PLOS PATHOGENS

RESEARCH ARTICLE

The proximal proteome of 17 SARS-CoV-2 proteins links to disrupted antiviral signaling and host translation

Jordan M. Meyers[1]◉, Muthukumar Ramanathan[1]◉, Ronald L. Shanderson[1,2]◉, Aimee Beck[3], Laura Donohue[1,4], Ian Ferguson[1,2], Margaret G. Guo[1,5], Deepti S. Rao[1], Weili Miao[1], David Reynolds[1], Xue Yang[1,2], Yang Zhao[1], Yen-Yu Yang[6], Catherine Blish[3], Yinsheng Wang[6], Paul A. Khavari[1,2]*

1 Program in Epithelial Biology, Stanford University, Stanford, California, United States of America,
2 Program in Cancer Biology, Stanford University, Stanford, California, United States of America,
3 Department of Medicine, Division of Infectious Diseases and Geographic Medicine, Stanford University School of Medicine, Stanford, California, United States of America, 4 Program in Genetics, Stanford University, Stanford, California, United States of America, 5 Program in Biomedical Informatics, Stanford University, Stanford, California, United States of America, 6 Department of Chemistry, University of California, Riverside, California, United States of America

◉ These authors contributed equally to this work.
* khavari@stanford.edu

Abstract

Viral proteins localize within subcellular compartments to subvert host machinery and promote pathogenesis. To study SARS-CoV-2 biology, we generated an atlas of 2422 human proteins vicinal to 17 SARS-CoV-2 viral proteins using proximity proteomics. This identified viral proteins at specific intracellular locations, such as association of accessary proteins with intracellular membranes, and projected SARS-CoV-2 impacts on innate immune signaling, ER-Golgi transport, and protein translation. It identified viral protein adjacency to specific host proteins whose regulatory variants are linked to COVID-19 severity, including the TRIM4 interferon signaling regulator which was found proximal to the SARS-CoV-2 M protein. Viral NSP1 protein adjacency to the EIF3 complex was associated with inhibited host protein translation whereas ORF6 localization with MAVS was associated with inhibited RIG-I 2CARD-mediated *IFNB1* promoter activation. Quantitative proteomics identified candidate host targets for the NSP5 protease, with specific functional cleavage sequences in host proteins CWC22 and FANCD2. This data resource identifies host factors proximal to viral proteins in living human cells and nominates pathogenic mechanisms employed by SARS-CoV-2.

Author summary

SARS-CoV-2 is the latest pathogenic coronavirus to emerge as a public health threat. We create a database of proximal host proteins to 17 SARS-CoV-2 viral proteins. We validate that NSP1 is proximal to the EIF3 translation initiation complex and is a potent inhibitor of translation. We also identify ORF6 antagonism of RNA-mediate innate immune

signaling. We produce a database of potential host targets of the viral protease NSP5, and create a fluorescence-based assay to screen cleavage of peptide sequences. We believe that this data will be useful for identifying roles for many of the uncharacterized SARS-CoV-2 proteins and provide insights into the pathogenicity of new or emerging coronaviruses.

Introduction

Coronaviruses comprise a diverse family of large positive-sense single stranded (+ss)RNA enveloped viruses that cause respiratory and gastrointestinal disease. In addition to common seasonal coronaviruses, a number of strains can cause severe disease, as seen in the Severe Acute Respiratory Syndrome Coronavirus (SARS-CoV) virus outbreak in 2003 [1], the Middle East Respiratory Syndrome Coronavirus (MERS-CoV) virus outbreak in 2012 [2] and the 2019 outbreak of SARS-CoV-2 [3]. This viral family has large (26 to 32kb) genomes that encode tens of viral proteins. All coronaviruses have a similar organization consisting of a large open reading frame (ORF1) encoding two overlapping polyproteins, pp1A and pp1AB. These polyproteins are cleaved by one of two viral proteases, nonstructural protein 3 (NSP3) and NSP5, with the resulting protein products sequentially numbered NSP1-NSPX. ORF1AB is invariably followed by genes encoding the structural proteins, including the Spike protein (S), Envelope protein (E), Membrane protein (M), and Nucleocapsid protein (N). SARS-CoV-2 encodes NSP1-16 as well as the accessory proteins ORF3a, ORF3b, ORF6, ORF7a, ORF7b, ORF8, ORF9b, ORF9c, ORF10, and ORF14, though it is not known if each open reading frame encodes for a functional protein product. The function of many SARS-CoV-2 accessory proteins is either unknown or highly variable across differing coronaviruses, underscoring the need to begin mapping their putative localizations and functions.

Proximity proteomics (BioID) uses enzymes, such as the modified bacterial biotin ligase, BirA, to biotinylate nearby proteins on lysine residue-containing proteins within a radius of 10-20nm [4]. When fused to a protein of interest it labels not only proteins that directly bind the fused protein but also those adjacent to it, enabling rapid isolation of biotinylated proteins whose identity can provide clues about the localization and function of the protein studied. When coupled to mass spectrometry it provides an alternative to traditional tandem affinity purification and mass spectrometry (TAP-MS) [5]. Whereas, TAP-MS can isolate protein complexes that stably bind the protein of interest in a manner robust enough to survive protein extraction, BioID-MS labels both transient and stable interactors in living cells, particularly those stabilized by cellular membranes that can be destroyed in traditional TAP-MS experiments. In this way, BioID may localize the cellular "neighborhoods" of a given fusion protein. We recently generated a biotin ligase derived from *Bacillus subtilis*, which has 50 times greater activity than the original *E. coli* BirA [4,6], allowing decreased labeling times and increased signal-to-noise ratios. Applying proximity proteomics to SARS-CoV-2 viral proteins in human cells may facilitate insight into their localization and putative functions.

The actions of specific SARS-CoV-2-encoded proteins are only partially understood at present. The replication transcription complex, which includes the RNA-dependent RNA polymerase and other factors, and the N protein, which are necessary for protecting the newly synthesized genomes and assembling the viral particles, comprise the core viral replication machinery. Other viral gene products, generally termed accessory factors, are believed to be dedicated to manipulating the host environment to foster viral replication [7]. One of the main functions of accessory factors is to block host antiviral response [8] though many viral proteins are multifunctional and structural proteins can also affect host signaling pathways [9].

Other coronaviruses have also been shown to block host translation [10,11], inhibit interferon signaling [12, 13], antagonize viral RNA sensing [14,15], and degrade host mRNAs [16]. The degree of homology between SARS-CoV-2 and other coronaviruses, suggests the existence of both shared and divergent host protein interactions between its viral proteins and those of the other members of the coronavirus family.

Here we used proximity proteomics to identify the human proteins vicinal to 17 major SARS-CoV-2 proteins and, from that data and validation studies, to predict their likely location and function. We examined the intersection of the resulting atlas of human factors adjacent to SARS-CoV-2 viral proteins with risk loci associated with severe COVID-19 by genome wide association studies (GWAS). This nominated specific, viral protein-adjacent host candidates whose natural variation in expression may contribute to differences in COVID-19 susceptibility in the population. We also demonstrated that multiple SARS-CoV-2 products can affect host translation and host innate immune signaling and define a list of potential host targets and pathways for the NPS5 protease. Taken together, these resource data plot the location of the 17 major SARS-CoV-2 within the cell, define an atlas of human host proteins adjacent to them, and offer insight into potential pathogenic mechanisms engaged by SARS-CoV-2.

Results

Host proteins proximal to viral proteins and their subcellular localization

To identify the human host proteins vicinal to the 17 major SARS-CoV-2 encoded viral proteins, HA epitope tagged fusions of BASU-BirA [6] were generated with each of these 17 viral ORFs (**Fig 1A**). BASU was introduced at the N and C terminus to minimize disruption as previously described [17]. Samples were prepared from plasmid-transfected 293T cells after 2 hours of biotin labeling and the biotinylated proteins were then isolated using streptavidin. Samples were divided for LC-MS/MS and immunoblotting (**S1 Fig**). MS data search was performed and protein lists were analyzed and scored using the Significance Analysis of Interactome (SAINT) method [18]. Using a cutoff of a SAINT score of 0.9 generated a list of 2421 host proteins compromising 1119 different host proteins. (**Figs 1B and S2, and S1 Table**) across the 17 viral proteins studied, 513 of which were unique to a specific viral protein. These data comprise a compendium of candidate human proteins adjacent to SARS-CoV-2-encoded proteins.

The identity of these 1119 human proteins provided clues to SARS-CoV-2 biology. Gene ontology (GO) term analysis (**Fig 1B, 1C, and 1D and S2 Table**) identified processes associated with SARS-CoV-2 viral protein impacts. This included translation initiation, RNA binding, the 26S proteasome, signaling, and SNARE-associated intracellular transport. It also identified adjacencies to major histocompatibility (MHC) proteins and components of the nuclear pore complex (NPC). A number of these processes, such as protein translation, are known processes affected by coronaviruses, while others, such as RNA-binding, are less well characterized.

To begin to map putative localizations for the 17 studied SARS-CoV-2 proteins within the cell, cellular component GO-term enrichment analysis was performed (**Fig 2A**), which pointed to possible intracellular localizations for each viral protein based on curated knowledge of the host proteins identified adjacent to each viral protein. To validate and extend this, protein fractions were prepared from cells expressing each SARS-CoV-2 protein studied. These included four overlapping fractions: a) cytoplasm b) cytoplasm/membrane c) nucleus/membrane, and d) nucleus (**Fig 2B**). Integrating GO-term analysis with immunoblotting of these fractions enabled predictions of the likely intracellular localization of each viral protein (**Fig 2C and 2D**). Many SARS-CoV-2 accessory proteins concentrate in the ER or in ER-proximal
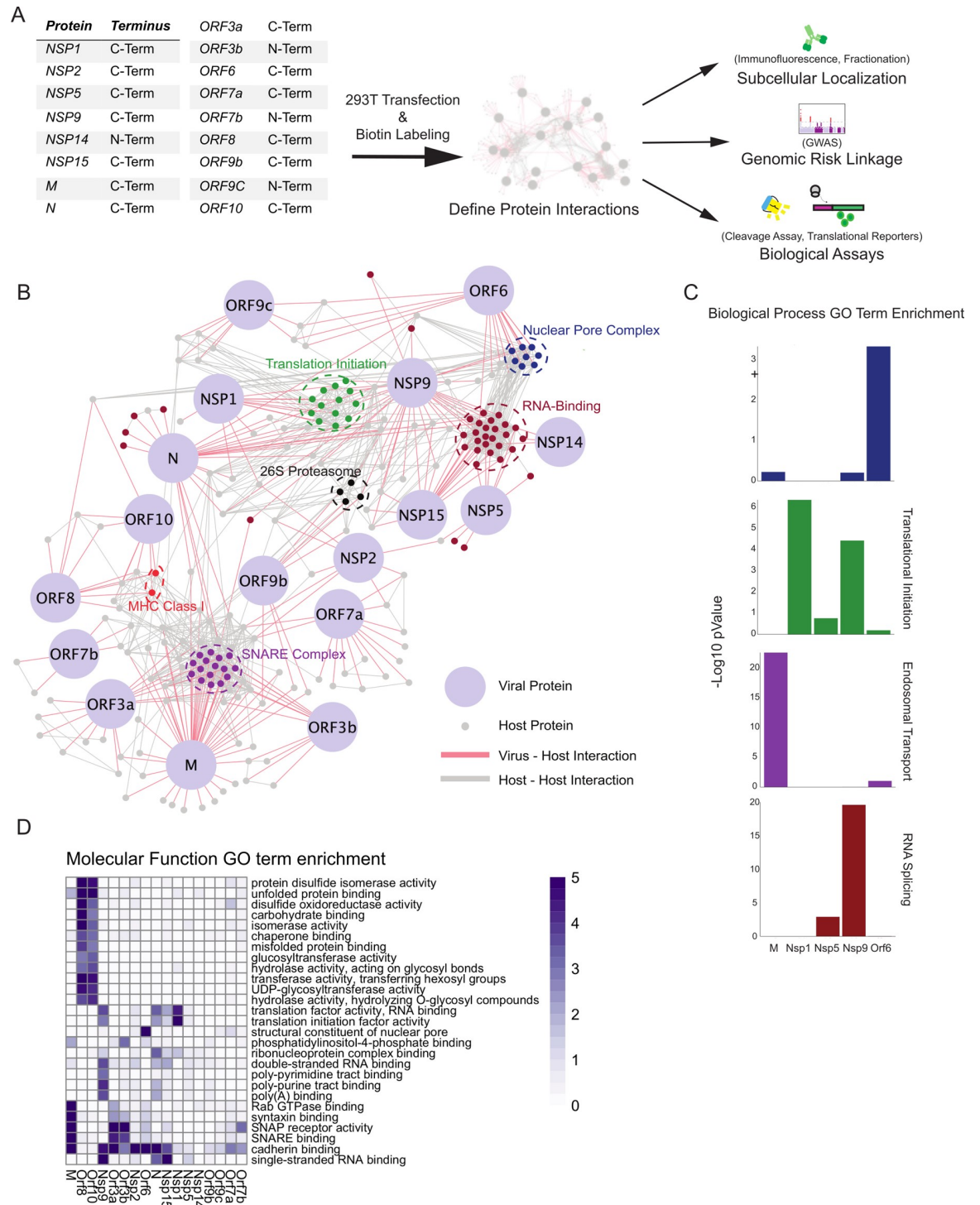
Fig 1. Proximal Interactome of 17 SARS-CoV-2 proteins. A) Schematic of BioID workflow. B) Curated network of SARS-CoV2 virus-host protein associations (SAINT ≥ 0.9) obtained from BASU BioID. Coronavirus proteins are labeled in light blue and virus-host interactions are connected by red edges, while host-host protein interactions obtained from high confidence STRING interactions are labeled in grey. Highlighted node clusters of similar function, including 26S proteasome components (black), MHC Class I (red), nuclear pore (dark blue), RNA-binding (maroon), SNARE complex (purple), translation initiation complex (green) proteins were selected based on GO term analysis. C) Selected biological process GO term enrichment; enrichment scores are given as -Log10 p-values. Selected GO terms are nuclear pore organization, translational initiation, endosomal transport, and RNA splicing. D) Heatmap of molecular function GO term enrichment of SARS-CoV-2 proteins. All presented GO terms have a -Log10 p-value >3 for the Nucleoprotein, the listed non-structural proteins, or the listed open reading frames or a -Log10 p-Value >5 for the M membrane protein.

