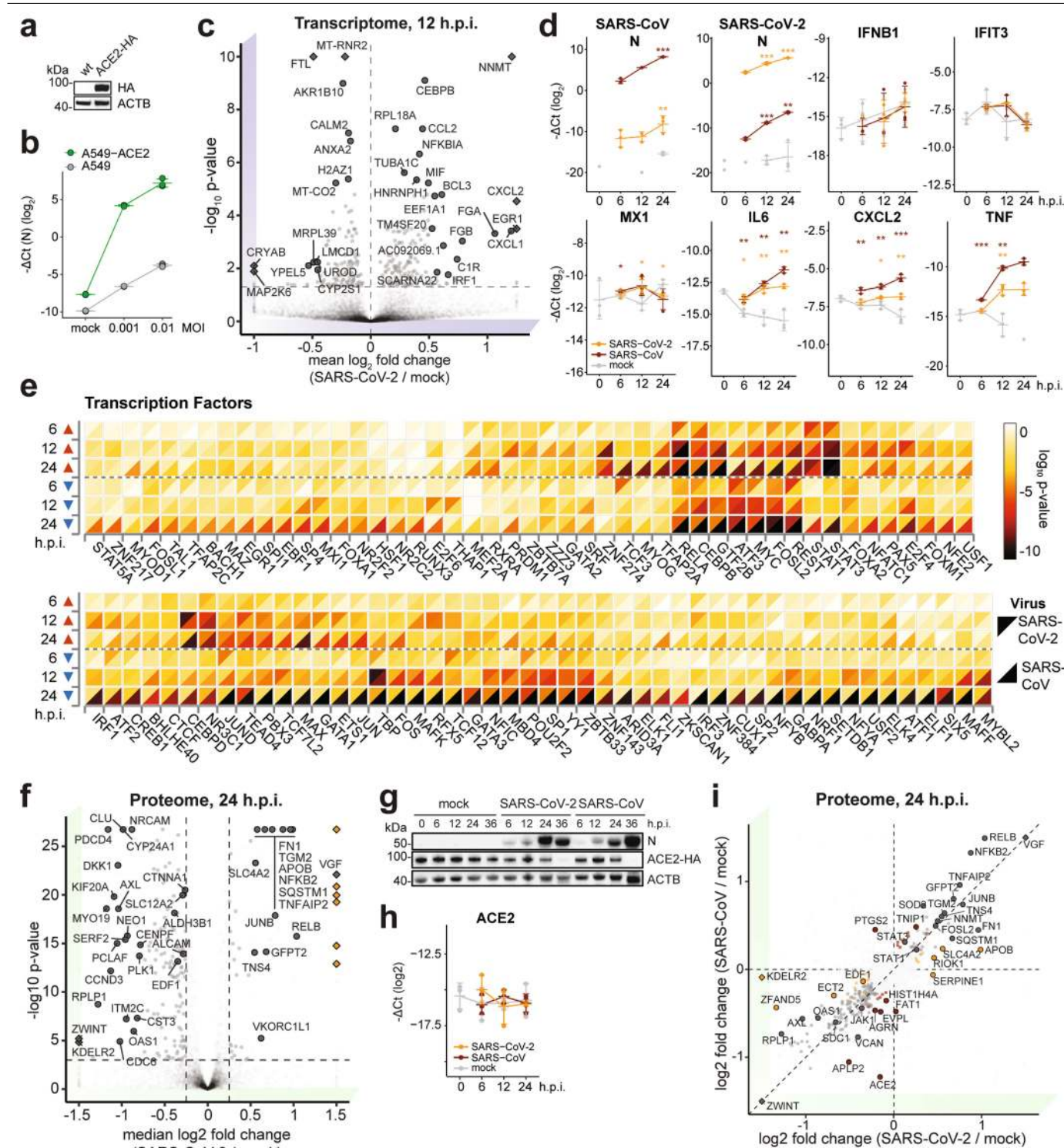
Extended Data Fig. 4 | See next page for caption.

Extended Data Fig. 4 | Tracking of virus-specific changes in infected A549-ACE2 cells by transcriptomics and proteomics. a, Western blot showing ACE2-HA expression levels in A549 cells untransduced (wild-type) or transduced with ACE2-HA-encoding lentivirus ($n = 2$ independent experiments). For gel source data, see Supplementary Fig. 1. **b**, mRNA expression levels of SARS-CoV-2 N relative to $RPLPO$ as measured by qRT–PCR upon infection of wild-type A549 and A549-ACE2 cells at the indicated MOIs. Error bars represent mean and standard deviation ($n = 3$ independent experiments). **c**, Volcano plot of mRNA expression changes of A549-ACE2 cells, infected with SARS-CoV-2 at an MOI of 2 in comparison to mock infection at 12 hpi. Significant hits are highlighted in grey (moderated t-test fasle discovery rate-corrected two-sided P-value, $n = 3$ independent experiments). Diamonds indicate that the actual log$_2$ fold change or P-value were truncated to fit into the plot. **d**, Expression levels, as measured by qRT–PCR, of SARS-CoV-2/SARS-CoV N and host transcripts relative to $RPLPO$ in infected (MOI of 2) A549-ACE2 cells with SARS-CoV-2 (orange) and SARS-CoV (brown) at indicated time points. Error bars correspond to mean and standard deviation (Two-sided student t-test, unadjusted P-value, $n = 3$ independent experiments). *P-value ≤ 0.05; **P-value ≤ 0.01; ***P-value ≤ 10^{-3}. **e**, Analysis of transcription factors, whose targets are significantly enriched among up- (red arrow) and down- (blue arrow) regulated genes of A549-ACE2 cells infected with SARS-CoV-2 (upper triangle) and SARS-CoV (lower triangle) for indicated time points (Fisher's exact test unadjusted one-sided P-value ≤ 10^{-4}). **f**, Volcano plot of SARS-CoV-2-induced protein abundance changes at 24 hpi in comparison to mock. Viral proteins are highlighted in orange, selected significant hits are marked in black (Bayesian linear model-based unadjusted two-sided P-value ≤ 10^{-3}, |median log$_2$ fold change| ≥ 0.25, $n = 4$ independent experiments). Diamonds indicate that the actual log$_2$ fold change was truncated to fit into the plot. **g**, Western blot showing the total levels of ACE2-HA protein at 6, 12, 24 and 36 hpi (mock, SARS-CoV-2 and SARS-CoV infections); N viral protein as infection and ACTB as loading controls ($n = 3$ independent experiments). For gel source data, see Supplementary Fig. 1. **h**, Stable expression of $ACE2$ mRNA transcript relative to $RPLPO$, as measured by qRT–PCR, after SARS-CoV-2 and SARS-CoV infections (MOI of 2) of A549-ACE2 cells at indicated hpi (error bars show mean and standard deviation, $n = 3$ independent experiments). **i**, Scatter plots comparing the host proteome of SARS-CoV-2 (x-axis) and SARS-CoV (y-axis) infection at 24 hpi (log$_2$ fold change in comparison to the mock infection samples at the same time point). Significantly regulated proteins (Bayesian linear model-based unadjusted two-sided P-value ≤ 10^{-3}, |log$_2$ fold change| ≥ 0.25, $n = 4$ independent experiments), are colored according to their specificity in both infections. Diamonds indicate that the actual log$_2$ fold change was truncated to fit into the plot.
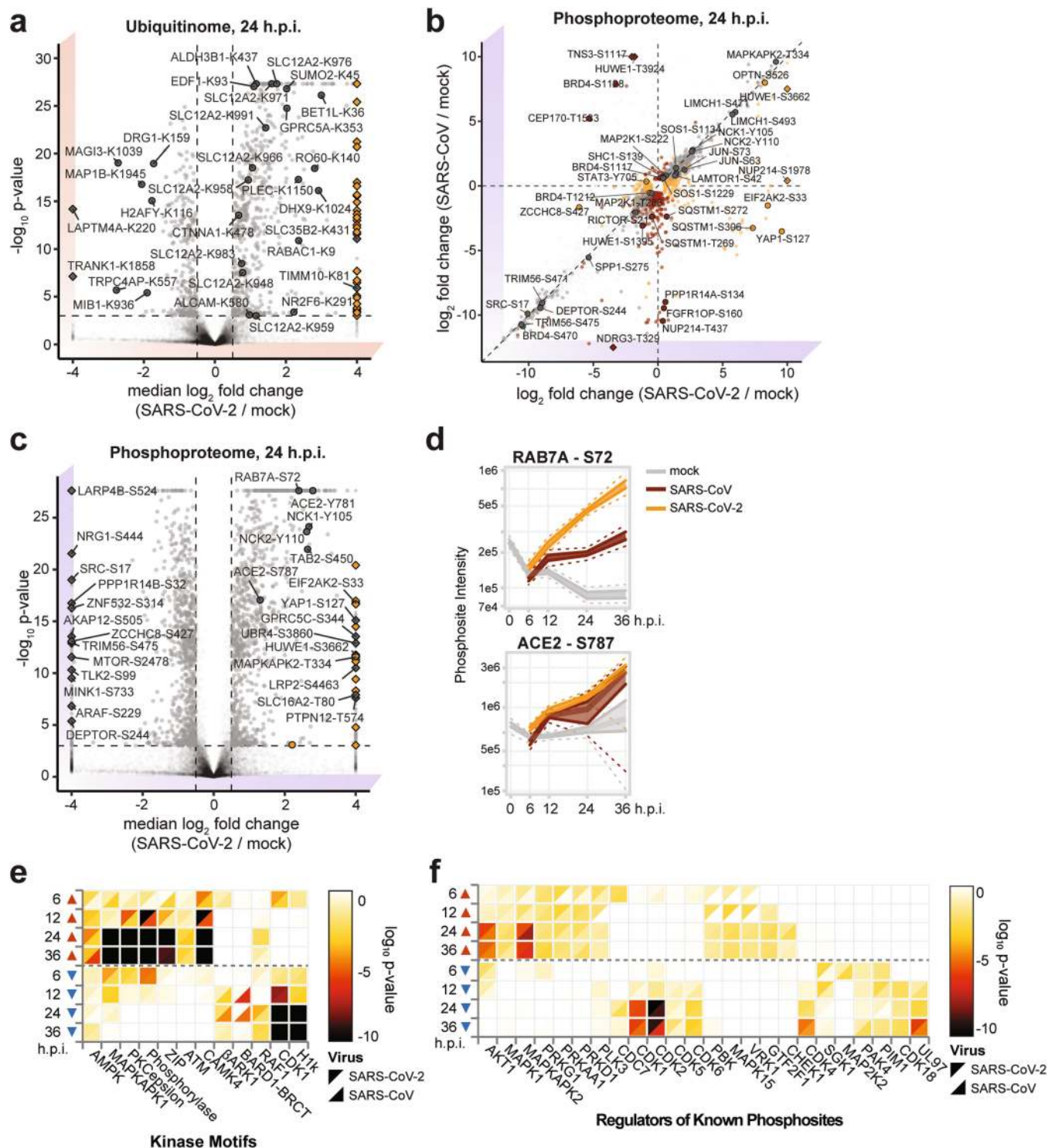
Extended Data Fig. 5 | Post-translational modifications modulated during SARS-CoV-2 or SARS-CoV infection. a, Volcano plots of SARS-CoV-2-induced ubiquitination changes at 24 hpi in comparison to mock. The viral PTM sites are highlighted in orange and selected significant hits in black. **b**, Scatter plots comparing the host phosphoproteome of SARS-CoV-2 (*x*-axis) and SARS-CoV (*y*-axis) infection at 24 hpi ($\log_2$ fold change in comparison to the mock infection samples at the same time point). Significantly regulated sites are colored according to their specificity in both infections. **c**, Volcano plots of SARS-CoV-2-induced phosphorylation changes at 24 hpi in comparison to mock. The viral PTM sites are highlighted in orange and selected significant hits in black. For **a**–**c**, a change is defined significant if its Bayesian linear model-based unadjusted two-sided *P*-value $\leq 10^{-3}$ and $|\log_2$ fold change$| \geq 0.5$, $n = 3$ independent experiments for ubiquitination and $n = 4$ independent