A Cell Component GO Terms

B

C
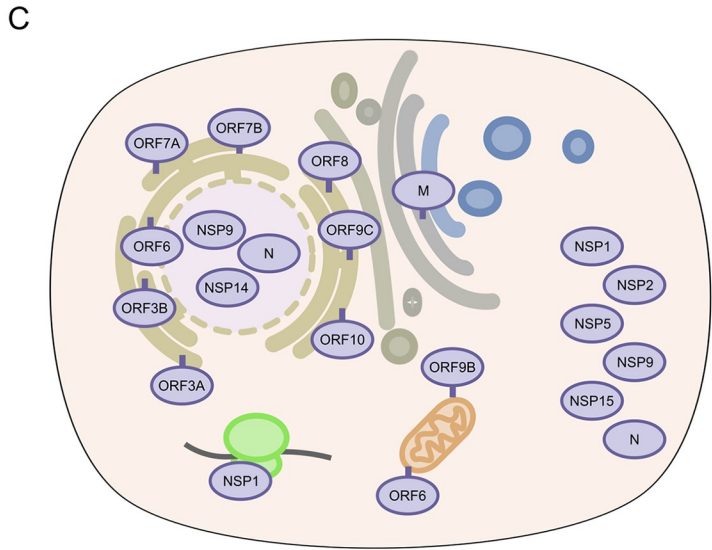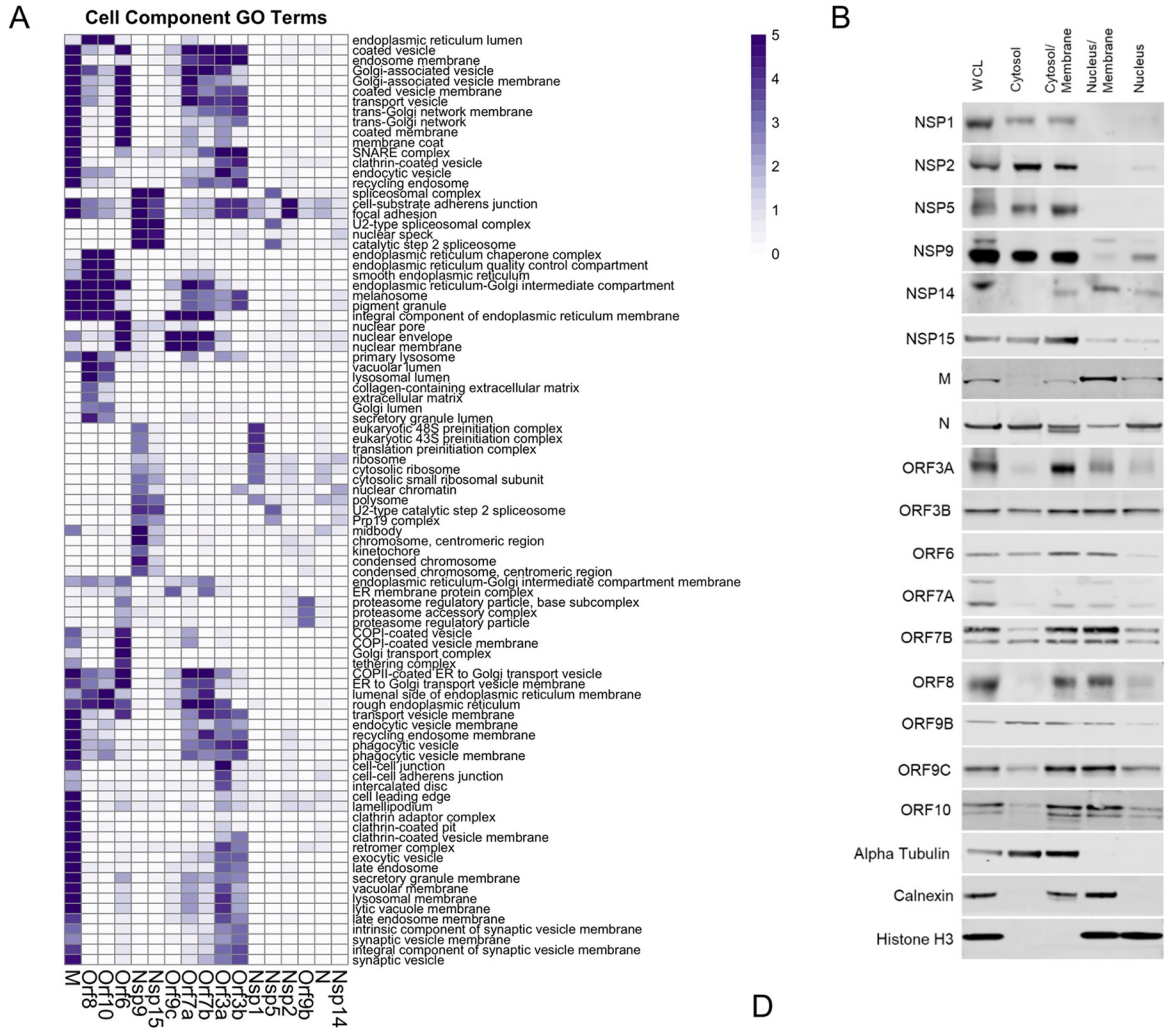
D

Viral Protein	Predicted Location
NSP1	Cytoplasmic
NSP2	Cytoplasmic
NSP5	Cytoplasmic
NSP9	Cytoplasmic, Nucleus
NSP14	Nucleus
NSP15	Diffuse
M	ER, Golgi, Endosomal
N	Diffuse
ORF3A	ER, Golgi, Endosomal
ORF3B	ER, Golgi, Endosomal
ORF6	ER, Nuclear Membranes
ORF7A	ER, Golgi
ORF7B	ER, Golgi
ORF8	ER, Golgi
ORF9B	Diffuse
ORF9C	Nuclear Membranes
ORF10	ER, Golgi

Fig 2. Localization of SARS-CoV-2 proteins. A) Heatmap of cell component GO term enrichment of SARS-CoV-2 proteins. B) Western blots of SARS-CoV-2 viral protein-expressing HEK293T cell fractions; whole cell lysate (WCL), cytosol, cytosol/membrane, nucleus/membrane and nucleus fractions. Alpha-tubulin, calnexin, and histone H3 were used as fractionation controls for cytosol, membrane, and nucleus respectively. Schematic C) and table D) depicting the predicted location of all SARS-CoV-2 proteins surveyed in this study based on both the BioID and fractionation analysis.

membranes (M, ORF3a, ORF3b, ORF6, ORF7a, ORF7b, ORF8, and ORF10). A number, however, appear to be predominantly cytoplasmic (NSP1, NSP2, NSP5, NSP15, ORF9b) and, interestingly, several appear to localize in part to the nucleus or nuclear membrane (NSP9, NSP14, ORF6, ORF9c). The localization predicted from these data is consistent with observations from other recent work [17,19]. Of the membrane localized proteins, subtle differences in location could be inferred. In the case of M protein, association with membranes in the endocytic pathway as well as lysosomal membranes was predicted. ORF8 and ORF10 clustered similarly with enrichment for ER interactions in the lumen.

Additionally, we sought to confirm the location of select viral proteins during infection. This would allow us to both assess if localization matched in conditions where all viral proteins are expressed and at normal levels, and also to rule out that localization is not different during transient overexpression of our BASU-tagged constructs. Specifically, A549 cells expressing the SARS-CoV-2 receptor, ACE2, were infected with SARS-CoV-2 at an MOI of 0.1, fixed with 4% paraformaldehyde, and stained with antibodies [20] against SARS-CoV-2 proteins (NSP1, NSP14, ORF6) along with host proteins identified from the BIOID (EIF3A, Rae1, Nup98, MCM3AP) and Hoescht for nuclear staining (**Fig 3**). NSP1 displayed punctate cytoplasmic (**Fig 3A**), while NSP14 stained both perinuclear and nuclear (**Fig 3B**), and ORF6 stained mostly perinuclear (**Fig 3C**). The staining patterns of infected cells largely mirrored that which was observed from the BIOID and fractionation studies. These data indicate that specific SARS-CoV-2 may display increased localization to a variety of intracellular sites, including the cytoplasm, nucleus and distinct endomembranes.

As there exists an orthogonal proteomics dataset [17], our results were compared with the previously published dataset. Rather than using proximity labeling to identify interactors, this study used TAP-MS to identify co-purified interactors of SARS-CoV-2 viral proteins. These two approaches have been found to be complementary [21,22]; where the first approach enriches for stable complexes, while the latter enriches for neighbor proteins in the endogenous environments of the bait. To address this, the identified viral proteomes of the two studies were compared. To do so we first reanalyzed the data from the TAP-MS study [17] to call all hits with a SAINT score ≥ 0.9 as was to call high-confidence interactors equally. This generated a list of 695 host proteins. Comparison between the two studies yielded 960 proteins unique to the current study, 536 unique to the TAP-MS study, and 159 proteins shared between the two (**S3 Fig**). GO-term analysis for molecular function, biological process, and cellular component were performed on the hits from the TAP-MS data (**S4, S5, and S6** Figs). Proximity labeling methods are particularly suited for identifying the location of bait proteins [23,24] therefore, cell component analysis was compared between the two lists (**S7 Fig**). Given the high amount of unique hits, it was not surprising to see major differences between cell component enrichment between viral proteins. NSP1 uniquely enriched for hits associated with preinitiation complex and ribosomal components by BIOID, but not by TAP-MS. Contrastingly, NSP9 enriched for similar factors only by TAP-MS. Despite the differences, there were important similarities between the two approaches. For example, ORF6 was found to enrich for nuclear pore components using either method. Taken together for SARS-CoV-2 biology, the two approaches appear to give complementary insights into potential targets and pathways and both are valid for discovery.
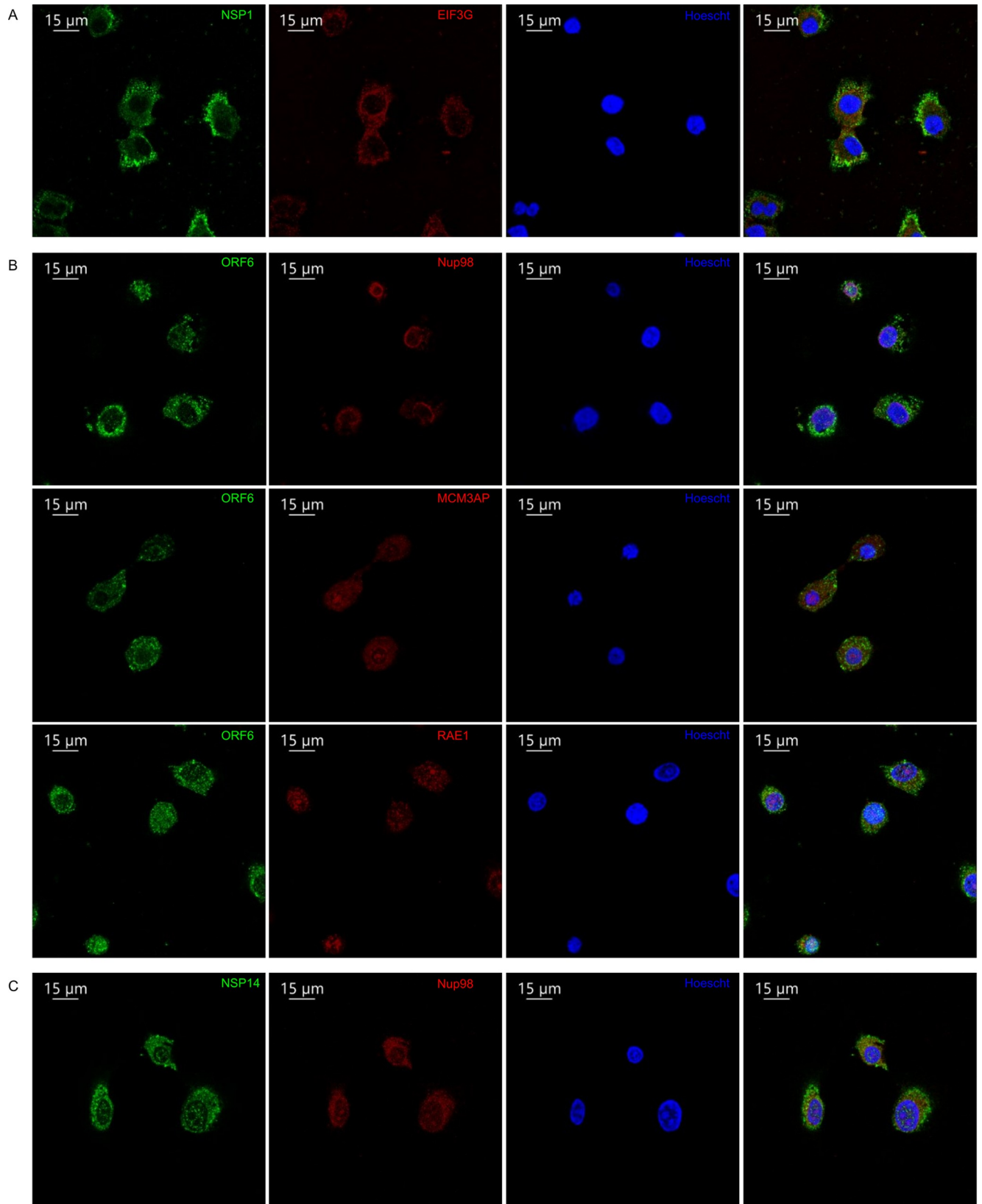
Fig 3. Immunofluorescent staining of SARS-CoV-2 proteins in infected cells. A549 cells expressing ACE2 were infected for 48 hours with SARS-CoV-2 WA1 strain, fixed, and stained with the indicated viral (green) and host proteins (red) as well as nuclear staining by Hoescht (blue). A) NSP1, B) ORF6, C) NSP14.

https://doi.org/10.1371/journal.ppat.1009412.g003

Viral proximal interactors include drug targetable host genes

There is a lack of SARS-CoV-2 specific antiviral therapies or against coronaviruses generally. Many current and experimental therapeutics were developed for activity against other viruses and are being tested for cross efficacy against SARS-CoV-2. Others are therapies known to have broad antiviral effects. There is significant interest in developing drugs that directly target SARS-CoV-2 viral proteins, but research and development may take years before use in patients. Another approach is using drugs against host genes critical to virus infection and replication. For example, drugs targeting ACE-2, the main receptor for SARS-CoV-2, or ACE-2 expression and function have been pursued. To expand the list of possible drugs beyond entry inhibitors, we compared the viral proximal proteome generated in this study against the "druggable" genome, which include databases of the gene targets of available drugs. This generated a list of 47 host genes (**S8 Fig** and **S3 Table**) and highlights, as previously reported [17], a group of cellular kinases associated with N protein. The viral nucleocapsid has been shown to be phosphorylated and phosphorylation is suggested to be important for its function [25,26]. This highlights cellular kinase inhibitors as drugs with possible activity against SARS-CoV-2.

GWAS-linked host proteins in the viral proximal proteome

The genetic basis for the wide spectrum of COVID-19 severity in different individuals within the human population is not fully understood. A number of recent genome wide association studies (GWAS) studies have endeavored to map genetic risk loci associated with SARS-CoV-2 infection and COVID-19 clinical severity [27,28]. These studies leverage large numbers of patients to identify SNPs that are correlated with outcomes such as infection and severity of disease, including hospitalization and mortality. Such linkage studies have identified a number of non-coding variants that may perform a regulatory function, for example, by altering expression of effect genes (eGenes) important in host susceptibility to SARS-CoV-2.

To determine if any putative COVID-19 risk-linked regulatory variants might control the expression of host proteins proximal to SARS-CoV-2 viral proteins, the following analysis was performed. Using publicly available data from GWAS studies [27,28], all single nucleotide polymorphisms (SNPs) associated with increased risk of COVID disease that reside in non-coding DNA were identified. These were filtered for variants localized to open chromatin, characteristic of regulatory DNA, in cell types relevant to COVID-19 pathogenesis, including immune and pulmonary cells. The resulting disease risk-linked variants were further distilled to those identified as expression quantitative trail loci (eQTLs) for specific putative eGene targets (**Fig 4A**). These eGenes, which represent a set of genes whose expression may be controlled by natural variants in the human population linked to COVID-19 risk, were then intersected with the atlas of host factors identified as adjacent to SARS-CoV-2 viral proteins by proximity proteomics. Publicly available protein interaction data was then integrated to project the connectedness of resulting gene set (**Fig 4B**). The resulting network was notable for host proteins implicated in cytokine signaling, cell cycle control, transcription, and translation, suggesting that genetic susceptibility to COVID-19 may link to variations in the expression of proteins that mediate these processes.

Among proteins identified by this analysis was TRIM4, a RING E3 ligase, that activates type I interferon signaling through activation of the cytosolic RNA sensor RIG-I. TRIM4 was significantly associated with SARS-CoV-2 M protein in proximity proteomics data (**S1 Table**) and,
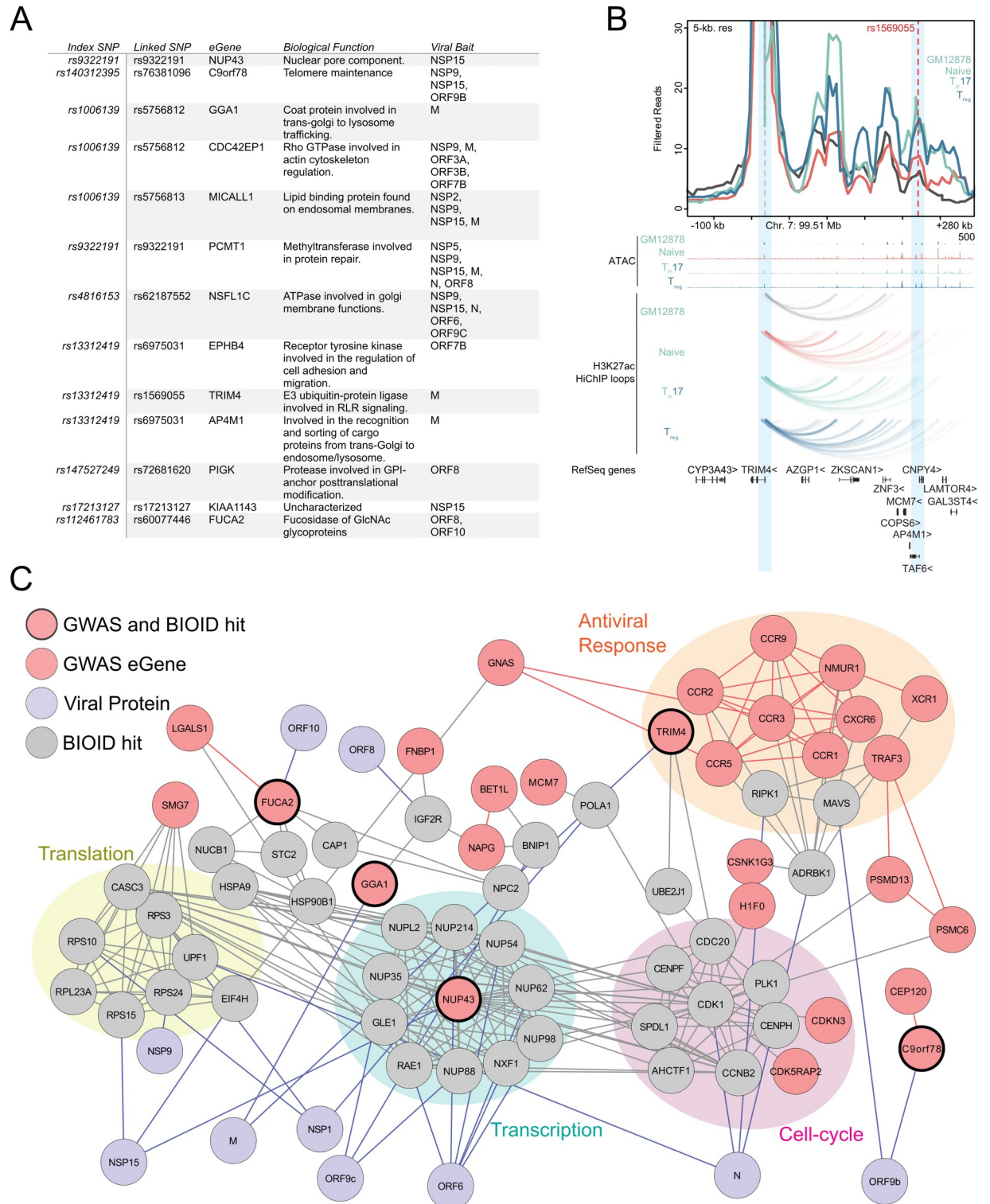
A

Index SNP	Linked SNP	eGene	Biological Function	Viral Bait
rs9322191	rs9322191	NUP43	Nuclear pore component.	NSP15
rs140312395	rs76381096	C9orf78	Telomere maintenance	NSP9, NSP15, ORF9B
rs1006139	rs5756812	GGA1	Coat protein involved in trans-golgi to lysosome trafficking.	M
rs1006139	rs5756812	CDC42EP1	Rho GTPase involved in actin cytoskeleton regulation.	NSP9, M, ORF3A, ORF3B, ORF7B
rs1006139	rs5756813	MICALL1	Lipid binding protein found on endosomal membranes.	NSP2, NSP9, NSP15, M
rs9322191	rs9322191	PCMT1	Methyltransferase involved in protein repair.	NSP5, NSP9, NSP15, M, N, ORF8
rs4816153	rs62187552	NSFL1C	ATPase involved in golgi membrane functions.	NSP9, NSP15, N, ORF6, ORF9C
rs13312419	rs6975031	EPHB4	Receptor tyrosine kinase involved in the regulation of cell adhesion and migration.	ORF7B
rs13312419	rs1569055	TRIM4	E3 ubiquitin-protein ligase involved in RLR signaling.	M
rs13312419	rs6975031	AP4M1	Involved in the recognition and sorting of cargo proteins from trans-Golgi to endosome/lysosome.	M
rs147527249	rs72681620	PIGK	Protease involved in GPI-anchor posttranslational modification.	ORF8
rs17213127	rs17213127	KIAA1143	Uncharacterized	NSP15
rs112461783	rs60077446	FUCA2	Fucosidase of GlcNAc glycoproteins	ORF8, ORF10

B

C

Fig 4. COVID disease risk eGenes proximal to viral proteins. A) Table of GWAS risk SNPs which also scored as BioID hit. B) Map of connectedness of eGenes (Mauve) with BioID interactors (Gray) and the corresponding viral proteins (Purple). eGenes also identified by BioID are outlined in black. GWAS-identified eGenes-associated with antiviral response, cell-cycle, transcription, and translation are also highlighted. C) Virtual 4-C plot showing chromatin contact between TRIM4 promoter and linked COVID disease risk SNP (rs1569055). Genome tracks showing ATAC peaks and contact loops of GM12878, Naïve T, Th-17, and T-reg cells.

using eQTLgen [29], a regulatory SNP (rs1569055) approximately 230kb downstream of the TRIM4 promoter was recently associated with increased COVID severity in patients [28]. High-C chromatin immunoprecipitation (HiChIP) data from the immortalized B-cell line GM12878 as well as primary T-cell populations demonstrated chromatin looping from the SNP to the TRIM4 promoter (**Fig 4C**). Looping of this SNP increased contact strength in naïve, T-regs and Th-17 T-cells. TRIM4 is one of a group of ubiquitin ligases [30–33] that can activate RIG-I during RNA-sensing and subsequent antiviral signaling. Altered expression coupled with disruption by potential association with the SARS-CoV-2 M protein supports a model with the following features; a) individuals with this regulatory variant may express less TRIM4 b) physical association with SARS-CoV-2 M protein further reduces functional TRIM4 c) a relative reduction in biologically active TRIM4 leads to reduced innate immune signaling d) this reduction leads to increased susceptibility to SARS-CoV-2 pathogenesis. Integrating proximity proteomics data with genetic risk eQTL variants may help identify such candidate susceptibility mechanisms for natural variations in disease outcomes within the population.

In addition to GWAS studies, several groups have undertaken genome wide CRISPR screens to identify host factors required for viral replication [34–36]. The intersection of significant hits from these CRISPR screens and the viral proteome from this study yielded few hits. Given the role of accessory factors as being non-essential for viral replication, the lack of overlap is not surprising. Nevertheless, NXF1 and MCM3AP, known mediators of mRNA export [37], were found as hits in both CRISPR studies and as proximal interactors of ORF6. NXF1 and MCM3AP are part of an RNA export complex [38] that docks at nuclear pores.

Predicted viral antagonism of host protein translation and antiviral response

NSP1 is a part of the viral polyprotein ORF1 and during normal viral replication is cleaved and liberated by the viral protease NSP3. Earlier work has identified NSP1 of SARS-CoV-1 as a potent inhibitor of translation in a mechanism that involves interactions with the host ribosomes [10,13]. Recently other groups have shown that NSP1 of SARS-CoV-2 similarly blocks translation through interaction with the 40s ribosome[39, 40]. High confidence proteins proximal to NSP1 included EIF3A, EIF3B, EIF2G, and EIF4G2 (**Fig 5A**) of which the first 3 are components of the EIF3 translation initiation complex. Interestingly, members of the EIF3 complex were not identified as high-confidence interactors by traditional TAP-MS studies [17,19]. Having confirmed through immunofluorescent microscopy that NSP1 predominantly resides in the cytosol of infected cells but only weakly co-localized with EIF3A, a member of the translation initiation complex, we sought to test if SARS-CoV-2 NSP1 inhibits host translation, NSP1 was expressed in HEK293T cells followed 24 hours later by transfection of in-vitro transcribed capped and polyadenylated mRNA expressing luciferase. NSP1 reduced luciferase signal by (49.7%) as compared to GFP control (**Fig 5B**), demonstrating that NSP1 can inhibit host cap-dependent translation, consistent with data reported by others [39,40]. To determine if NSP1 could inhibit translation of host-derived 5' UTRs and host IRES elements, two host UTRs (IFIT1 and ISG15) were subcloned separately upstream of luciferase along with two host IRES sequences (XIAP1 and APAF1) and luciferase measured in cells transfected with or without NSP1 construct. NSP1 reduced luciferase signal of both 5' UTRs (IFIT1 = 55.2%, ISG15 = 53.1%) and IRES elements (XIAP1 = 55.0%, APAF1 = 40.0), indicating a block in translation of these elements (**Fig 5B**). Lastly, NSP1 effects were tested on the SARS-CoV-2 5' UTR and the Cricket Paralysis Virus (CRPV) IRES. CRPV IRES is a minimal viral-derived
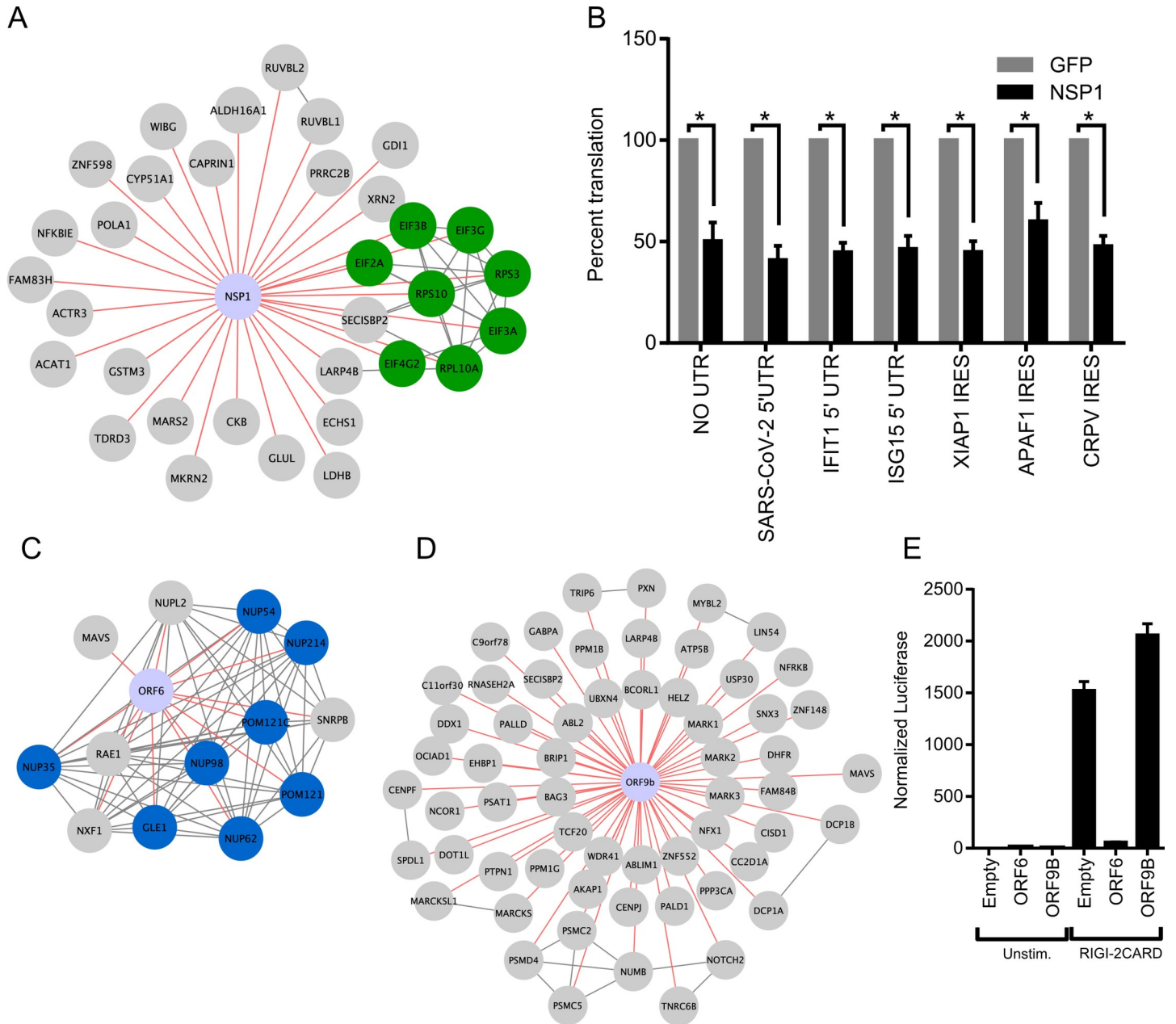
Fig 5. NSP1 and ORF6 disruption of host translation and innate immune signaling. A) Curated map of NSP1 proximal interactors. Highlighted are host proteins involved in translation initiation. B) Effect of NSP1 on translation of in vitro transcribed, capped polyadenylated transcripts containing 5' UTRs from SARS CoV-2, *IFIT1*, and *ISG15* as well as IRES elements from *XIAP1*, *APAF1*, and CRPV. Data shown is the average of three independent experiments and significance was calculated using Student's T Test where * indicates p value<0.005. Curated maps of ORF6 C) and ORF9b D) proximal proteins showing nuclear pore protein complex association with ORF6 and MAVS association with both ORF6 and ORF9b. E) Effect of ORF6 and ORF9b on *IFNB1* promoter activity after RIG-I 2-CARD induction. Normalized luciferase shown is the ratio of nano luciferase/firefly luciferase normalized to empty vector control. Data shown is the average of three independent experiments.