experiments for phosphorylation data. Diamonds in **a**–**c** indicate that the actual median $\log_2$ fold change was truncated to fit into the plot. **d**, Profile plots showing the time-resolved phosphorylation of ACE2 (S787) and RAB7A (S72) with indicated median, 50% and 95% confidence intervals, $n = 4$ independent experiments. **e**, The enrichment of host kinase motifs among the significantly regulated phosphorylation sites of SARS-CoV-2 (upper triangle) and SARS-CoV-infected (lower triangle) A549-ACE2 cells (MOI of 2) at the indicated time points (Fisher's exact test, unadjusted one-sided *P*-value $\leq 10^{-3}$). **f**, The enrichment of specific kinases among the ones known to phosphorylate significantly regulated sites at the indicated time points and annotated in PhosphoSitePlus database (Fisher's exact test, unadjusted one-sided *P*-value $\leq 10^{-2}$).

a VIM

b Ubiquitination, Phosphorylation, Phosphorylation, VIM - Protein Abundance

c Ubiquitination sites

d Membrane protein (M)

e Spike protein (S)

f Phosphorylation sites

g Nucleocapsid protein (N)

h Nucleocapsid protein (N) - C-terminal domain

Extended Data Fig. 6 | See next page for caption.

Article
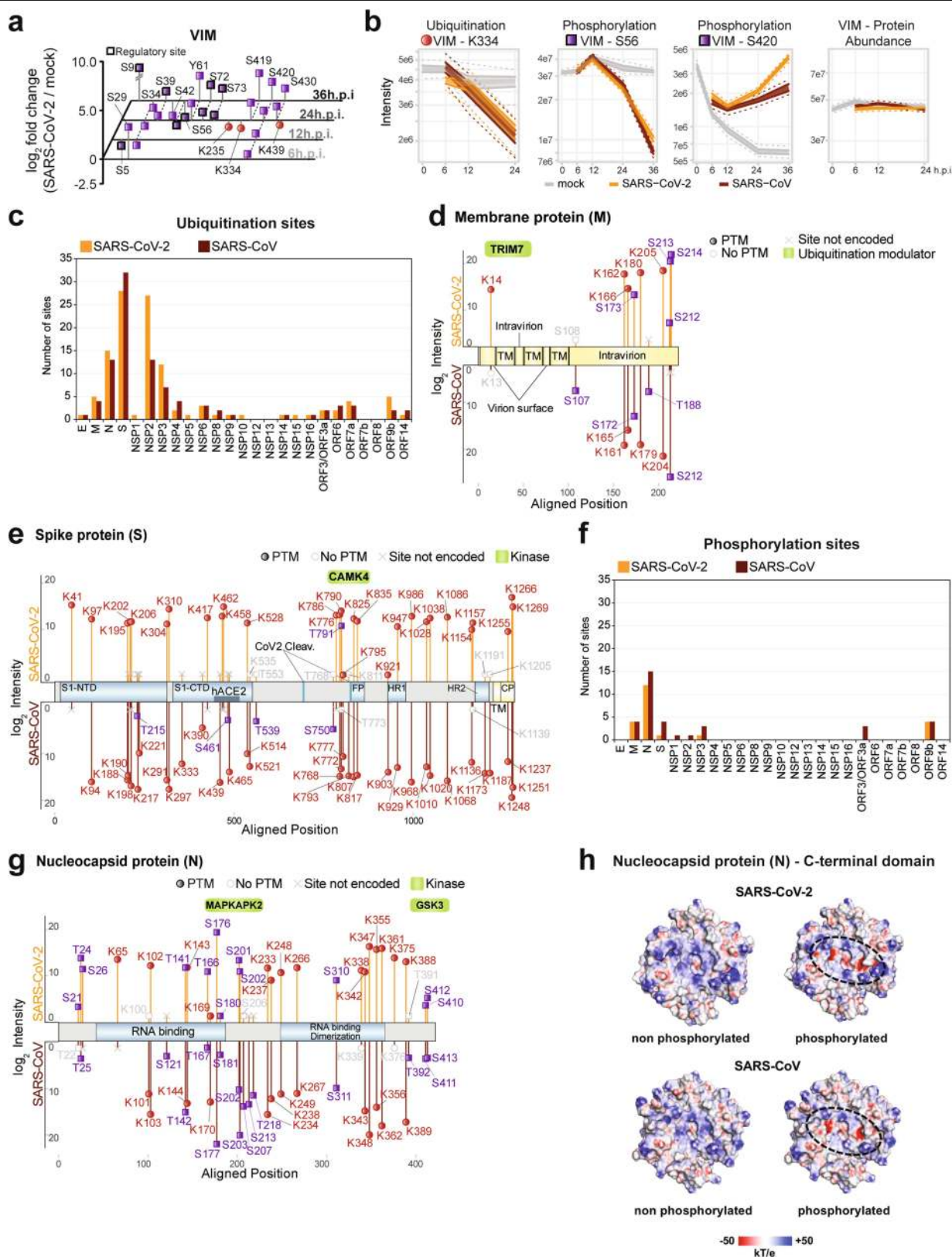
Extended Data Fig. 6 | Integration of multi-omics data from SARS-CoV-2 and SARS-CoV infection identified co-regulation of host and viral factors. **a**, Phosphorylation (purple square) and ubiquitination (red circles) sites on vimentin (VIM) regulated upon SARS-CoV-2 infection. The plot shows the medians of $\log_2$ fold changes compared to mock at 6, 12, 24 and 36 hpi, regulatory sites are indicated with a thick black border. **b**, Profile plots of VIM K334 ubiquitination, S56 and S72 phosphorylation, and total protein levels in SARS-CoV-2 or SARS-CoV infected A549-ACE2 cells at indicated times after infection, with indicated median, 50% and 95% confidence intervals, $n = 3$ (ubiquitination) or $n = 4$ (total protein levels, phosphorylation) independent experiments. **c**, Number of ubiquitination sites identified on each SARS-CoV-2 or SARS-CoV proteins in infected A549-ACE2 cells. **d**, **e**, Mapping the ubiquitination and phosphorylation sites of SARS-CoV-2/SARS-CoV M and S proteins on their aligned sequence showing median $\log_2$ intensities in infected A549-ACE2 cells at 24 hpi ($n = 4$ independent experiments for phosphorylation and $n = 3$ independent experiments for ubiquitination data) with functional (blue) and topological (yellow) domains highlighted. Ubiquitin modifying enzymes binding to both M proteins and the host kinases that potentially recognize motifs associated with the reported sites and overrepresented among cellular motifs enriched upon infection (Extended Data Fig. 5e, f) or interacting with given viral protein (Extended Data Fig. 1b) are indicated (green). **f**, Number of phosphorylation sites identified on each SARS-CoV-2 or SARS-CoV proteins in infected A549-ACE2 cells. **g**, Mapping the ubiquitination (red circle) and phosphorylation (purple square) sites of SARS-CoV-2/SARS-CoV N protein on their aligned sequence showing median $\log_2$ intensities in A549-ACE2 cells infected with the respective virus at 24 hpi ($n = 4$ independent experiments) with functional domains highlighted in blue. The host kinases that potentially recognize motifs associated with the reported sites and overrepresented among cellular motifs enriched upon infection (Extended Data Fig. 5e, f) or interacting with given viral protein (Extended Data Fig. 1b) (green). **h**, Electrostatic surface potential analysis of non-phosphorylated and phosphorylated SARS-CoV and SARS-CoV-2 N CTD dimers; red, white and blue regions represent areas with negative, neutral and positive electrostatic potential, respectively (scale from −50 to +50 kT e^{-1}). NTD, N-terminal domain; hACE2, binding site of human ACE2; FP, fusion peptide; HR1/2, Heptad region 1/2; CP, cytoplasmic region. CoV2 Cleav., SARS-CoV-2 cleavage sites.
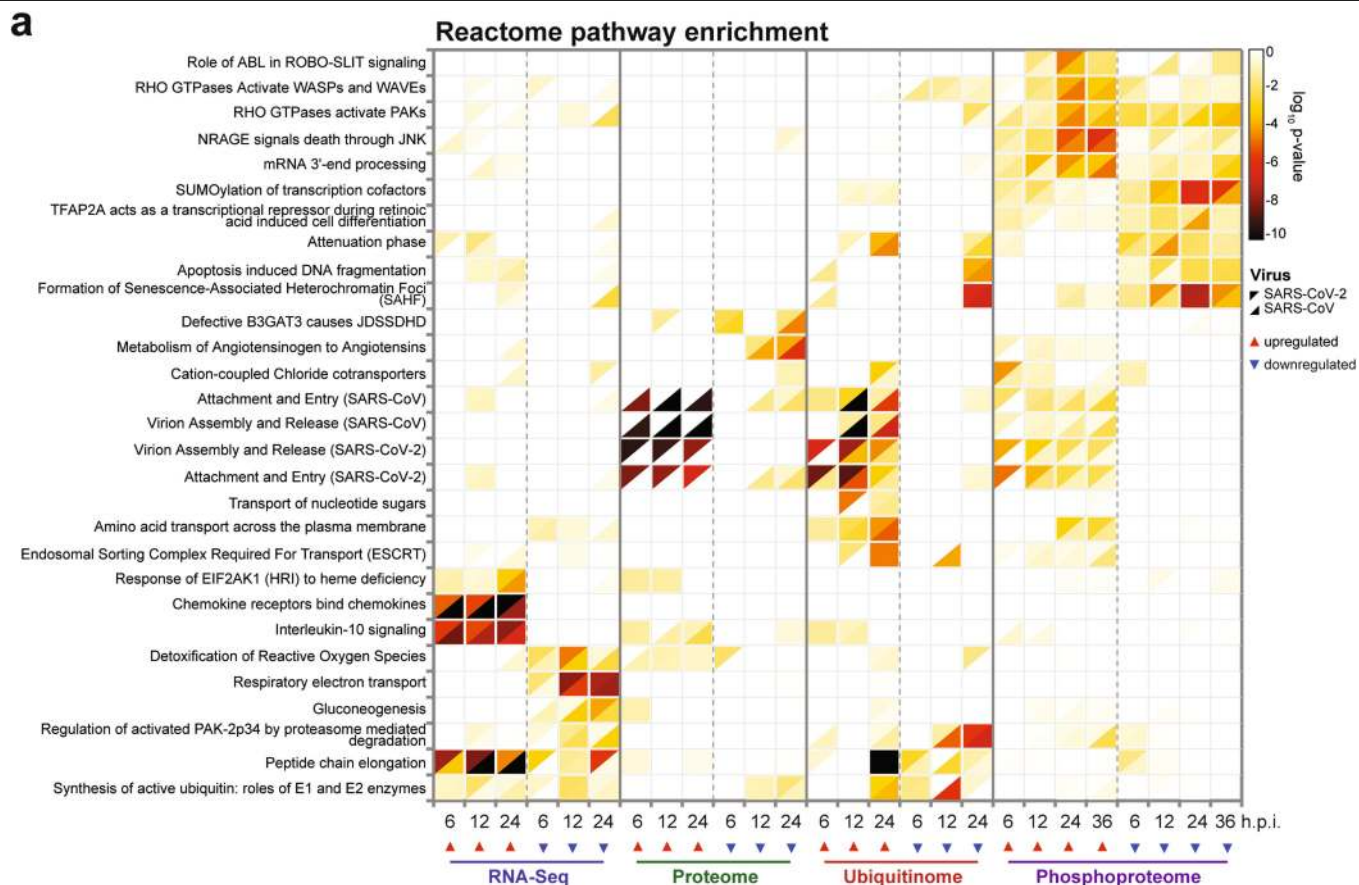
Extended Data Fig. 7 | Reactome pathways enrichment in multi-omics data of SARS-CoV-2 and SARS-CoV infection. a, Reactome pathways enriched in up- (red arrow) or downregulated (blue arrow) transcripts, proteins, ubiquitination and phosphorylation sites (Fisher's exact test unadjusted P-value $\leq 10^{-4}$) in SARS-CoV-2 or SARS-CoV-infected A549-ACE2 cells at indicated times after infection.

Extended Data Fig. 8 | See next page for caption.