IRES that initiates translation completely independent of EIF3. Surprisingly, NSP1 blocked both viral elements (SARS-CoV-2 = 59.1%, CRPV = 52.2%) compared to GFP control. NSP1 therefore exhibits broad translation inhibition of mRNAs containing various regulatory elements, suggesting NSP1 action on the initiating ribosome, however, additional actions, such as mRNA cleavage[10, 16], may also be operative. This highlights the utility of using proximal

proteomic labeling to identify, although they be weak or transient interactions, functional partners of viral proteins that cannot be readily identified by TAP-MS.

Host innate immune detection and signaling pathways are heavily targeted by viral proteins, especially accessory proteins [41]. Mitochondrial Activation of Viral Signaling (MAVS) is a critical signaling adaptor for RIG-I like receptors (RLR) cytosolic sensing pathway [42–45]. It recruits activated RLR sensors RIG-I and MDA-5 at mitochondrial and mitochondrial-proximal membranes and leads to the activation of both IRF3 and NF-κB and expression of type-I interferons [46]. RIG-I and MDA-5 recognize various types of non-host or aberrant RNA species and are critical for host defense against RNA viruses [47]. MAVS was found as a high confidence protein proximal to two SARS-CoV-2 proteins: ORF6 and ORF9b (**Fig 5C and 5D**). ORF6 has been found to inhibit type-I interferon SARS-CoV-2 [48,49] and the closely related SARS-CoV-1 [12,50]. One study demonstrated that ORF6 inhibition of type-1 interferon expression was linked to ORF6 binding to nuclear import complex Nup98/Rae1 [49], both of which were also captured as not only proximal interactors, but also were also observed to be localized in similar proximity to ORF6, but not ORF9B. We tested the ability of our SARS-CoV-2 ORF6 and ORF9b constructs to inhibit RLR signaling by co-transfecting constructs expressing ORF6 or ORF9B along with a reporter expressing nanoluciferase under the control of the *IFNB1* promoter for interferon β1 along with a second reporter constitutively expressing firefly luciferase. To activate RLR signaling we transfected in a plasmid expressing a truncated version of RIG-I only containing the 2 CARD domains. This truncation is constitutively recruited to MAVS and initiates signaling in absence of any RNA stimulus and will test the viral proteins' ability to block any signaling downstream of sensing. ORF6 significantly inhibited RIG-I 2CARD activation of *IFNB1* promoter activity by 96 percent (**Fig 5E**) while ORF9b showed no effect on inhibiting *IFNB1* promoter activity. These data demonstrating ORF6 proximity to MAVS, along with ORF6 inhibition of *IFNB1* promoter induction, implicate ORF6 impairment of MAVS in the RLR innate immune signaling pathway.

NSP5 proteomics prediction of potential host cleavage targets

NSP5 is one of two critical proteases encoded by SARS-CoV-2 and is also known as SARS-CoV-2 3CLpro due to its similarity to picornavirus 3C proteases and a number of other +ssRNA viruses. These proteases all contain chymotrypsin-like folds and a triad of residues harboring the critical cysteine residue [51]. 3CLpro-like proteases are considered important therapeutically since they are essential for cleaving large polyprotein products produced by +ssRNA viruses and chemical protease inhibitors may act broadly across members of a given virus family [52,53]. In addition to their necessity in the virus life cycle, many viral proteases can target host proteins and specifically affect antiviral responses or other cellular processes [54–56]. Complementing previous efforts to infer targets of the NSP5 protease, we identified 34 host proteins in the NSP5 proximal proteome (**S2C Fig** and **S1 Table**). To nominate possible host targets of NSP5 whose levels are decreased upon protease expression, we performed SILAC mass spectrometry comparing wild type SARS-CoV-2 NSP5 to the catalytically-inactive NSP5^{C145A} mutant [17,57]. Residue 145 is the critical catalytic cysteine and mutation to alanine prevents protease activity [58]. Cell death was observed following transient expression of wild type NPS5, but not NSP5^{C145A}, so cells were collected 24 hours post-transfection to minimize those effects. Immunoblotting for NSP5 and C145A in whole cell lysates confirmed expression of the proteases in samples from both heavy and light media (**S9 Fig**). A number of host proteins showed significant depletion in cells expressing wild type NSP5, but not protease-inactive NSP5^{C145A} (**Fig 6A**). Combining both data generated identified an additional 26 candidates resulting in a pool of 60 potential host protein targets for NSP5 (**Fig 6B**).
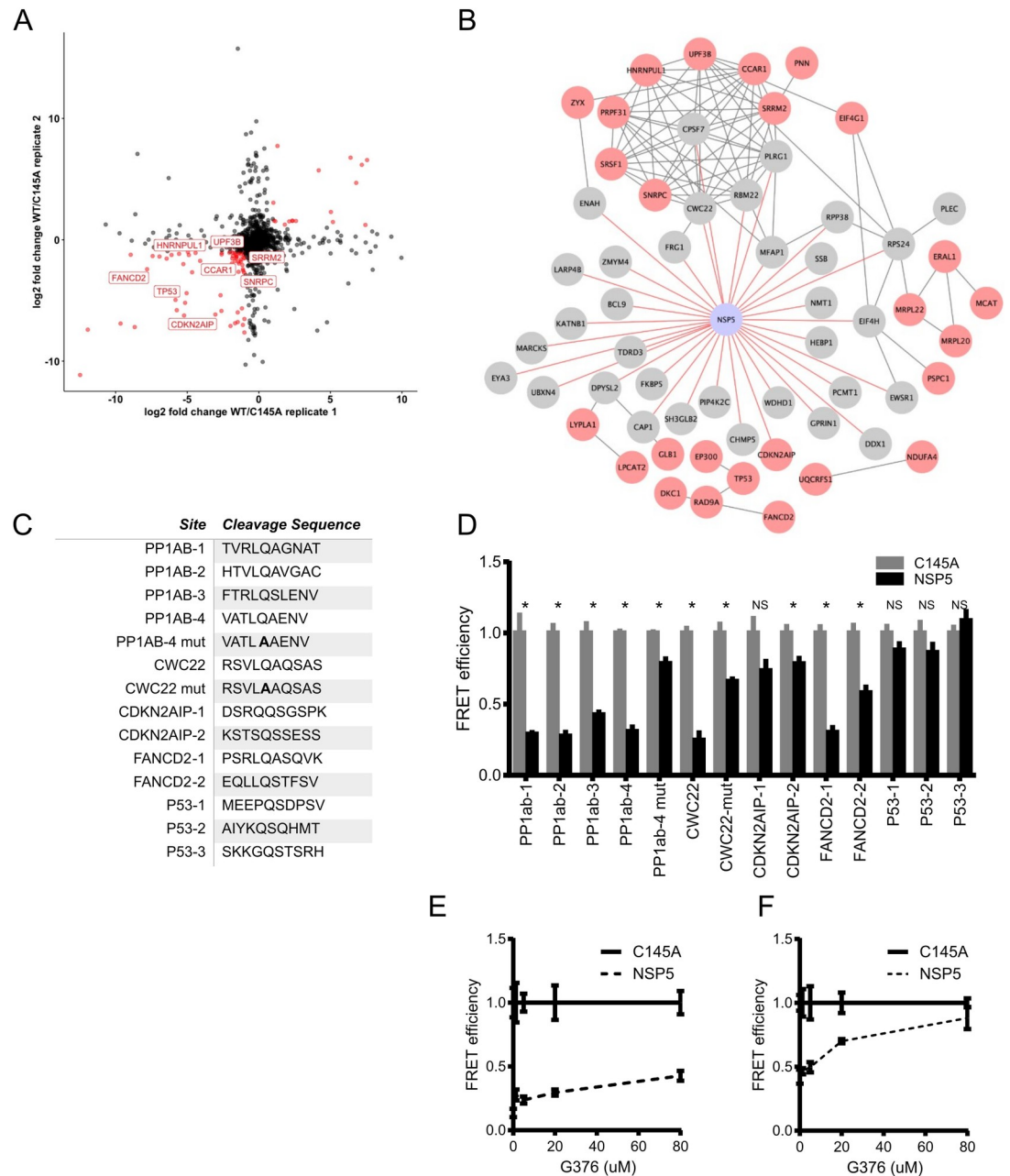
Fig 6. BioID and SILAC MS identify candidate targets for the viral protease NSP5. A) Comparison of two biological replicates of protein abundance in HEK293T cells expressing either NSP5 wild type (WT) or the catalytically inactive NSP5^C145A mutant by $\log_2$ fold change. B) Map of NSP5 proximal interactome (Gray) overlaid with host proteins decreased in abundance in SILAC (Red). CDNK2AIP was detected as both a BioID hit and decreased in abundance in SILAC. C) Peptide cleavage assay of four sequences from SARS-CoV-2 polyprotein PP1AB and the indicated host genes: CWC22, CDNK2AIP, FANCD2, P53. Normalized FRET signal is shown comparing HEK293T cells expressing either wild type NSP5 or NSP5^C145A. PP1ab and CWC22 mutant (mut) sequences contain nucleic acid changes that result in QS→AS mutation in the peptide sequence. Data shown is representative of three independent experiments and significance was calculated using Student's T Test. * Indicates p value <0.05, NS not significant. Dose-dependent effect of coronavirus protease inhibitor G376 on cleavage of peptide encoded by PP1ab-2 D) and FANCD2-2 E) sequences.

To begin to examine potential cleavage of these candidate proteins by NSP5, we searched their peptide sequences for potential cleavage sites using a published a cleavage prediction algorithm [59]. We then took these peptide sequences and tested them for cleavage by NSP5 using a loss of fluorescence resonance energy transfer (FRET) fluorescence assay. In brief, potential cleavage sites were inserted between a FRET pair and then this construct was co-transfected along with plasmids expressing either NSP5 or the NSP5^{C145A}, with loss of FRET signal only after wild type NSP5 expression as indicative of cleavage. This assay can be used to screen potential peptides encoded by host target sequences. To this end four sequences taken from SARS-CoV-2-pp1AB polyprotein, which is normally cleaved by NSP5, were cleaved as expected and as demonstrated by loss of FRET signal (**Fig 6C and 6D**). Testing of sequences from human CDKN2AIP, CWC22, FANCD2, and P53 proteins indicated NSP5 cleavage of one CWC22 and two FANCD2 peptide sequences (**Fig 6C and 6D**). Neither CDKN2AIP nor P53 sequences tested were cleavable by NSP5 in our assay and their depletion in the SILAC data may represent indirect effects of NSP5 activity. CWC22 is a component of the RNA spliceosome required for pre-mRNA splicing via promotion of exon-junction complex assembly [60,61]. FANCD2 is activated by ATM and localizes at BRCA1 foci during DNA damage [62]. These data suggest that SARS-CoV-2 may target host RNA splicing and DNA damage pathways via NSP5-mediated reduction in key proteins, namely CWC22 and FANCD2, that are involved in these processes.

As noted, viral protease inhibitors are a powerful class of drugs that potently block viral replication by preventing processing of viral polyproteins into functional subunits. Inhibition of viral proteases should also prevent cleavage of host proteins which may serve to blunt toxic effects on infected cells. GC376 [63] is a NSP5 protease inhibitor developed against feline coronavirus, the causative agent of fatal feline infectious peritonitis. Recent reports showed GC376 to be effective against SARS-CoV-2 NSP5. We tested the effect of GC376 on the cleavage of the ORF1ab-2 and FANCD2-2 peptide sequences. Using a range of concentrations up to 80 μM, ORF1ab-2 showed a modest inhibition of NSP5 as compared to NSP5^{C145A} in the FRET assay. FANCD2-2 showed a dramatic reduction in cleavage by NSP5 even at concentrations of 20 μM (**Fig 6E and 6F**). These data support GC376 inhibition of SARS-CoV-2 NSP5 action on viral and human host protein sequences cleavable by the viral protease.

Discussion

Here we present a compendium of human host proteins adjacent to 17 SARS-CoV-2 viral proteins, with a goal to offer insight into potential mechanisms that these viral proteins may engage during pathogenesis. These data encompass the less well understood SARS-CoV-2 accessory factors and predict the localization of each these viral proteins as well as identify significant adjacencies to proteins that mediate core cellular processes, including translation, signaling, RNA interactions, and intracellular transport. Further validation of viral protein localization was confirmed by immunofluorescent microscopy for select proteins in infected cells. For translation, SARS-CoV-2 NSP1 was found to be adjacent to subunits of the EIF3 translation initiation complex and proved a broad inhibitor of translation. For innate immune signaling, viral ORF6 was found proximal to the RLR pathway component, MAVS, with ORF6 potently inhibiting induction of the *IFNB1* promoter. Integration of GWAS data in COVID-19 identified SNPs associated with natural variation in the expression of specific genes, including the viral M protein-proximal TRIM4 activator of type I interferon, that may contribute to disease susceptibility differences in the human population. Comparing wild type NSP5 with its catalytically inactive point mutant helped identify proteins whose levels were decreased by this viral protease and nominated cleavage sequences in human CWC22 and FANCD2,

implicating specific candidates for viral disruption of normal pre-mRNA splicing and DNA damage pathways, respectively. We also observed a number of SARS-CoV-2 proteins (M, NSP2, NSP9, NSP15, ORF6, ORF7a, ORF7b, ORF8, ORF9c, and ORF10) vicinal to nuclear pore proteins. Given that coronavirus replication takes place exclusively in the cytosol of cells, these interactions, if functional, might point to a viral role in disrupting nuclear import/export. This is further supported by several lines of genetic evidence. GWAS data suggest the importance of nuclear pore component Nup43 and overlap of our proteomics data with whole genome CRISPR screen hits from recently studies [34–36] suggests that the mRNA export factors MCM3AP and NXF1 are important for viral replication. Taken together, these data indicate potential intracellular locations and candidate functions of the SARS-CoV-2 viral proteins studied and provide a resource for future studies of pandemic coronaviruses.

SARS-CoV-2, as the etiological agent of COVID-19, joins SARS-CoV and MERS-CoV as an important coronavirus pathogen. Very minor mutations in the viral spike protein [64–67] along with a number of animal reservoirs in endemic regions represent a significant risk for new pandemic coronavirus strains to emerge [68], underscoring the need to understand coronaviral accessory protein functions and virus-host interactions. Comparative studies that analyze multiple coronaviruses [19,69], including both highly pathogenic and nonpathogenic, will be very beneficial to understanding what can identify new possibly pandemic virus strains. Such resources may allow the research community to not only address current concerns and also provide insight to address with newly emerging coronaviruses in the future. Currently, the ability of S proteins capable of binding to human ACE2 receptor has been seen both in highly pathogenic coronaviruses such as SARS-CoV and SARS-CoV-2 as well as in relatively non-pathogenic coronaviruses such as HCoV-NL63 [70]. Thus, comparing the actual molecular interactions and effects of viral proteins on the host between pathogenic and non-pathogenic virus strains may provide actual insight on what makes certain coronaviruses more dangerous and highlight critical virus-host interactions that may be targeted to reduce disease.

The viral envelope of SARS-CoV-2 must contain the proper structural components comprised of S, E, M, and N with a completed viral genome [71] and replication of both subgenomic and genomic RNA occurs in membranous compartments [72]. Accordingly, coronaviruses devote substantial portions of its large genome to manipulating host processes involved in ER-Golgi transport and endocytic and exocytic activity, which was captured in the proximal interactome. We also found evidence of the interaction of SARS-CoV-2 with MHC class I molecules with M, ORF7a, ORF7b, ORF8, and ORF10. Down-regulation of surface expressed proteins has been reported for SARS [73]. It is still an open question to how SARS-CoV-2 affects surface expression of important host receptors, which viral proteins affect this process, and the effects on virus replication and disease.

Translation inhibition is a general strategy utilized by many virus families including other RNA viruses like orthomyxoviruses [74], picornaviruses [75,76], rhabdoviruses [77], and togaviruses [78]. Host translational blockade may broadly block antiviral responses and can also cause affect the viability of the infected cell. Some but not all viruses have strategies to overcome translational shutoff, biasing translation of viral mRNA, including the use of IRES elements [79]. Lung tissue from COVID patients, in particular, displayed proteomic changes associated with translation inhibition [80]. NSP1 from both SARS-CoV-1 [10] and SARS-CoV-2 [39,40] have been shown to be potent inhibitors of host translation and are thought to do so using at least two mechanisms: binding to and inhibition of EIF3 translation initiation complex and direct cleavage of host mRNAs. Cryo-EM studies place a domain of NSP1 as sitting in the mRNA channel of the 40S ribosome. Our proximity proteomics data shows NSP1 of SARS-CoV-2 binding to a significant number of EIF3 complex subunits and we demonstrate that NSP1 is able to block translation of capped transcripts as well as transcripts

containing host and viral IRES elements. We also observe, as another study has shown[39], that NSP1-induced translational shutoff affects host and viral transcripts containing the leader sequence found in the viral 5' UTR. This element exists on all genomic and subgenomic viral RNAs [81]. Whether other SARS-CoV-2 factors are necessary to overcome NSP1 translational inhibition or if, during viral replication, the large number of viral transcripts simply outcompetes host transcripts, as seen in vesicular stomatitis virus [82], remains to be determined. A recent study [80] from autopsies of COVID patients characterized whole proteome changes in multiple organs.

Innate immune signaling is a central mechanism of host cell response to viral infection. ORF6 of SARS-CoV-1 [12,50] and SARS-CoV-2 [49] were shown to be potent inhibitors of such antiviral signaling. One proposed mechanism is that ORF6, through association with specific NPCs (Rae1-Nup98), blocks import of activated transcription factors needed to induce *IFNB1* transcripts and other primary interferon-stimulated genes. In this regard, we identified MAVS proximal to ORF6 and ORF9b. We observed that ORF6, but not ORF9b, inhibited RLR signaling downstream of RIG-I RNA-binding. However, Nup98/Rae1 is most commonly associated with export of RNAs from the nucleus [83,84]. This study has identified both NSP15 and ORF6 as proximal interactors with many components of the nuclear RNA export pathway including Nup43, Nup98, Rae1, MCM3AP, NXF1, and others. Additional evidence, from both GWAS studies and genome-wide CRISPR screens has identified RNA export as important for severe disease and viral replication. It will be important to reexamine the mechanism of ORF6 in impairment of interferon signaling as well as examining the effects NSP15 may have on RNA export. Thus, targeting the ability of the virus to disrupt host RNA export pathways may be an attractive therapeutic target and warrants more attention. The effects of SARS-CoV-2 on host processes identified in this study are summarized in **Fig 7**.

Viral proteases, such as SARS-CoV-2 NSP5 studied here, have been shown to be potent antiviral targets [85]. These proteases are essential for viral replication and escape has proven difficult in resistance studies [86]. Coronaviruses encode two proteases NSP3 and NSP5, with NSP5 classified as the main protease. They are both necessary for the processing of the ORF1ab polyprotein containing the viral replicase proteins. NSP5 shows similarity to proteases found in picornaviruses and noroviruses [87]. Beyond their importance in viral replication, these viral proteases can target host proteins containing their target residues [88]. NSP5 recognizes certain glutamine-serine/alanine/glycine residues, with added specificity being determined by two to three flanking residues [59]. Picornavirus virulence has been shown to be mediated in part by 3C protease cleavage of host proteins [55]. Using both BioID and SILAC metabolic labeling followed by mass spectrometry, we sought to identify candidate host proteins and use a modified FRET-based cleavage assay to determine if these candidates contained sequences cleavable by NSP5. We identified human CWC22 and FANCD2 as candidates; both proteins contained sequences that could be cleaved by NSP5 in an assay used here which can be used to rapidly assess other potential host targets. The proteomic studies also identified clusters of host factors involved in DNA damage and repair and RNA splicing. Furthermore, we show the effects of GC376 [63], a protease inhibitor of feline coronavirus, displays evidence of inhibition of NSP5 cleavage activity. Consistent with this, GC376 has been shown to block viral replication of SARS-CoV-2 in early studies [89] and we observe that this protease inhibitor blocks NSP5 cleavage of both host and viral target peptide sequences.

The global impacts of the SARS-CoV-2 pandemic have focused attention on identifying new treatments and interventions. Given both the newness of the virus and the relative dearth of research into human coronaviruses, it is important that many resources are generated to better understand aspects of the virus-host interaction. The present work contains a proximal proteomic resource for 17 SARS-CoV-2 viral proteins and combining such proximal
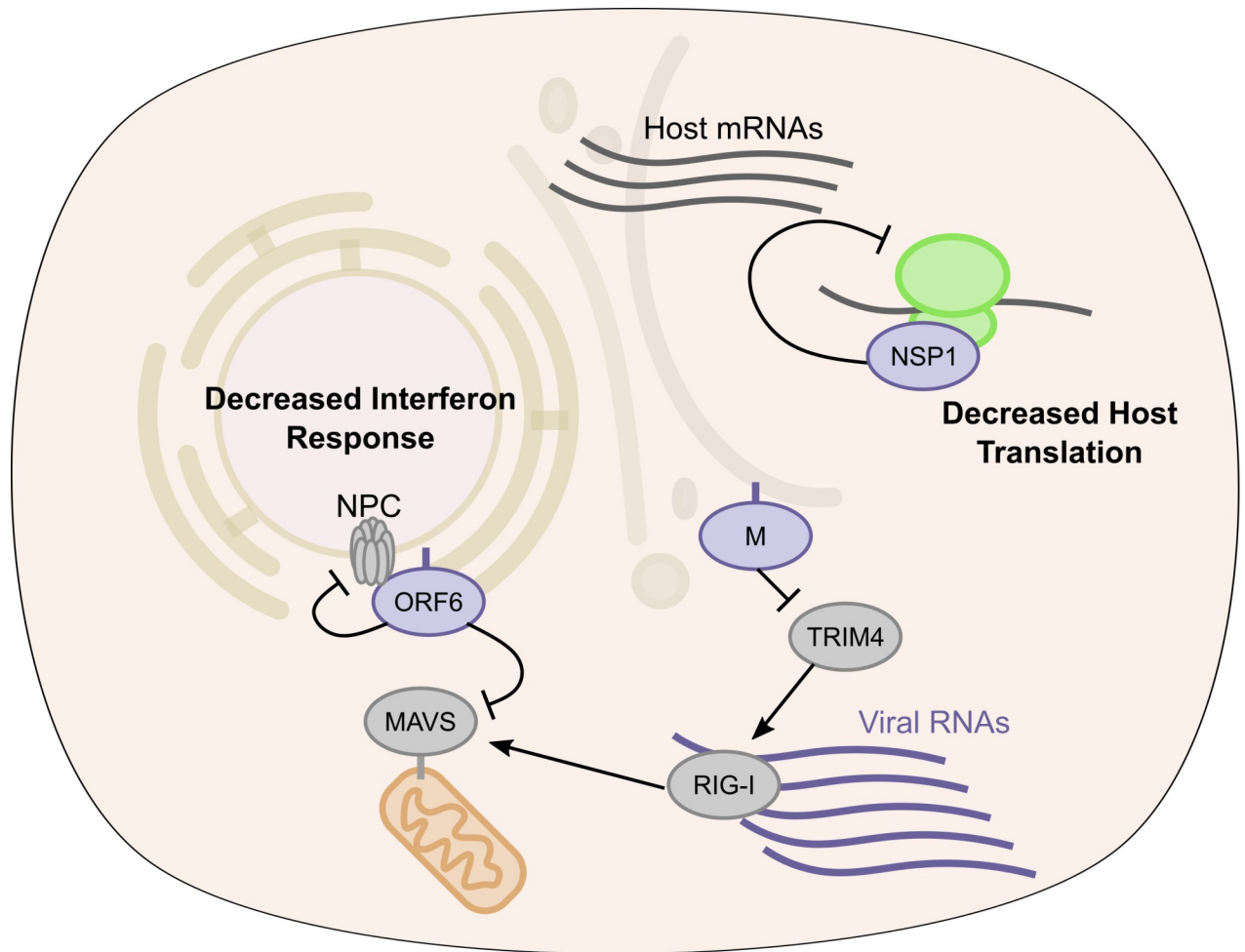
Fig 7. SARS-CoV-2 proximal proteins in translation and interferon activation. Model of SARS-CoV-2 antagonism of host antiviral response. ORF6 protein inhibits RLR signaling leading to decreased type I interferon and ISG transcription, M protein through TRIM4 interactions may also alter host response. NSP1 disrupts host translation of transcripts containing both ISG 5' UTR and stress responsive IRES elements.

proteomics with TAP-based proteomics may be helpful in leveraging the strengths associated with each technique. While proximity proteomics can identify transient, indirect, and weak binding events, including those dependent on intact membranes, TAP-based approaches can focus attention on complexes of proteins that stably associated with each other. We validate the quality of the present proximity data set by corroborating spatial insights with biochemical fractionation experiments. Taken together with other efforts to generate high-quality resources, these data should prove helpful in both generating hypotheses and better understanding dynamics of virus-host interactions in regards to human disease.

Materials and methods

Cell culture

HEK293T were obtained from Takara Bio and were cultured on DMEM 10% FBS, 1% Penicillin/Streptomycin and grown at 37C, 5% CO_2. For SILAC experiments [90], the cells were cultured in a medium containing [13C6,15N2]-lysine and [13C6]-arginine for at least 2 weeks to promote complete incorporation of the stable isotope-labeled amino acids. Cells were

tested for mycoplasma prior to experiments using MycoAlert Mycoplasma Detection kit (Lonza).

Transfection, biotin labeling, and streptavidin pulldown

All viral expression constructs were obtained from Addgene [17]. A summary of all plasmids used in this study is listed in Table 1. HA-BASU was cloned in frame with either an N-terminal or C-terminal linker as indicated. For BioID experiments 5x106 HEK293T were plated and transfected with 5ug of each viral expression plasmid. 24 hours post transfection, biotin was added (50 uM final concentration) for 4 hours, then media was exchanged twice with DPBS and the cells harvested and lysed in RIPA buffer (Thermo Scientific) supplemented with protease inhibitors. Lysates were sonicated and then, using the Kingfisher Flex automated purification, incubated for six hours with 100 uL of ReSYN (ReSYN Biosciences) streptavidin microparticles and then washed sequentially with 2% LDS buffer, Triton X-100 buffer (1% Triton X-100 0.1%, Sodium Deoxycholate 500mM, 1mM EDTA, 50mM HEPES pH 7.5), Igepal Wash Buffer (0.5% Igepal, 0.5% Sodium Deoxycholate, 10mM TRIS pH 7.5, 333.3mM LiCL, 20mM EDTA), and deposited into 50mM TRIS pH 7.4. Samples were washed with automated mixing for 30 minutes for each step. A portion of the whole cell lysate was saved and ran on

Table 1. Plasmids used in this study.

Plasmid	Origin	Source
eGFP-HA_BASU	This study	Addgene 153996
SARS_COV2_nsp1_HA_BASU	This study	Addgene 154071
SARS_COV2_nsp2_HA_BASU	This study	Addgene 153988
SARS_COV2_nsp5_HA_BASU	This study	Addgene 153987
SARS_COV2_nsp9_HA_BASU	This study	Addgene 153986
SARS_COV2_nsp14_HA_BASU	This study	Addgene 153993
SARS_COV2_nsp15_HA_BASU	This study	Addgene 153985
SARS_COV2_M_HA_BASU	This study	Addgene 153990
SARS_COV2_N_HA_BASU	This study	Addgene 153989
SARS_COV2_ORF3a_HA_BASU	This study	Addgene 153994
SARS_COV2_ORF3b_HA_BASU	This study	Addgene 153992
SARS_COV2_ORF6_HA_BASU	This study	Addgene 153984
SARS_COV2_ORF7a_HA_BASU	This study	Addgene 153983
SARS_COV2_ORF7b_HA_BASU	This study	Addgene 153995
SARS_COV2_ORF8_HA_BASU	This study	Addgene 153982
SARS_COV2_ORF9b_HA_BASU	This study	Addgene 153981
SARS_COV2_ORF9c_HA_BASU	This study	Addgene 153991
SARS_COV2_ORF10_HA_BASU	This study	Addgene 153980
pRF-CRPV-IGR-IRES		
pRF-ISG15-CRPV-IGR-IRES	This study	
pRF-IFIT1-CRPV-IGR-IRES	This study	
pRF-APAF1-CRPV-IGR-IRES	This study	
pRF-XIAP1-CRPV-IGR-IRES	This study	
pRF-5'SARS2-CRPV-IGR-IRES	This study	
ECFP-TevS-YPET		Addgene 100097
pMODnano-IFNBpro	This study	
pGL4.13[luc2/TK]	Promega	Cat# E5061

Table 1. A list of plasmids used in this study along with source for plasmid where available.

https://doi.org/10.1371/journal.ppat.1009412.t001

Table 2. Antibodies Used in This Study.