Extended Data Fig. 8 | SARS-CoV-2 uses a multi-pronged approach to perturb host-pathways at several levels. a, The host subnetwork perturbed by SARS-CoV-2 M predicted by the network diffusion approach. Edge thickness reflects the transition probability in random walk with restart, directed edges represent the walk direction, and ReactomeFI connections are highlighted in black. **b**, Selection of the optimal threshold for the network diffusion model of SARS-CoV-2 M-induced proteome changes. The plot shows the relationship between the minimal allowed edge weight of the random walk graph (*x*-axis) and the mean inverse length of the path from the regulated proteins to the host targets of the viral protein along the edges of the resulting filtered subnetwork (*y*-axis). The red curve represents the metric for the network diffusion analysis of the actual data. The grey band shows 50% confidence interval, and dashed lines correspond to 95% confidence interval for the average inverse path length distribution for 1,000 randomized datasets. Optimal edge weight threshold that maximizes the difference between the metric based on the real data and its 3rd quartile based on randomized data are highlighted by the red vertical line. **c**, **d**, Subnetworks of the network diffusion predictions linking host targets of SARS-CoV-2 ORF7b (**c**) to the factors involved in innate immunity and ORF8 (**d**) to the factors involved in TGF-β signalling. **e**, **f**, Western blot showing the accumulation of the autophagy-associated factor MAP1LC3B upon SARS-CoV-2 ORF3 expression in HEK293-R1 cells (*n* = 3 independent experiments) (**e**) and SARS-CoV-2/SARS-CoV infection of A549-ACE2 cells (*n* = 3 independent experiments) (**f**). For gel source data, see Supplementary Fig. 1. **g**, **h**, Profile plots showing the time-resolved ubiquitination of the autophagy regulators MAP1LC3A, GABARAP, VPS33A and VAMP8 (*n* = 3 independent experiments) (**g**), as well as an increase in total protein abundance of APOB with indicated median, 50% and 95% confidence intervals (*n* = 4 independent experiments) (**h**). **i**, Overview of perturbations to host-cell innate immunity-related pathways, induced by distinct proteins of SARS-CoV-2, derived from the network diffusion model and overlaid with transcriptional, ubiquitination and phosphorylation changes upon SARS-CoV-2 infection. **j**, Heat map showing the effects of the indicated SARS-CoV-2 proteins on type-I IFN expression levels, ISRE and GAS promoter activation in HEK293-R1. Accumulation of type-I IFN in the supernatant was evaluated by testing supernatants of PPP-RNA (IVT4) stimulated cells on MX1-luciferase reporter cells, ISRE promoter activation−by luciferase assay after IFN-α stimulation, and GAS promoter activation−by luciferase assay after IFN-γ stimulation in cells expressing SARS-CoV-2 proteins as compared to the controls (ZIKV NS5 and SMN1) (*n* = 3 independent experiments).

Article
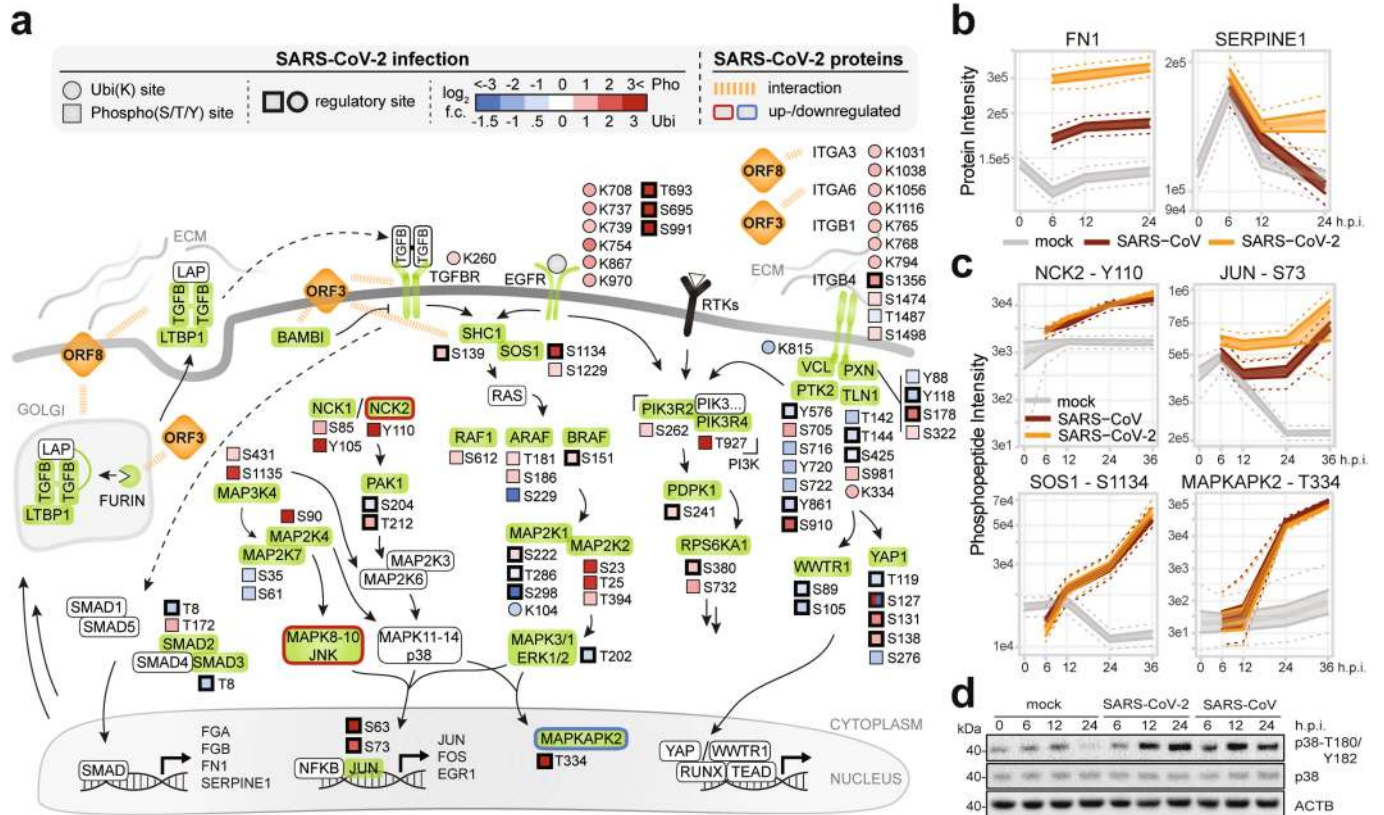
Extended Data Fig. 9 | Perturbation of host integrin-TGF-β-EGFR-receptor tyrosine kinase signalling by SARS-CoV-2. a, Overview of perturbations to host-cell Integrin-TGF-β-EGFR-receptor tyrosine kinase signalling, induced by distinct proteins of SARS-CoV-2, derived from the network diffusion model and overlaid with transcriptional, ubiquitination and phosphorylation changes upon SARS-CoV-2 infection. **b**, Profile plots of total protein levels of SERPINE1 and FN1 in SARS-CoV-2 or SARS-CoV-infected A549-ACE2 cells at 6, 12, and 24 hpi, with indicated median, 50% and 95% confidence intervals (*n* = 4