Antibody	Origin	Catalog Number
HA-tag (C29F4) (Rabbit)	Cell Signaling Technologies	Cat# 3724; RRID:AB_1549585
HA-tag (Mouse)	Abcam	Cat# ab130275; RRID:AB_11156884
HA-tag (Mouse)	Cell Signaling Technologies	Cat # 2367
β-Actin	Sigma-Aldrich	Cat# A1978; RRID:AB_476692
Calnexin	Cell Signaling Technologies	Cat# 2679
Lamin A/C	Cell Signaling Technologies	Cat# 4777
NSP1	mrcppureagents	Cat# SARS-CoV2 NSP1
NSP14	mrcppureagents	Cat# NSP14 SARS-CoV2 PC
ORF6	mrcppureagents	Cat# SARS-CoV2 ORF6
EIF3G	Proteintech	Cat# 11165-1-AP
Nup98	Atlas Antibodies	Cat# HPA074810
Rae1	Atlas Antibodies	Cat# HPA048795
MCM3AP	Atlas Antibodies	Cat# HPA021527
Goat Anti-Sheep 488	Thermofisher	Cat# A-11015
Goat Anti-Rabbit 647	Thermofisher	Cat# A32795
800CW Goat anti-Mouse IgG (H + L)	LI-COR	Cat# 926–32210; RRID:AB_621842
800CW Goat anti-Rabbit IgG (H + L)	LI-COR	Cat# 926–32211; RRID:AB_621843
680RD Goat anti-Mouse IgG (H + L)	LI-COR	Cat# 926–68070; RRID:AB_10956588
800CW Streptavidin	LI-COR	Cat# 926–32230;
		RRID:AB_2877131
680RD Streptavidin	LI-COR	Cat# 926–68079; RRID:AB_2877132
680RD Goat anti-Rabbit IgG (H + L)	LI-COR	Cat# 926–68071; RRID: AB_10956166

Table 2. A list of primary and secondary antibodies used in this study with source and catalog number included where applicable.

SDS-PAGE gel, transferred to PVDF, and then probed with anti-HA antibody with 800CW anti-rabbit (LICOR) secondary along with 800CW streptavidin dye (LICOR) to confirm viral protein expression and total biotinylation. A summary of antibodies used in this study is listed in Table 2. For SILAC experiments with NSP5 and C145A, approximately 2×106 cells were harvested, washed with ice-cold PBS for three times, and lysed by incubating on ice for 30 min with CelLytic M (Sigma) cell lysis reagent containing 1% protease inhibitor cocktail. The cell lysates were centrifuged at 7,000g and at 4˚C for 15 min, and the resulting supernatants collected.

Sample preparation for mass spectrometry

After wash and purification samples contained bound proteins on beads in TRIS buffer. The protein on the beads were reduced with dithiothreitol, and alkylated with iodoacetamide. The processed proteins were subsequently digested with Trypsin/Lys-C (Promega) at an enzyme/substrate ratio of 1:100 in 50 mM NH4HCO3 (pH 8.5) at 37˚C for overnight.

For SILAC samples, the protein lysates prepared from cells with WT or mutant NSP5 were combined at 1:1 ratio (by mass), and 30 μg of the mixed protein lysate was loaded onto a 10% SDS-PAGE gel. After electrophoresis, the gel lanes were cut into 11 slices according to apparent molecular weight ranges of proteins (< 20, 20–25, 25–30, 30–37, 37–42, 42–50, 50–62, 62–75, 75–100, 100–150, >150 kDa), reduced in-gel with dithiothreitol, and alkylated with iodoacetamide. The processed proteins were subsequently digested in-gel with Trypsin/Lys-C (Promega) at an enzyme/substrate ratio of 1:100 in 50 mM NH4HCO3 (pH 8.5) at 37˚C for overnight. Subsequently, peptides were recovered from gels with a solution containing 5%

acetic acid in H2O and then with a solution containing 2.5% acetic acid in an equi-volume mixture of CH3CN and H2O.

All the resulting peptide mixture was subsequently dried in a Speed-vac, and desalted by employing OMIX C18 pipet tips (Agilent Technologies, Santa Clara, CA). LC-MS/MS experiments were conducted on a Q Exactive Plus mass spectrometer equipped with an UltiMate 3000 UPLC system (Thermo Fisher Scientific).

LC-MS/MS analysis

Samples were automatically loaded at 3 μL/min onto a precolumn (150 μm i.d. and 3.5 cm in length) packed with ReproSil-Pur 120 C18-AQ stationary-phase material (5 μm in particle size, 120 Å in pore size, Dr. Maisch). The precolumn was connected to a 20-cm fused-silica analytical column (PicoTip Emitter, New Objective, 75 μm i.d.) packed with 3 μm C18 beads (ReproSil-Pur 120 C18-AQ, Dr. Maisch). The peptides were then resolved using a 180-min gradient of 2–45% acetonitrile in 0.1% formic acid, and the flow rate was maintained at 300 nL/min.

The mass spectrometer was operated in a data-dependent acquisition mode. Full-scan mass spectra were acquired in the range of m/z 350–1500 using the Orbitrap analyzer at a resolution of 70,000 at m/z 200. Up to 25 most abundant ions found in MS with a charge state of 2 or above were sequentially isolated and collisionally activated in the HCD cell with a normalized collision energy of 28 to yield MS/MS.

Database search

Maxquant, Version 1.5.2.8, was used to analyze the LC-MS and MS/MS data for protein identification and quantification [91]. The database we used for search was human IPI database, version 3.68, which contained 87,061 protein entries. The maximum number of miss-cleavages for trypsin was two per peptide. Cysteine carbamidomethylation and methionine oxidation were set as fixed and variable modifications, respectively. The tolerances in mass accuracy were 20 ppm for both MS and MS/MS. The maximum false discovery rates (FDRs) were set at 0.01 at both peptide and protein levels, and the minimum required peptide length was 6 amino acids. Spectral match assignment files were collapsed to the gene level and false positive matches and contaminants were removed. SAINT analysis [18] was run with the following parameters: 10,000 iterations, LowMode ON, Normalize ON and the union of MinFold ON and OFF. Minimum interactome inclusion criteria were SAINT $\geq$ 0.9, fold change over matched cell type control $\geq$ 4. Low normalized spectral count proteins were removed.

Gene ontology

Gene Ontology (GO) term analyses were produced using the clusterProfiler R package [92]. Proteins with SAINT score $\geq$0.9 were classified as likely interactions and used to identify enriched GO terms for the individual SARS-CoV2 protein interactomes. Highly redundant GO terms were removed for readability. Bar plots and heatmaps were produced with the ggplot2 [93] and heatmap [94] packages respectively in R.

Host-virus interaction network

Host-virus interaction network produced from BASU BioID interactions with a SAINT score $\geq$ 0.9 in Cytoscape [95]. The network was further curated to emphasize the significantly enriched GO terms for each SARS-CoV-2 protein. Edges denoting Host-virus protein interactions are indicated in red. Host-host interactions were determined from high confidence

(>0.700) STRING database interactions obtained from experimental evidence and database interactions for all of the curated proteins. Cell endogenous protein interactions are denoted by grey edges. Clusters were highlighted based on highly enriched GO terms for SARS-CoV-2 proteins.

Cellular fractionation

Cellular fractionations were generated using a previous protocol [96] with minimal modification. Cells were transfected as described previously. Cell pellets were split into three separate samples. The first was lysed using RIPA buffer (Thermo Scientific) and was labeled whole cell lysate (WCL). The second sample was resuspended in buffer containing 0.3% Igepal, 10mM HEPES, 10mM KCl, 1.5mM MgCl2. Sample was pelleted at 1500g and supernatant was collected and labeled cytoplasm/membrane fraction. The remaining pellet was washed once and then lysed in RIPA and labeled nuclear fraction. The third sample was lysed in buffer containing 100ug/mL Digitonin, 50mM HEPES, and 150mM NaCl. Sample was pelleted at 2400g and supernatant was collected and labeled cytoplasm fraction. The remaining pellet was washed once and lysed in RIPA and labeled nuclear/membrane fraction. Equal volumes of each fraction along with 20ug of WCL were loaded and ran in a 4–12% Tris-Bis Polyacrylamide Gel (Invitrogen). Samples were transferred to PVDF and blotted for HA (Viral Proteins), Alpha tubulin (cytoplasm control), calnexin (membrane control), Histone H3 (nuclear control) (Table 2).

Viral infection and immunofluorescent microscopy

SARS-CoV-2 (USA-WA1/2020) viral stocks were obtained from BEI and propagated on VeroE6 cells (ATCC) in DMEM with 2% FBS. The virus was passaged 3 times on VeroE6 cells and titers were calculated via plaque assay. Viral stocks were subject to sequencing to confirm that no furin cleavage site or other mutations were acquired during passaging. A549-ACE2 cells were donated from Ralf Bartenschlager's lab and cultured in DMEM supplemented with 10% FBS and 2% Geneticin (Thermo Fisher). Cells were plated on 8-well chamber slides (ibidi) 24 hours before infection. On the day of infection, cells were inoculated with SARS-CoV-2 at an MOI of 0.1 and incubated at 37˚C under 5% CO2. At 48 hours, chamber slides were washed with PBS and submerged in 4% PFA for 30 minutes according to BSL3 fixation and inactivation protocols. All SARS-CoV-2 live virus experiments were conducted under BSL3 conditions at Stanford University. Cells were then permeabilized with PBS containing 2% BSA and 0.2% Triton X-100. Staining of primary antibodies against both viral (NSP1, NSP14, ORF6) and host (EIF3A, Nup98, Rae1, MCM3AP) targets was done by incubating the cells with dilutions (1:500 for all viral targets, EIF3A 1:200, Nup98 1:100, Rae1 1:100, MCM3AP 1:100) at room temperature for 30 minutes followed by 3 washes with PBS containing 2% BSA and 0.2% Triton X-100. Goat anti-sheep conjugated to Alexa Fluor 488 was used to detect viral primary antibodies and Goat-anti rabbit conjugated to Alexa Fluor 657 was used to detect the host primary antibodies. Staining of secondary antibodies was done at 1:500 dilution for 30 minutes and followed by 3 washes and then Hoescht staining was performed at a final concentration of 1ug/mL for 5 minutes and followed by 3 washes with PBS. After the final PBS wash, fresh PBS was added to the chambers and images were taken on Zeiss 880 confocal microscope using the 63X oil immersion lens. Additional antibody information is listed in Table 2.

GWAS COVID risk SNP analysis

COVID GWAS datasets were sourced from COVID-19 Host Genome Initiative (https://www.covid19hg.org/), the Ellinghaus, Degenhardt, et al study [27], and from the UK Biobank

(https://grasp.nhlbi.nih.gov/Covid19GWASResults.aspx). GWAS hits were converted to hg19 coordinates and phenotypes for each GWAS study were noted. Gene locations are sourced from gencode v19 exon coordinates. The GWAS SNPs were then expanded by LD r2 > 0.8 with phase 1000 Genomes LD information using LDlinkR [97], and phase 1 1000 Genome LD information using HaploReg [98]. The expanded SNP list was then overlapped with GTEx lung, spleen, blood, cis-eQTL data, DICE cis-QTL data, and eQTLGen cis-eQTL data [29,99,100].

HiChIP data processing and virtual 4C visualization

HiChIP all valid pair matrices for GM12878, Naïve T cells, Th17 cells and Treg were downloaded from GEO (GSE101498,[101]). v4C plots were generated from HiChIP valid pair matrices. The interaction profile of a specific 5-kb bin containing the TRIM4 anchor was then plotted in R. H3K27ac ChIP-seq peaks for GM12878, Naïve T cells, Tregs and T helper cells were downloaded from ENCODE as 1d peak sets. FitHiChIP pipeline was used to call loops with 5kb bin, peak-to-all interaction type, loose background, and FDR < 0.01[102]. The merged significant interaction files from FitHiChIP pipeline along with corresponding ATAC–seq profiles were visualized in WashU Epigenome web browser. Browser shots from WashU track sessions were then included in the v4C and interaction map anecdote.

Luciferase assays

For NSP1 translation assays, in-vitro transcribed transcripts were generated by first PCR amplifying DNA containing T7 promoter followed by UTR or IRES elements and firefly or renilla luciferase. Second, using HiScribe T7 ARCA mRNA Kit (with tailing) (NEB) capped and polyadenylated transcripts were synthesized. 5x105 293T cells were transfected with 2ug of plasmids expressing either GFP or NSP1 and then incubated overnight. The next day the cells were transfected with 2ug of the corresponding IVT transcripts and were harvested 8 hours post second transfection. Cells were harvested with 400ul of Passive Lysis Buffer (Promega) and quadruplicate samples were plated on an opaque 96 well plate. 50ul of LARII firefly luciferase substrate (Promega) was added and the plate was read on the luminescence setting of the Spectramax i5 plate reader. 50ul of Stop & Glo renilla luciferase substrate was then added and the plates reread. For IFNB1 promoter activity assays, 2.5x105 293T cells were transfected with 2ug of plasmids expressing either ORF6 or ORF9b along with 1ug of plasmid containing nanoluciferase under the control of a the human IFNB1 promoter and 50ng of a plasmid containing firefly luciferase under the control of the constitutive TK promoter and either 1ug of empty vector or 1ug of a plasmid expressing the 2-CARD domain of RIG-I. 24 hours post transfection cells were harvested in 200ul of Passive Lysis Buffer and triplicate samples were plated on an opaque 96 well plate. Nano-Glo Dual-Luciferase Reporter Assay System (Promega) was used to obtain a firefly luciferase reading for IFN-beta promoter activity normalized to the firefly luciferase transfection control.

NSP5 cleavage site prediction

Protein sequences for hits from SILAC and BASU-BioID proteomics experiments were run through the NetCorona algorithm [59] using the web application: (https://services.healthtech. dtu.dk/service.php?NetCorona-1.0). For Coronavirus Polyprotein controls, the SARS-COV-2 ORF1ab protein sequence (from Uniprot Fasta UP000464024) was run through the NetCorona web application. A previously tested SARS-COV-1 sequence [59], VATLQAENV, was found to be shared in the SARS-COV-2 protein sequence and was also used as a control.

FRET-based NSP5 cleavage assay

Predicted NSP5 cleavage sites were cloned into ECFP-TevS-YPET (Addgene Plasmid #100097) [103] Briefly, the plasmid was re-cloned to put the Tev Protease Site between an XbaI site and a BsiWI site. Cleavage sequences were cloned in between XbaI and BsiWI sites, restoring the XbaI and BsiWI sites, with one Glycine on each side of the predicted NSP5 cleavage sequences. This approach was based on a cloning strategy used previously to study norovirus protease cleavage sites [104]. For protease cleavage assay, 3x10^4 HEK 293T cells were plated in DMEM + 10% FBS into 96 well black, clear bottom plates (Greiner). 24 hours later, cells were transfected in quadruplicate with 0.1ug of FRET plasmid containing the NSP5 cut-site and 0.1ug of either WT-NSP5 or mutant NSP5C145A expression plasmids [17] with Lipofectamine 3000 following manufacturer protocol for 96-well plates. 24 hours later, media was removed and PBS was added to wells, and wells were imaged on a Spectramax i5 instrument (Molecular Devices) with the following wavelengths: 420/485 nm for ECFP, 485/535 nm for YPET and 420/535 nm for FRET as previously described [103]. After background subtraction of un-transfected wells, FRET efficiency was calculated as FRET/ECFP.

Quantification and statistical analysis

Gene ontology adjusted p-values were produced using the Benjamini-Hochberg method. For heatmaps, a threshold was set whereby at least one protein had a significant score for the presented GO terms. A -Log10 p-Value threshold of >5 for M proteins and >1.3 for all others was used for 'non-stringent' heatmaps (with the exception of molecular function heatmaps, which uses an M protein threshold of 3) and -Log10 p-Value > 5 for M proteins and >3 for all others was used for more stringent heat maps presented in Figs 1D and 2A.

Graphed data are expressed as mean ± SEM and sample size (N) represents independent experiments as noted. Statistical analysis was performed in GraphPad Prism 7 and described in the figure legends.