independent experiments). **c**, Profile plots showing intensities of indicated phosphosites on NCK2, JUN, SOS1 and MAPKAPK2 in SARS-CoV-2 or SARS-CoV-infected A549-ACE2 cells at 6, 12, 24 and 36 hpi, with indicated median, 50% and 95% confidence intervals (*n* = 4 independent experiments). **d**, Western blot showing phosphorylated (T180/Y182) and total protein levels of p38 in SARS-CoV-2 or SARS-CoV infected A549-ACE2 cells (*n* = 3 independent experiments). For gel source data, see Supplementary Fig. 1.
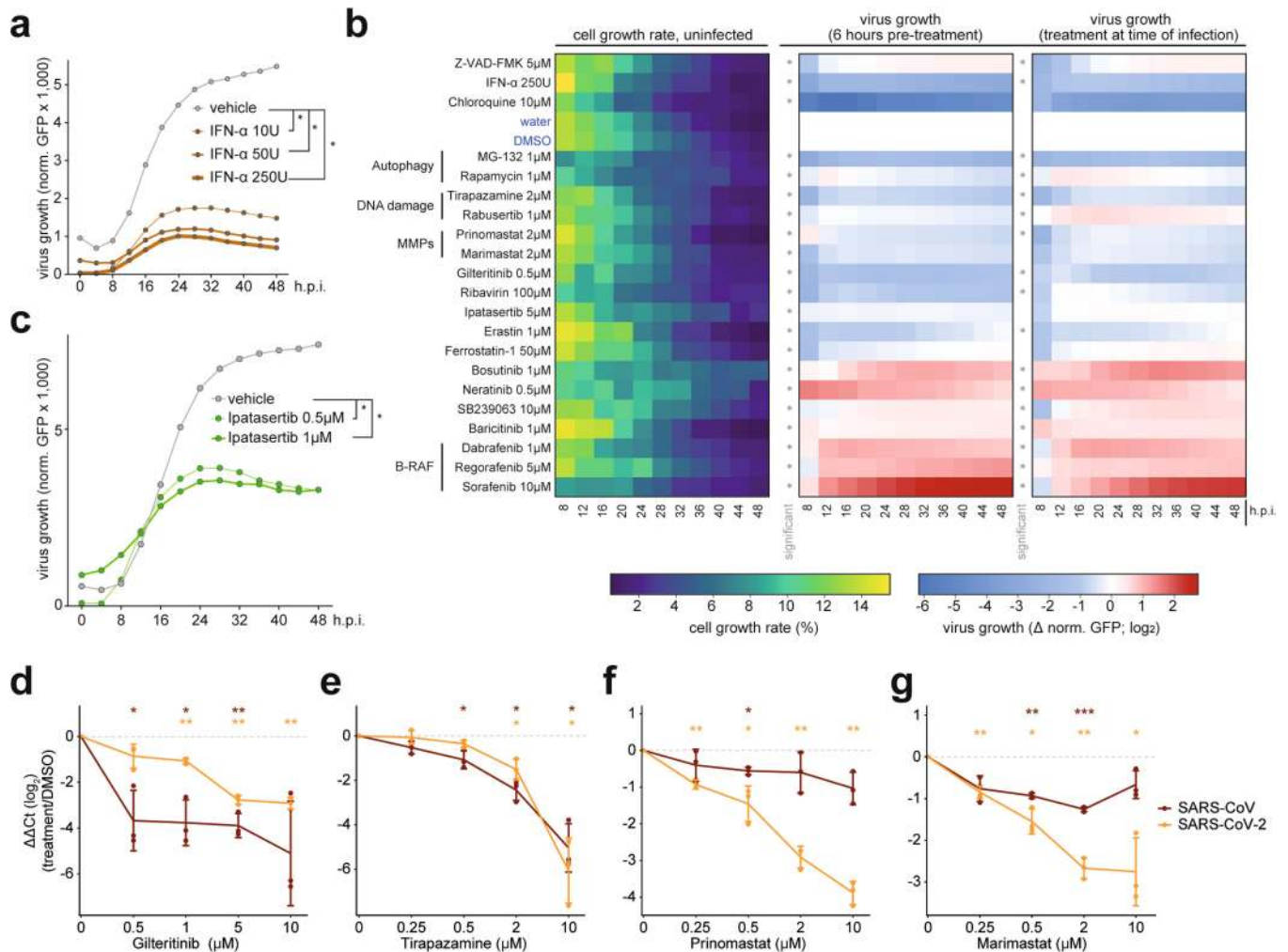
Extended Data Fig. 10 | Drug repurposing screen, focusing on pathways perturbed by SARS-CoV-2, reveals potential candidates for use in antiviral therapy. a, A549-ACE2 cells exposed for 6 h to the specified concentrations of IFN-α and infected with SARS-CoV-2-GFP reporter virus (MOI of 3). GFP signal and cell confluency were analysed by live-cell imaging for 48 hpi. Time-courses show virus growth over time as the mean of GFP-positive area normalized to the total cell area (n = 4 independent experiments). **b**, A549-ACE2 cells were pre-treated for 6 h or treated at the time of infection with SARS-CoV-2-GFP reporter virus (MOI of 3). GFP signal and cell growth were tracked for 48 hpi by live-cell imaging using an Incucyte S3 platform. Left heat map: the cell growth rate (defined as the change of cell confluence between t_i and t_{i-1} timepoints divided by cell confluence at t_{i-1}) over time in drug-treated uninfected conditions. Middle (6 h of pre-treatment) and right (treatment at the time of infection) heat maps: treatment-induced changes in virus growth over time (GFP signal normalized to total cell confluence $\log_2$ fold change between the treated and control (water, DMSO) conditions). Only non-cytotoxic treatments with significant effects on SARS-CoV-2-GFP are shown. Asterisks indicate significance of the difference to the control treatment (Wilcoxon test; unadjusted two-sided P-value ≤ 0.05, n = 4 independent experiments). **c**, A549-ACE2 cells exposed for 6 h to the specified concentrations of ipatasertib and infected with SARS-CoV-2-GFP reporter virus (MOI of 3). GFP signal and cell confluency were analysed by live-cell imaging for 48 hpi. Time-courses show virus growth over time as the mean of GFP-positive area normalized to the total cell area (n = 4 independent experiments). **d–g**, mRNA expression levels at 24 hpi of SARS-CoV-2 (orange) and SARS-CoV (brown) N relative to $RPLPO$, compared to DMSO-treated cells, as measured by qRT–PCR in infected A549-ACE2 cells (MOI of 1) pre-treated for 6 h with gilteritinib (**d**), tirapazamine (**e**), prinomastat (**f**) or marimastat (**g**). Error bars represent mean and standard deviation (Student t-test, two-sided, unadjusted P-value, n = 3 independent experiments). *P-value ≤ 0.05; **P-value ≤ 0.01; ***P-value ≤ 10^{-3}.

Article

Extended Data Table 1 | Functional annotations of the protein–protein interaction network of SARS-CoV-2 and SARS-CoV (AP–MS)