Student's t test was performed comparing the means between GFP controls and the experimental conditions where N is three independent experiments. For IFNB1 reporter assays, empty vector without RIG-I 2-CARD was set to one and all other conditions are relative to that empty control and is the average of three independent experiments.

NSP5 FRET-cleavage was calculated as described previously and Student's t test was performed comparing the means between NSP5C145A mutant and NSP5 WT where N is three independent experiments.

Supporting information

S1 Fig. Western Blots of samples used for BIOID-MS. A portion of each biological replicate (depicted as 1, 2, or 3) for all SARS-CoV-2 BASU fusion expression constructs were run on SDS-PAGE gels and samples were blotted for HA, Streptavidin, or Actin as a loading control. GFP controls (high, med, low) were used as background and for blot-to-blot comparisons. (TIF)

S2 Fig. Interaction Maps for SARS-CoV-2 proteins. Interaction maps for all SARS-CoV-2 proteins(A-Q) from this study. All host proteins with SAINT scores ≥ 0.9 were included. (PDF)

S3 Fig. Comparison of BioID-MS and TAP-MS. Comparison of hits (SAINT score ≥0.9) between the current study and AP-MS (Gordon et al. 2020). (TIF)

S4 Fig. Proteomic Hits from TAP-MS Molecular Function. GO-term analysis Molecular Function (MF) (full list) of proteomic hits from Gordon et al. 2020.
(TIF)

S5 Fig. Proteomic Hits from TAP-MS Biological Processes. GO-term analysis Biological Processes (BP) (full list) of proteomic hits from Gordon et al. 2020.
(TIF)

S6 Fig. Proteomic Hits from TAP-MS Cell Component. GO-term analysis Cell Component (CC) (full list) of proteomic hits from Gordon et al. 2020.
(TIF)

S7 Fig. Comparison of GO-term cell component analysis. Comparison of cell component GO-term analysis between BIOID and TAP-MS (IP) data sets. Enrichment scores are given as -Log10 p-values.
(TIF)

S8 Fig. SARS-CoV-2 association with druggable host genes. Interaction maps of SARS-CoV-2 with druggable proximal host proteins. Intersection of SARS-CoV-2 proximal network generated in this study with the druggable genome databases from https://pharos.nih.gov/.
(TIF)

S9 Fig. Western blot of SILAC samples expressing WT or mutant (C145A) NSP5. Western blot of samples used for SILAC whole proteome enrichment/depletion. Light (L) or Heavy (H) isotope labelled amino acids is indicated. Cells transiently expressing NSP5, but not C145A mutant, displayed cell toxicity as early as 24 hours post transfection.
(TIF)

S1 Table. Proximal proteomic interactors of SARS-CoV-2 Proteins. List of proximal proteomic interactors of SARS-CoV-2 proteins. Combined list contains all hits with SAINT score ≥0.9. Fold change values represent fold change relevant to two closest GFP control samples in regards to expression levels. Table of unique interactors lists host proteins only found as proximal interactors to a single SARS-Cov-2 protein. Remaining tables list full protein lists of each SARS-CoV-2 protein with peptide counts found for each host proteins and all GFP controls.
(XLSX)

S2 Table. GO-term enrichment analysis. Complete list for each viral protein of Molecular Function, Biological Process, and Cellular component GO-term analysis. Enrichment scores are given as -Log10 p-values
(XLSX)

S3 Table. Druggable Gene List of SARS-CoV-2 Proximal Proteome. Complete table of host genes found as proximal interactors of SARS-CoV-2 proteins. Drugs listed are reported inhibitors of the indicated host protein.
(XLSX)

S4 Table. Key Oligos Used in This Study. Table of oligos used in this study for confirmation of BASU-tagged constructs, sequences used in translation reporter assay, and peptide sequences used in FRET-cleavage assay.
(XLSX)

Acknowledgments

Author Contributions

Conceptualization: Jordan M. Meyers, Muthukumar Ramanathan.

Formal analysis: Laura Donohue, Ian Ferguson, Margaret G. Guo, Xue Yang, Yang Zhao.

Investigation: Jordan M. Meyers, Muthukumar Ramanathan, Ronald L. Shanderson, Aimee Beck, Deepti S. Rao, Weili Miao, David Reynolds, Xue Yang, Yen-Yu Yang.

Methodology: Jordan M. Meyers, Muthukumar Ramanathan, Ronald L. Shanderson.

Project administration: Jordan M. Meyers.

Supervision: Catherine Blish, Yinsheng Wang, Paul A. Khavari.

Visualization: Ronald L. Shanderson, Laura Donohue, Ian Ferguson, Margaret G. Guo.

Writing – original draft: Jordan M. Meyers.

Writing – review & editing: Jordan M. Meyers, Muthukumar Ramanathan, Ronald L. Shanderson, Yinsheng Wang, Paul A. Khavari.

References

1. Ksiazek TG, Erdman D, Goldsmith CS, Zaki SR, Peret T, Emery S, et al. A novel coronavirus associated with severe acute respiratory syndrome. N Engl J Med. 2003; 348(20):1953–66. https://doi.org/10.1056/NEJMoa030781 PMID: 12690092

2. Zaki AM, van Boheemen S, Bestebroer TM, Osterhaus AD, Fouchier RA. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. N Engl J Med. 2012; 367(19):1814–20. https://doi.org/10.1056/NEJMoa1211721 PMID: 23075143

3. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet. 2020; 395(10223):497–506. https://doi.org/10.1016/S0140-6736(20)30183-5 PMID: 31986264

4. Roux KJ, Kim DI, Raida M, Burke B. A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. J Cell Biol. 2012; 196(6):801–10. https://doi.org/10.1083/jcb.201112098 PMID: 22412018

5. Rigaut G, Shevchenko A, Rutz B, Wilm M, Mann M, Seraphin B. A generic protein purification method for protein complex characterization and proteome exploration. Nat Biotechnol. 1999; 17(10):1030–2. https://doi.org/10.1038/13732 PMID: 10504710

6. Ramanathan M, Majzoub K, Rao DS, Neela PH, Zarnegar BJ, Mondal S, et al. RNA-protein interaction detection in living cells. Nat Methods. 2018; 15(3):207–12. https://doi.org/10.1038/nmeth.4601 PMID: 29400715

7. Narayanan K, Huang C, Makino S. SARS coronavirus accessory proteins. Virus Res. 2008; 133 (1):113–21. https://doi.org/10.1016/j.virusres.2007.10.009 PMID: 18045721

8. Totura AL, Baric RS. SARS coronavirus pathogenesis: host innate immune responses and viral antagonism of interferon. Curr Opin Virol. 2012; 2(3):264–75. https://doi.org/10.1016/j.coviro.2012.04.004 PMID: 22572391

9. Kasuga Y, Zhu B, Jang KJ, Yoo JS. Innate immune sensing of coronavirus and viral evasion strategies. Exp Mol Med. 2021; 53(5):723–36. https://doi.org/10.1038/s12276-021-00602-1 PMID: 33953325

10. Lokugamage KG, Narayanan K, Huang C, Makino S. Severe acute respiratory syndrome coronavirus protein nsp1 is a novel eukaryotic translation inhibitor that represses multiple steps of translation initiation. J Virol. 2012; 86(24):13598–608. https://doi.org/10.1128/JVI.01958-12 PMID: 23035226

11. Kamitani W, Huang C, Narayanan K, Lokugamage KG, Makino S. A two-pronged strategy to suppress host protein synthesis by SARS coronavirus Nsp1 protein. Nat Struct Mol Biol. 2009; 16(11):1134–40. https://doi.org/10.1038/nsmb.1680 PMID: 19838190

12. Frieman M, Yount B, Heise M, Kopecky-Bromberg SA, Palese P, Baric RS. Severe acute respiratory syndrome coronavirus ORF6 antagonizes STAT1 function by sequestering nuclear import factors on the rough endoplasmic reticulum/Golgi membrane. J Virol. 2007; 81(18):9812–24. https://doi.org/10.1128/JVI.01012-07 PMID: 17596301

13. Wathelet MG, Orr M, Frieman MB, Baric RS. Severe acute respiratory syndrome coronavirus evades antiviral signaling: role of nsp1 and rational design of an attenuated strain. J Virol. 2007; 81 (21):11620–33. https://doi.org/10.1128/JVI.00702-07 PMID: 17715225

14. Daffis S, Szretter KJ, Schriewer J, Li J, Youn S, Errett J, et al. 2'-O methylation of the viral mRNA cap evades host restriction by IFIT family members. Nature. 2010; 468(7322):452–6. https://doi.org/10.1038/nature09489 PMID: 21085181

15. Cruz JL, Sola I, Becares M, Alberca B, Plana J, Enjuanes L, et al. Coronavirus gene 7 counteracts host defenses and modulates virus virulence. PLoS Pathog. 2011; 7(6):e1002090. https://doi.org/10.1371/journal.ppat.1002090 PMID: 21695242

16. Kamitani W, Narayanan K, Huang C, Lokugamage K, Ikegami T, Ito N, et al. Severe acute respiratory syndrome coronavirus nsp1 protein suppresses host gene expression by promoting host mRNA degradation. Proc Natl Acad Sci U S A. 2006; 103(34):12885–90. https://doi.org/10.1073/pnas.0603144103 PMID: 16912115

17. Gordon DE, Jang GM, Bouhaddou M, Xu J, Obernier K, White KM, et al. A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. Nature. 2020; 583(7816):459–68. https://doi.org/10.1038/s41586-020-2286-9 PMID: 32353859

18. Choi H, Larsen B, Lin ZY, Breitkreutz A, Mellacheruvu D, Fermin D, et al. SAINT: probabilistic scoring of affinity purification-mass spectrometry data. Nat Methods. 2011; 8(1):70–3. https://doi.org/10.1038/nmeth.1541 PMID: 21131968

19. Gordon DE, Hiatt J, Bouhaddou M, Rezelj VV, Ulferts S, Braberg H, et al. Comparative host-coronavirus protein interaction networks reveal pan-viral disease mechanisms. Science. 2020. https://doi.org/10.1126/science.abe9403 PMID: 33060197

20. Rihn SJ, Merits A, Bakshi S, Turnbull ML, Wickenhagen A, Alexander AJT, et al. A plasmid DNA-launched SARS-CoV-2 reverse genetics system and coronavirus toolkit for COVID-19 research. PLoS Biol. 2021; 19(2):e3001091. https://doi.org/10.1371/journal.pbio.3001091 PMID: 33630831

21. Yao Z, Darowski K, St-Denis N, Wong V, Offensperger F, Villedieu A, et al. A Global Analysis of the Receptor Tyrosine Kinase-Protein Phosphatase Interactome. Mol Cell. 2017; 65(2):347–60. https://doi.org/10.1016/j.molcel.2016.12.004 PMID: 28065597

22. Yadav L, Tamene F, Goos H, van Drogen A, Katainen R, Aebersold R, et al. Systematic Analysis of Human Protein Phosphatase Interactions and Dynamics. Cell Syst. 2017; 4(4):430–44 e5. https://doi.org/10.1016/j.cels.2017.02.011 PMID: 28330616

23. Lee SY, Kang MG, Park JS, Lee G, Ting AY, Rhee HW. APEX Fingerprinting Reveals the Subcellular Localization of Proteins of Interest. Cell Rep. 2016; 15(8):1837–47. https://doi.org/10.1016/j.celrep.2016.04.064 PMID: 27184847

24. Go CD, Knight JDR, Rajasekharan A, Rathod B, Hesketh GG, Abe KT, et al. A proximity-dependent biotinylation map of a human cell. Nature. 2021; 595(7865):120–4. https://doi.org/10.1038/s41586-021-03592-2 PMID: 34079125

25. Khan A, Tahir Khan M, Saleem S, Junaid M, Ali A, Shujait Ali S, et al. Structural insights into the mechanism of RNA recognition by the N-terminal RNA-binding domain of the SARS-CoV-2 nucleocapsid phosphoprotein. Comput Struct Biotechnol J. 2020; 18:2174–84. https://doi.org/10.1016/j.csbj.2020.08.006 PMID: 32837710

26. Dinesh DC, Chalupska D, Silhan J, Koutna E, Nencka R, Veverka V, et al. Structural basis of RNA recognition by the SARS-CoV-2 nucleocapsid phosphoprotein. PLoS Pathog. 2020; 16(12):e1009100. https://doi.org/10.1371/journal.ppat.1009100 PMID: 33264373

27. Severe Covid GG, Ellinghaus D, Degenhardt F, Bujanda L, Buti M, Albillos A, et al. Genomewide Association Study of Severe Covid-19 with Respiratory Failure. N Engl J Med. 2020; 383(16):1522–34. https://doi.org/10.1056/NEJMoa2020283 PMID: 32558485

28. Initiative C-HG. The COVID-19 Host Genetics Initiative, a global initiative to elucidate the role of host genetic factors in susceptibility and severity of the SARS-CoV-2 virus pandemic. Eur J Hum Genet. 2020; 28(6):715–8. https://doi.org/10.1038/s41431-020-0636-6 PMID: 32404885

29. Võsa U, Claringbould A, Westra H-J, Bonder MJ, Deelen P, Zeng B, et al. Unraveling the polygenic architecture of complex traits using blood eQTL metaanalysis. bioRxiv. 2018: 447367.