annotation_label	annotation_category	annotation_genes
Cell adhesion and motility	cellular_process	AMIGO2 CDH16 CDH17 CLDN12 DSC3 EPCAM FAT1 LRFN4 NECTIN2 NECTIN3 PCDH9 PCDHA12 PCDHA4 PCDHAC2 PCDHGC3 PTPRF PTPRS PVR NRP2 PLXNA1 PLXND1 SEMA4B SEMA4C
Endolysosomal trafficking	cellular_process	RAB13 RAB14 RAB1A RAB21 RAB2A RAB31 RAB32 RAB34 RAB3D RAB5A RAB5B RAB7A RAB8A RAB9A
ER quality control	cellular_process	CANX ERLEC1 FBXO6 OS9 UGGT1 UGGT2
ER stress	cellular_process	HSPA1A HSPA2 HSPA6 HSPA8 HSPA9 HSPH1 G3BP1 G3BP2 CAPRIN1
ER to cytosol trafficking	cellular_process	FAF2 NPLOC4 UFD1
ER-Golgi protein trafficking	cellular_process	AREG KDELR1 LMAN1 LMAN2 PIEZO1 TMED2 TMED7 TMED9 TMEM199 ARFIP1 SCAMP1 SCAMP2 SCAMP3 SCAMP4 CUX1 GOLIM4
Glycolysis	cellular_process	L2HGDH OGDH PDHX PDPR
Glycolysis	cellular_process	ACO2 FH MDH1
GPI anchor	cellular_process	GPAA1 PIGS PIGU
Ion transport by ATPases	cellular_process	ATP11C ATP12A ATP13A1 ATP13A3 ATP2A3 ATP2B4 ATP6AP1 ATP6V0A2 ATP6V1B1 ATP7B ATP8B1 ATP8B2
Lipid oxidation	cellular_process	ACAD10 ACADS ACSF2 PCCA PCCB ECI1
mRNA processing	cellular_process	HNRNPM MYEF2 DICER1 TARBP2 MBNL1
Nuclear import/export	cellular_process	IPO8 TNPO1 TNPO2 XPO5 XPO6 XPO7 XPOT
Oxidoreduction	cellular_process	ALDH2 ALDH5A1
Glycosylation	cellular_process	B4GALT7 POMGNT1 ALG11 ALG13 ALG14 B3GALT6 B3GAT3 EXT1 EXTL2 EXTL3 GLCE XXYLT1 DAD1 TMEM258 GALNT1 GALNT10 GALNT12 ALG5 ALG8 FUT8 LMAN1 OSTC STT3A
Glycosylation	cellular_process	FUCA2 GANAB GBA GUSB
Palmitoylation	cellular_process	SELENOK ZDHHC20 SPTLC2 ZDHHC13 ZDHHC18 ZDHHC21 ZDHHC3 ZDHHC6 ZDHHC9 GOLGA7 ZDHHC5
Transcription elongation	cellular_process	GTF2F2 SETD2
tRNA charging	cellular_process	IARS2 NARS2 PPA2 SARS2 TARS2 HARS2
tRNA splicing	cellular_process	FAM98A RTCB RTRAF
Ubiquitin-like ligase activity	cellular_process	MGRN1 RNF130 RNF149 RNF19A STUB1 WWP1 WWP2 ZNRF3 HUWE1 MDM2 TRIM47
ATP synthase	complex_compartment	ATP5F1B ATP5F1D ATP5F1E ATP5PB ATP5PD MT-ATP6 ATP5PF
COG complex	complex_compartment	COG1 COG2 COG3 COG4 COG5 COG6 COG7 COG8
Condensin II complex	complex_compartment	NCAPD3 NCAPH2 NCAPG2
ECM regulators and metalloproteases	complex_compartment	ADAM17 ADAM9 CLTRN CNDP2 CPD ECE1 MMP15 RNPEP ADAM10 ADAM15
Endocytosis via AP-2 complex	complex_compartment	AP2A1 AP2M1 AP2S1 EPN2
ER membrane protein complex	complex_compartment	EMC10 EMC2 EMC3 EMC4 EMC8
Golgi membrane	complex_compartment	B4GAT1 CSGALNACT1 ENTPD4 QSOX1 QSOX2 SAMD8 STEAP2 TVP23C
HOPS complex	complex_compartment	HOOK3 VPS11 VPS16 VPS18 VPS39 VPS41
Integrator complex	complex_compartment	INTS1 INTS12 INTS2 INTS4 INTS5 INTS8
Integrins	complex_compartment	ITGA3 ITGB4 ITGB5
MHC-I complex	complex_compartment	B2M HLA-A HLA-C HLA-E HLA-G HFE
Mitochondrial metalloproteases	complex_compartment	NLN PITRM1 PMPCA PMPCB
Mitochondrial respiratory chain	complex_compartment	NDUFA10 NDUFS2 NDUFS8
Nuclear inner membrane	complex_compartment	DPY19L2 DPY19L3 DPY19L4 LEMD3 PSEN2 ZMPSTE24
Nuclear pore	complex_compartment	NUP188 NUP205 NUP93
Peroxisome	complex_compartment	GNPAT MAVS MGST1 PEX10 PEX13 PEX2
Proteasome core	complex_compartment	PSMA4 PSMA5
Proteasome regulatory proteins	complex_compartment	PSMC2 PSMC4 PSMC5 PSMD11 PSMD12 PSMD4 PSME3
Sarcoglycan complex	complex_compartment	SGCB SGCD SGCE
Septin complex	complex_compartment	SEPTIN10 SEPTIN11 SEPTIN2 SEPTIN7 SEPTIN8 SEPTIN9
SNARE complex	complex_compartment	BET1 GOSR1 GOSR2 NAPA NAPG SNAP25 STX10 STX12 STX16 STX2 STX4 STX5 STX6 STX7 VAMP2 VAMP3 VAMP4 VAMP7 VTI1A
Solute carriers	complex_compartment	SLC12A4 SLC12A6 SLC12A7 SLC15A4 SLC16A4 SLC18B1 SLC19A2 SLC20A1 SLC22A5 SLC23A2 SLC25A2 SLC25A52 SLC12A9 SLC26A2 SLC29A3 SLC25A24 SLC2A6 SLC29A4 SLC30A1 SLC35D2 SLC35F5 SLC30A5 SLC33A1 SLC35A1 SLC35A2 SLC23A1 SLC30A7 SLC35F2 SLC35F6 SLC39A1 SLC39A14 SLC6A6 SLC7A6 SLC35B4 SLC37A4 SLC38A2 SLC45A1 SLC46A1 SLC47A2 SLC4A10 SLC4A2 SLC4A4 SLC6A15 SLC9A1
Cytokine receptors signaling	signaling	CD44 IFNGR1 IL10RB IL13RA1 IL6ST OSMR JAK1 ACVR1 ACVR1B ACVR2A BAMBI BMPR1A BMPR2 FKBP1A TGFBR1 TGFBR2 EIF2A FKBP1A SHC1
EphrinB-EPHB pathway	signaling	EPHB2 EPHB3
ErbB receptor signaling	signaling	ERBB2 ERBB3 NRG1
GPCRs signaling	signaling	GPR39 OPN3 S1PR3 S1PR5 GNA13
Inflammatory response	signaling	AHR AXL CD70 DCBLD2 IFITM1 LDLR LPAR1 SELENOS TNFSF15 TNFSF9 TPBG
NEDD4-ITCH complex	signaling	NDFIP1 NDFIP2 ITCH
Notch signaling	signaling	NOTCH1 NOTCH2 NOTCH3
p53 signaling	signaling	FAS TP53 BNIP3L EPHA2 FAS STEAP3 MET NDRG1 MDM2 IGF2R BAG3
Receptor tyrosine phosphatases	signaling	PTPMT1 PTPN11 PTPRA PTPRF PTPRJ PTPRM PTPRS
TGF-β and integrins signaling	signaling	FGA FGB PROS1 SERPINE1 TGFB1 TGFB2 LTBP1 TGFBI IGFBP3 IGFBP4 SERPINE1 PLAU
TNF receptors superfamily	signaling	TNFRSF10A TNFRSF10B TNFRSF10D TNFRSF1A

Proteins identified as SARS-CoV-2 and/or SARS-CoV host binders via AP–MS (Fig. 1b) grouped based on functional enrichment analysis of GOBP, GPCC, GPMF and Reactome terms (Supplementary Table 2).

Corresponding author(s): Andreas Pichlmair

Last updated by author(s): Feb 15, 2021

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted *Give P values as exact values whenever suitable.*
☐	☒	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's d, Pearson's r), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	MS data acquisition: XCalibur software (v.3.0 (DDA) and v.4.4 (DIA); Thermo Fisher) Peptide identification and quantification: MaxQuant (v1.6.14; https://maxquant.org), Spectronaut (v.13&14; Biognosys, commercial), maxquantUtils R package (https://doi.org/10.5281/zenodo.4536603), in-house Julia scripts (https://doi.org/10.5281/zenodo.4541090), Live imaging data: IncuCyte software (v2019b, Sartorius; commercial) Sequencing data: Dropseq (v1.12, http://mccarrolllab.org/dropseq/)
Data analysis	Statistical analysis of MS data: R (v3.6.0), Julia(v.1.5), Stan (v2.19; https://mc-stan.org), msglm R package (https://doi.org/10.5281/zenodo.4536605), Perseus (v1.6.14.0), RStudio (v1.2.1335). Statistical analysis of RNA-Seq data: R (v3.4.4), Limma R package (v3.46.0). Network Diffusion Analysis: HierarchicalHotNet.jl Julia package (https://doi.org/10.5281/zenodo.4536590). Gene Set Enrichment Analysis: OptEnrichedSetCover.jl Julia package (https://doi.org/10.5281/zenodo.4536596), VegaLite.jl package (https://github.com/queryverse/VegaLite.jl). Network and pathway visualization: yEd (v.3.20, yWorks), cytoscape (v3.8.1), Ingenuity Pathway Analysis (https://www.qiagenbioinformatics.com/products/ingenuity-pathway-analysis; Qiagen). Viral PTMs alignment: BioAlignments.jl Julia package (v.2.0, https://github.com/BioJulia/BioAlignments.jl). Structural alignment and visualisation: Chimera (v1.4) - MatchMaker tool, PyMOL (v2.4). Electrostatic surface potential: PBEQ Solver tool - CHARMM-GUI server (http://www.charmm-gui.org/?doc=input/pbeqsolver). A collection of in-house R, Julia and Python scripts (https://doi.org/10.5281/zenodo.4541082).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about <u>availability of data</u>

All manuscripts must include a <u>data availability statement</u>. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The raw sequencing data for this study have been deposited in the European Nucleotide Archive (ENA) at EMBL-EBI under accession number PRJEB38744 (https://www.ebi.ac.uk/ena/data/view/PRJEB38744).
The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD022282, PXD020461 and PXD020222.
The protein interactions from this publication have been submitted to the IMEx (http://www.imexconsor-tium.org) consortium through IntAct with the identifier IM-28109.

The following public data sets were used in the study:
- Gene Ontology and Reactome annotations (http://download.baderlab.org/EM_Genesets/April_01_2019/Human/UniProt/Human_GO_AllPathways_with_GO_iea_April_01_2019_UniProt.gmt),
- IntAct Protein Interactions (https://www.ebi.ac.uk/intact/, v2019.12),
- IntAct Protein Complexes (https://www.ebi.ac.uk/complexportal/home, v2019.12),
- CORUM Protein Complexes (http://mips.helmholtz-muenchen.de/corum/download/allComplexes.xml.zip, v2018.3),
- Reactome Functional Interactions (https://reactome.org/download/tools/ReatomeFIs/FIsInGene_020720_with_annotations.txt.zip),
- Human (v2019.10), Human-CoV, SARS-CoV-2 and SARS-CoV (v2020.08) protein sequences: https://uniprot.org.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences   ☐ Behavioural & social sciences   ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see <u>nature.com/documents/nr-reporting-summary-flat.pdf</u>

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	The sample sizes were chosen from past knowledge on the good sample size to ensure adequate power. Sample sizes are always indicated in figure legends or related "Methods" section.
Data exclusions	Due to low number of protein identifications, 1 out of 304 MS runs, and 6 out of 276 MS raw files were excluded from the statistical analysis of AP-MS and DIA viral protein overexpression data, respectively, reducing the number of replicates for specific conditions to n=3.
Replication	For Mass spectrometry, in vitro viral replication experiments, (co-IP-)WB analysis, reporter assay and IFN bioassay, a minimum of three biological experiments were performed independently. All replications were successful.
Randomization	N/A. No randomization was used given the small number of samples and the lack of influence of randomization on the experimental design and experimental approach used. (no animal experiments were performed in this study).
Blinding	N/A. Investigators were not blinded to experimental groups (in vitro experiments required prior knowledge for data interpretation).

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	*Briefly describe the study type including whether data are quantitative, qualitative, or mixed-methods (e.g. qualitative cross-sectional, quantitative experimental, mixed-methods case study).*
Research sample	*State the research sample (e.g. Harvard university undergraduates, villagers in rural India) and provide relevant demographic information (e.g. age, sex) and indicate whether the sample is representative. Provide a rationale for the study sample chosen. For studies involving existing datasets, please describe the dataset and source.*
Sampling strategy	*Describe the sampling procedure (e.g. random, snowball, stratified, convenience). Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient. For qualitative data, please indicate whether data saturation was considered, and what criteria were used to decide that no further sampling was needed.*
Data collection	*Provide details about the data collection procedure, including the instruments or devices used to record the data (e.g. pen and paper,*

Data collection	computer, eye tracker, video or audio equipment) whether anyone was present besides the participant(s) and the researcher, and whether the researcher was blind to experimental condition and/or the study hypothesis during data collection.
Timing	Indicate the start and stop dates of data collection. If there is a gap between collection periods, state the dates for each sample cohort.
Data exclusions	If no data were excluded from the analyses, state so OR if data were excluded, provide the exact number of exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.
Non-participation	State how many participants dropped out/declined participation and the reason(s) given OR provide response rate OR state that no participants dropped out/declined participation.
Randomization	If participants were not allocated into experimental groups, state so OR describe how participants were allocated to groups, and if allocation was not random, describe how covariates were controlled.

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	Briefly describe the study. For quantitative data include treatment factors and interactions, design structure (e.g. factorial, nested, hierarchical), nature and number of experimental units and replicates.
Research sample	Describe the research sample (e.g. a group of tagged Passer domesticus, all Stenocereus thurberi within Organ Pipe Cactus National Monument), and provide a rationale for the sample choice. When relevant, describe the organism taxa, source, sex, age range and any manipulations. State what population the sample is meant to represent when applicable. For studies involving existing datasets, describe the data and its source.
Sampling strategy	Note the sampling procedure. Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient.
Data collection	Describe the data collection procedure, including who recorded the data and how.
Timing and spatial scale	Indicate the start and stop dates of data collection, noting the frequency and periodicity of sampling and providing a rationale for these choices. If there is a gap between collection periods, state the dates for each sample cohort. Specify the spatial scale from which the data are taken
Data exclusions	If no data were excluded from the analyses, state so OR if data were excluded, describe the exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.
Reproducibility	Describe the measures taken to verify the reproducibility of experimental findings. For each experiment, note whether any attempts to repeat the experiment failed OR state that all attempts to repeat the experiment were successful.
Randomization	Describe how samples/organisms/participants were allocated into groups. If allocation was not random, describe how covariates were controlled. If this is not relevant to your study, explain why.
Blinding	Describe the extent of blinding used during data acquisition and analysis. If blinding was not possible, describe why OR explain why blinding was not relevant to your study.

Did the study involve field work? ☐ Yes ☐ No

Field work, collection and transport

Field conditions	Describe the study conditions for field work, providing relevant parameters (e.g. temperature, rainfall).
Location	State the location of the sampling or experiment, providing relevant parameters (e.g. latitude and longitude, elevation, water depth).
Access & import/export	Describe the efforts you have made to access habitats and to collect and import/export your samples in a responsible manner and in compliance with local, national and international laws, noting any permits that were obtained (give the name of the issuing authority, the date of issue, and any identifying information).
Disturbance	Describe any disturbance caused by the study and how it was minimized.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	HA-HRP (Sigma-Aldrich; H6533; 1:2500 dilution), ACTB-HRP (Santa Cruz; sc-47778; 1:5000 dilution), MAP1LC3B (Cell Signaling; 3868; 1:1000 dilution), MAVS (Cell Signaling; 3993; 1:1000 dilution), HSPA1A (Cell Signaling; 4873; 1:1000 dilution), TGFβ (Cell Signaling; 3711; 1:1000 dilution), phospho-p38 (T180/Y182) (Cell Signaling; 4511; 1:1000 dilution), p38 (Cell Signaling; 8690; 1:1000 dilution) and SARS-CoV-2/SARS-CoV N protein (Sino Biological; 40143-MM05; 1:1000 dilution) antibodies were used Secondary antibodies detecting mouse (Cell Signaling; 7076; 1:5000 dilution/Jackson ImmunoResearch; 115-035-003; 1:5000 dilution), rat (Invitrogen; 31470; 1:5000 dilution), and rabbit IgG (Cell Signaling; 7074; 1:5000 dilution) were horseradish peroxidase (HRP)-coupled.
Validation	Antibodies were validated by either knock-down (MAP1LC3B, MAVS, ACTB), over-expression (HA-HRP, TGFβ), or infection experiments in this study (SARS-CoV-2/SARS-CoV N protein). Additionally, antibodies have been validated by the manufacturers.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	Vero E6 cells (CRL-1586) were purchased from ATCC. A549 and HEK293T cells were a gift from Georg Kochs (Universitäts Klinikum, Freiburg). HEK293R1 cells were a gift from Andrew Bowie (Trinity College, Dublin)
Authentication	The identity of all the immortalized cell lines used in this study was confirmed by STR-profiling (Eurofins Medigenomix). Official certification can be provided upon request.
Mycoplasma contamination	All cell lines were tested to be mycoplasma free by standard PCR-based assay.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell line was used in this study.

Palaeontology and Archaeology

Specimen provenance	*Provide provenance information for specimens and describe permits that were obtained for the work (including the name of the issuing authority, the date of issue, and any identifying information).*
Specimen deposition	*Indicate where the specimens have been deposited to permit free access by other researchers.*
Dating methods	*If new dates are provided, describe how they were obtained (e.g. collection, storage, sample pretreatment and measurement), where they were obtained (i.e. lab name), the calibration program and the protocol for quality assurance OR state that no new dates are provided.*

☐ Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.

Ethics oversight	*Identify the organization(s) that approved or provided guidance on the study protocol, OR state that no ethical approval or guidance was required and explain why not.*

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	*For laboratory animals, report species, strain, sex and age OR state that the study did not involve laboratory animals.*
Wild animals	*Provide details on animals observed in or captured in the field; report species, sex and age where possible. Describe how animals were caught and transported and what happened to captive animals after the study (if killed, explain why and describe method; if released, say where and when) OR state that the study did not involve wild animals.*
Field-collected samples	*For laboratory work with field-collected samples, describe all relevant parameters such as housing, maintenance, temperature, photoperiod and end-of-experiment protocol OR state that the study did not involve samples collected from the field.*

| Ethics oversight | *Identify the organization(s) that approved or provided guidance on the study protocol, OR state that no ethical approval or guidance was required and explain why not.* |

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Human research participants

Policy information about studies involving human research participants

Population characteristics	*Describe the covariate-relevant population characteristics of the human research participants (e.g. age, gender, genotypic information, past and current diagnosis and treatment categories). If you filled out the behavioural & social sciences study design questions and have nothing to add here, write "See above."*
Recruitment	*Describe how participants were recruited. Outline any potential self-selection bias or other biases that may be present and how these are likely to impact results.*
Ethics oversight	*Identify the organization(s) that approved the study protocol.*

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration	*Provide the trial registration number from ClinicalTrials.gov or an equivalent agency.*
Study protocol	*Note where the full trial protocol can be accessed OR if not available, explain why.*
Data collection	*Describe the settings and locales of data collection, noting the time periods of recruitment and data collection.*
Outcomes	*Describe how you pre-defined primary and secondary outcome measures and how you assessed these measures.*

Dual use research of concern

Policy information about dual use research of concern

Hazards

Could the accidental, deliberate or reckless misuse of agents or technologies generated in the work, or the application of information presented in the manuscript, pose a threat to:

No	Yes	
☐	☒	Public health
☒	☐	National security
☒	☐	Crops and/or livestock
☒	☐	Ecosystems
☒	☐	Any other significant area

| Hazards | Wild-type and recombinant SARS-CoV-2 strains were used in this study |

For examples of agents subject to oversight, see the United States Government Policy for Institutional Oversight of Life Sciences Dual Use Research of Concern.

Experiments of concern

Does the work involve any of these experiments of concern:

No	Yes	
☐	☐	Demonstrate how to render a vaccine ineffective
☐	☐	Confer resistance to therapeutically useful antibiotics or antiviral agents
☐	☐	Enhance the virulence of a pathogen or render a nonpathogen virulent
☐	☐	Increase transmissibility of a pathogen
☐	☐	Alter the host range of a pathogen
☐	☐	Enable evasion of diagnostic/detection modalities
☐	☐	Enable the weaponization of a biological agent or toxin
☐	☐	Any other potentially harmful combination of experiments and agents

ChIP-seq

Data deposition

☐ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☐ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links *May remain private before publication.*	*For "Initial submission" or "Revised version" documents, provide reviewer access links. For your "Final submission" document, provide a link to the deposited data.*
Files in database submission	*Provide a list of all files available in the database submission.*
Genome browser session (e.g. UCSC)	*Provide a link to an anonymized genome browser session for "Initial submission" and "Revised version" documents only, to enable peer review. Write "no longer applicable" for "Final submission" documents.*

Methodology

Replicates	*Describe the experimental replicates, specifying number, type and replicate agreement.*
Sequencing depth	*Describe the sequencing depth for each experiment, providing the total number of reads, uniquely mapped reads, length of reads and whether they were paired- or single-end.*
Antibodies	*Describe the antibodies used for the ChIP-seq experiments; as applicable, provide supplier name, catalog number, clone name, and lot number.*
Peak calling parameters	*Specify the command line program and parameters used for read mapping and peak calling, including the ChIP, control and index files used.*
Data quality	*Describe the methods used to ensure data quality in full detail, including how many peaks are at FDR 5% and above 5-fold enrichment.*
Software	*Describe the software used to collect and analyze the ChIP-seq data. For custom code that has been deposited into a community repository, provide accession details.*

Flow Cytometry

Plots

Confirm that:

☐ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☐ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☐ All plots are contour plots with outliers or pseudocolor plots.

☐ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	*Describe the sample preparation, detailing the biological source of the cells and any tissue processing steps used.*
Instrument	*Identify the instrument used for data collection, specifying make and model number.*
Software	*Describe the software used to collect and analyze the flow cytometry data. For custom code that has been deposited into a community repository, provide accession details.*
Cell population abundance	*Describe the abundance of the relevant cell populations within post-sort fractions, providing details on the purity of the samples and how it was determined.*
Gating strategy	*Describe the gating strategy used for all relevant experiments, specifying the preliminary FSC/SSC gates of the starting cell population, indicating where boundaries between "positive" and "negative" staining cell populations are defined.*

☐ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

Magnetic resonance imaging

Experimental design

Design type	*Indicate task or resting state; event-related or block design.*

Design specifications	*Specify the number of blocks, trials or experimental units per session and/or subject, and specify the length of each trial or block (if trials are blocked) and interval between trials.*
Behavioral performance measures	*State number and/or type of variables recorded (e.g. correct button press, response time) and what statistics were used to establish that the subjects were performing the task as expected (e.g. mean, range, and/or standard deviation across subjects).*

Acquisition

Imaging type(s)	*Specify: functional, structural, diffusion, perfusion.*
Field strength	*Specify in Tesla*
Sequence & imaging parameters	*Specify the pulse sequence type (gradient echo, spin echo, etc.), imaging type (EPI, spiral, etc.), field of view, matrix size, slice thickness, orientation and TE/TR/flip angle.*
Area of acquisition	*State whether a whole brain scan was used OR define the area of acquisition, describing how the region was determined.*

Diffusion MRI ☐ Used ☐ Not used

Preprocessing

Preprocessing software	*Provide detail on software version and revision number and on specific parameters (model/functions, brain extraction, segmentation, smoothing kernel size, etc.).*
Normalization	*If data were normalized/standardized, describe the approach(es): specify linear or non-linear and define image types used for transformation OR indicate that data were not normalized and explain rationale for lack of normalization.*
Normalization template	*Describe the template used for normalization/transformation, specifying subject space or group standardized space (e.g. original Talairach, MNI305, ICBM152) OR indicate that the data were not normalized.*
Noise and artifact removal	*Describe your procedure(s) for artifact and structured noise removal, specifying motion parameters, tissue signals and physiological signals (heart rate, respiration).*
Volume censoring	*Define your software and/or method and criteria for volume censoring, and state the extent of such censoring.*

Statistical modeling & inference

Model type and settings	*Specify type (mass univariate, multivariate, RSA, predictive, etc.) and describe essential details of the model at the first and second levels (e.g. fixed, random or mixed effects; drift or auto-correlation).*
Effect(s) tested	*Define precise effect in terms of the task or stimulus conditions instead of psychological concepts and indicate whether ANOVA or factorial designs were used.*

Specify type of analysis: ☐ Whole brain ☐ ROI-based ☐ Both

Statistic type for inference (See Eklund et al. 2016)	*Specify voxel-wise or cluster-wise and report all relevant parameters for cluster-wise methods.*
Correction	*Describe the type of correction and how it is obtained for multiple comparisons (e.g. FWE, FDR, permutation or Monte Carlo).*

Models & analysis

n/a	Involved in the study
☐	☐ Functional and/or effective connectivity
☐	☐ Graph analysis
☐	☐ Multivariate modeling or predictive analysis

Functional and/or effective connectivity	*Report the measures of dependence used and the model details (e.g. Pearson correlation, partial correlation, mutual information).*
Graph analysis	*Report the dependent variable and connectivity measure, specifying weighted graph or binarized graph, subject- or group-level, and the global and/or node summaries used (e.g. clustering coefficient, efficiency, etc.).*
Multivariate modeling and predictive analysis	*Specify independent variables, features extraction and dimension reduction, model, training and evaluation metrics.*