30. Oshiumi H, Matsumoto M, Hatakeyama S, Seya T. Riplet/RNF135, a RING finger protein, ubiquitinates RIG-I to promote interferon-beta induction during the early phase of viral infection. J Biol Chem. 2009; 284(2):807–17. https://doi.org/10.1074/jbc.M804259200 PMID: 19017631

31. Gack MU, Shin YC, Joo CH, Urano T, Liang C, Sun L, et al. TRIM25 RING-finger E3 ubiquitin ligase is essential for RIG-I-mediated antiviral activity. Nature. 2007; 446(7138):916–20. https://doi.org/10.1038/nature05732 PMID: 17392790

32. Kuniyoshi K, Takeuchi O, Pandey S, Satoh T, Iwasaki H, Akira S, et al. Pivotal role of RNA-binding E3 ubiquitin ligase MEX3C in RIG-I-mediated antiviral innate immunity. Proc Natl Acad Sci U S A. 2014; 111(15):5646–51. https://doi.org/10.1073/pnas.1401674111 PMID: 24706898

33. Yan J, Li Q, Mao AP, Hu MM, Shu HB. TRIM4 modulates type I interferon induction and cellular antiviral response by targeting RIG-I for K63-linked ubiquitination. J Mol Cell Biol. 2014; 6(2):154–63. https://doi.org/10.1093/jmcb/mju005 PMID: 24755855

34. Wang R, Simoneau CR, Kulsuptrakul J, Bouhaddou M, Travisano KA, Hayashi JM, et al. Genetic Screens Identify Host Factors for SARS-CoV-2 and Common Cold Coronaviruses. Cell. 2021; 184 (1):106–19 e14. https://doi.org/10.1016/j.cell.2020.12.004 PMID: 33333024

35. Daniloski Z, Jordan TX, Wessels HH, Hoagland DA, Kasela S, Legut M, et al. Identification of Required Host Factors for SARS-CoV-2 Infection in Human Cells. Cell. 2021; 184(1):92–105 e16. https://doi.org/10.1016/j.cell.2020.10.030 PMID: 33147445

36. Schneider WM, Luna JM, Hoffmann HH, Sanchez-Rivera FJ, Leal AA, Ashbrook AW, et al. Genome-Scale Identification of SARS-CoV-2 and Pan-coronavirus Host Factor Networks. Cell. 2021; 184 (1):120–32 e14. https://doi.org/10.1016/j.cell.2020.12.006 PMID: 33382968

37. Wickramasinghe VO, Stewart M, Laskey RA. GANP enhances the efficiency of mRNA nuclear export in mammalian cells. Nucleus. 2010; 1(5):393–6. https://doi.org/10.4161/nucl.1.5.12351 PMID: 21326821

38. Wickramasinghe VO, Andrews R, Ellis P, Langford C, Gurdon JB, Stewart M, et al. Selective nuclear export of specific classes of mRNA from mammalian nuclei is promoted by GANP. Nucleic Acids Res. 2014; 42(8):5059–71. https://doi.org/10.1093/nar/gku095 PMID: 24510098

39. Schubert K, Karousis ED, Jomaa A, Scaiola A, Echeverria B, Gurzeler LA, et al. SARS-CoV-2 Nsp1 binds the ribosomal mRNA channel to inhibit translation. Nat Struct Mol Biol. 2020; 27(10):959–66. https://doi.org/10.1038/s41594-020-0511-8 PMID: 32908316

40. Thoms M, Buschauer R, Ameismeier M, Koepke L, Denk T, Hirschenberger M, et al. Structural basis for translational shutdown and immune evasion by the Nsp1 protein of SARS-CoV-2. Science. 2020; 369(6508):1249–55. https://doi.org/10.1126/science.abc8665 PMID: 32680882

41. Beachboard DC, Horner SM. Innate immune evasion strategies of DNA and RNA viruses. Curr Opin Microbiol. 2016; 32:113–9. https://doi.org/10.1016/j.mib.2016.05.015 PMID: 27288760

42. Kawai T, Takahashi K, Sato S, Coban C, Kumar H, Kato H, et al. IPS-1, an adaptor triggering RIG-I- and Mda5-mediated type I interferon induction. Nat Immunol. 2005; 6(10):981–8. https://doi.org/10.1038/ni1243 PMID: 16127453

43. Meylan E, Curran J, Hofmann K, Moradpour D, Binder M, Bartenschlager R, et al. Cardif is an adaptor protein in the RIG-I antiviral pathway and is targeted by hepatitis C virus. Nature. 2005; 437 (7062):1167–72. https://doi.org/10.1038/nature04193 PMID: 16177806

44. Seth RB, Sun L, Ea CK, Chen ZJ. Identification and characterization of MAVS, a mitochondrial antiviral signaling protein that activates NF-kappaB and IRF 3. Cell. 2005; 122(5):669–82. https://doi.org/10.1016/j.cell.2005.08.012 PMID: 16125763

45. Xu LG, Wang YY, Han KJ, Li LY, Zhai Z, Shu HB. VISA is an adapter protein required for virus-triggered IFN-beta signaling. Mol Cell. 2005; 19(6):727–40. https://doi.org/10.1016/j.molcel.2005.08.014 PMID: 16153868

46. Yoneyama M, Fujita T. RNA recognition and signal transduction by RIG-I-like receptors. Immunol Rev. 2009; 227(1):54–65. https://doi.org/10.1111/j.1600-065X.2008.00727.x PMID: 19120475

47. Nan Y, Nan G, Zhang YJ. Interferon induction by RNA viruses and antagonism by viral pathogens. Viruses. 2014; 6(12):4999–5027. https://doi.org/10.3390/v6124999 PMID: 25514371

48. Lei X, Dong X, Ma R, Wang W, Xiao X, Tian Z, et al. Activation and evasion of type I interferon responses by SARS-CoV-2. Nat Commun. 2020; 11(1):3810. https://doi.org/10.1038/s41467-020-17665-9 PMID: 32733001

49. Miorin L, Kehrer T, Sanchez-Aparicio MT, Zhang K, Cohen P, Patel RS, et al. SARS-CoV-2 Orf6 hijacks Nup98 to block STAT nuclear import and antagonize interferon signaling. Proc Natl Acad Sci U S A. 2020; 117(45):28344–54. https://doi.org/10.1073/pnas.2016650117 PMID: 33097660

50. Sims AC, Tilton SC, Menachery VD, Gralinski LE, Schafer A, Matzke MM, et al. Release of severe acute respiratory syndrome coronavirus nuclear import block enhances host transcription in human lung cells. J Virol. 2013; 87(7):3885–902. https://doi.org/10.1128/JVI.02520-12 PMID: 23365422

51. Gorbalenya AE, Donchenko AP, Blinov VM, Koonin EV. Cysteine proteases of positive strand RNA viruses and chymotrypsin-like serine proteases. A distinct protein superfamily with a common structural fold. FEBS Lett. 1989; 243(2):103–14. https://doi.org/10.1016/0014-5793(89)80109-7 PMID: 2645167

52. He J, Hu L, Huang X, Wang C, Zhang Z, Wang Y, et al. Potential of coronavirus 3C-like protease inhibitors for the development of new anti-SARS-CoV-2 drugs: Insights from structures of protease and inhibitors. Int J Antimicrob Agents. 2020; 56(2):106055. https://doi.org/10.1016/j.ijantimicag.2020.106055 PMID: 32534187

53. Rathnayake AD, Zheng J, Kim Y, Perera KD, Mackin S, Meyerholz DK, et al. 3C-like protease inhibitors block coronavirus replication in vitro and improve survival in MERS-CoV-infected mice. Sci Transl Med. 2020; 12(557). https://doi.org/10.1126/scitranslmed.abc5332 PMID: 32747425

54. Kuyumcu-Martinez NM, Van Eden ME, Younan P, Lloyd RE. Cleavage of poly(A)-binding protein by poliovirus 3C protease inhibits host cell translation: a novel mechanism for host translation shutoff. Mol Cell Biol. 2004; 24(4):1779–90. https://doi.org/10.1128/MCB.24.4.1779-1790.2004 PMID: 14749392

55. Barco A, Feduchi E, Carrasco L. Poliovirus protease 3C(pro) kills cells by apoptosis. Virology. 2000; 266(2):352–60. https://doi.org/10.1006/viro.1999.0043 PMID: 10639321

56. Shubin AV, Demidyuk IV, Lunina NA, Komissarov AA, Roschina MP, Leonova OG, et al. Protease 3C of hepatitis A virus induces vacuolization of lysosomal/endosomal organelles and caspase-independent cell death. BMC Cell Biol. 2015; 16:4. https://doi.org/10.1186/s12860-015-0050-z PMID: 25886889

57. Muramatsu T, Kim YT, Nishii W, Terada T, Shirouzu M, Yokoyama S. Autoprocessing mechanism of severe acute respiratory syndrome coronavirus 3C-like protease (SARS-CoV 3CLpro) from its polyproteins. FEBS J. 2013; 280(9):2002–13. https://doi.org/10.1111/febs.12222 PMID: 23452147

58. Yang H, Yang M, Ding Y, Liu Y, Lou Z, Zhou Z, et al. The crystal structures of severe acute respiratory syndrome virus main protease and its complex with an inhibitor. Proc Natl Acad Sci U S A. 2003; 100 (23):13190–5. https://doi.org/10.1073/pnas.1835675100 PMID: 14585926

59. Kiemer L, Lund O, Brunak S, Blom N. Coronavirus 3CLpro proteinase cleavage sites: possible relevance to SARS virus pathology. BMC Bioinformatics. 2004; 5:72. https://doi.org/10.1186/1471-2105-5-72 PMID: 15180906

60. Steckelberg AL, Boehm V, Gromadzka AM, Gehring NH. CWC22 connects pre-mRNA splicing and exon junction complex assembly. Cell Rep. 2012; 2(3):454–61. https://doi.org/10.1016/j.celrep.2012.08.017 PMID: 22959432

61. Zhou Z, Licklider LJ, Gygi SP, Reed R. Comprehensive proteomic analysis of the human spliceosome. Nature. 2002; 419(6903):182–5. https://doi.org/10.1038/nature01031 PMID: 12226669

62. Garcia-Higuera I, Taniguchi T, Ganesan S, Meyn MS, Timmers C, Hejna J, et al. Interaction of the Fanconi anemia proteins and BRCA1 in a common pathway. Mol Cell. 2001; 7(2):249–62. https://doi.org/10.1016/s1097-2765(01)00173-3 PMID: 11239454

63. Pedersen NC, Kim Y, Liu H, Galasiti Kankanamalage AC, Eckstrand C, Groutas WC, et al. Efficacy of a 3C-like protease inhibitor in treating various forms of acquired feline infectious peritonitis. J Feline Med Surg. 2018; 20(4):378–92. https://doi.org/10.1177/1098612X17729626 PMID: 28901812

64. Yang Y, Liu C, Du L, Jiang S, Shi Z, Baric RS, et al. Two Mutations Were Critical for Bat-to-Human Transmission of Middle East Respiratory Syndrome Coronavirus. J Virol. 2015; 89(17):9119–23. https://doi.org/10.1128/JVI.01279-15 PMID: 26063432

65. Sheahan T, Rockx B, Donaldson E, Sims A, Pickles R, Corti D, et al. Mechanisms of zoonotic severe acute respiratory syndrome coronavirus host range expansion in human airway epithelium. J Virol. 2008; 82(5):2274–85. https://doi.org/10.1128/JVI.02041-07 PMID: 18094188

66. Menachery VD, Yount BL Jr, Debbink K, Agnihothram S, Gralinski LE, Plante JA, et al. A SARS-like cluster of circulating bat coronaviruses shows potential for human emergence. Nat Med. 2015; 21 (12):1508–13. https://doi.org/10.1038/nm.3985 PMID: 26552008

67. Wan Y, Shang J, Graham R, Baric RS, Li F. Receptor Recognition by the Novel Coronavirus from Wuhan: an Analysis Based on Decade-Long Structural Studies of SARS Coronavirus. J Virol. 2020; 94(7). https://doi.org/10.1128/JVI.00127-20 PMID: 31996437

68. Shi Z, Hu Z. A review of studies on animal reservoirs of the SARS coronavirus. Virus Res. 2008; 133 (1):74–87. https://doi.org/10.1016/j.virusres.2007.03.012 PMID: 17451830

69. Gussow AB, Auslander N, Faure G, Wolf YI, Zhang F, Koonin EV. Genomic determinants of pathogenicity in SARS-CoV-2 and other human coronaviruses. Proc Natl Acad Sci U S A. 2020; 117 (26):15193–9. https://doi.org/10.1073/pnas.2008176117 PMID: 32522874

70. Hofmann H, Pyrc K, van der Hoek L, Geier M, Berkhout B, Pohlmann S. Human coronavirus NL63 employs the severe acute respiratory syndrome coronavirus receptor for cellular entry. Proc Natl Acad Sci U S A. 2005; 102(22):7988–93. https://doi.org/10.1073/pnas.0409465102 PMID: 15897467

71. de Haan CA, Rottier PJ. Molecular interactions in the assembly of coronaviruses. Adv Virus Res. 2005; 64:165–230. https://doi.org/10.1016/S0065-3527(05)64006-7 PMID: 16139595

72. van Hemert MJ, van den Worm SH, Knoops K, Mommaas AM, Gorbalenya AE, Snijder EJ. SARS-coronavirus replication/transcription complexes are membrane-protected and need a host factor for activity in vitro. PLoS Pathog. 2008; 4(5):e1000054. https://doi.org/10.1371/journal.ppat.1000054 PMID: 18451981

73. Minakshi R, Padhan K, Rani M, Khan N, Ahmad F, Jameel S. The SARS Coronavirus 3a protein causes endoplasmic reticulum stress and induces ligand-independent downregulation of the type 1 interferon receptor. PLoS One. 2009; 4(12):e8342. https://doi.org/10.1371/journal.pone.0008342 PMID: 20020050

74. Feigenblum D, Schneider RJ. Modification of eukaryotic initiation factor 4F during infection by influenza virus. J Virol. 1993; 67(6):3027–35. https://doi.org/10.1128/JVI.67.6.3027-3035.1993 PMID: 8098776

75. Sommergruber W, Ahorn H, Klump H, Seipelt J, Zoephel A, Fessl F, et al. 2A proteinases of coxsackie- and rhinovirus cleave peptides derived from eIF-4 gamma via a common recognition motif. Virology. 1994; 198(2):741–5. https://doi.org/10.1006/viro.1994.1089 PMID: 8291255

76. de Breyne S, Bonderoff JM, Chumakov KM, Lloyd RE, Hellen CU. Cleavage of eukaryotic initiation factor eIF5B by enterovirus 3C proteases. Virology. 2008; 378(1):118–22. https://doi.org/10.1016/j.virol.2008.05.019 PMID: 18572216

77. Connor JH, Lyles DS. Vesicular stomatitis virus infection alters the eIF4F translation initiation complex and causes dephosphorylation of the eIF4E binding protein 4E-BP1. J Virol. 2002; 76(20):10177–87. https://doi.org/10.1128/jvi.76.20.10177-10187.2002 PMID: 12239292

78. Ventoso I, Sanz MA, Molina S, Berlanga JJ, Carrasco L, Esteban M. Translational resistance of late alphavirus mRNA to eIF2alpha phosphorylation: a strategy to overcome the antiviral effect of protein kinase PKR. Genes Dev. 2006; 20(1):87–100. https://doi.org/10.1101/gad.357006 PMID: 16391235

79. Kieft JS. Viral IRES RNA structures and ribosome interactions. Trends Biochem Sci. 2008; 33(6):274–83. https://doi.org/10.1016/j.tibs.2008.04.007 PMID: 18468443

80. Nie X, Qian L, Sun R, Huang B, Dong X, Xiao Q, et al. Multi-organ proteomic landscape of COVID-19 autopsies. Cell. 2021; 184(3):775–91 e14. https://doi.org/10.1016/j.cell.2021.01.004 PMID: 33503446

81. Kim D, Lee JY, Yang JS, Kim JW, Kim VN, Chang H. The Architecture of SARS-CoV-2 Transcriptome. Cell. 2020; 181(4):914–21 e10. https://doi.org/10.1016/j.cell.2020.04.011 PMID: 32330414

82. Neidermyer WJ Jr, Whelan SPJ. Global analysis of polysome-associated mRNA in vesicular stomatitis virus infected cells. PLoS Pathog. 2019; 15(6):e1007875. https://doi.org/10.1371/journal.ppat.1007875 PMID: 31226162

83. Pritchard CE, Fornerod M, Kasper LH, van Deursen JM. RAE1 is a shuttling mRNA export factor that binds to a GLEBS-like NUP98 motif at the nuclear pore complex through multiple domains. J Cell Biol. 1999; 145(2):237–54. https://doi.org/10.1083/jcb.145.2.237 PMID: 10209021

84. Ren Y, Seo HS, Blobel G, Hoelz A. Structural and functional analysis of the interaction between the nucleoporin Nup98 and the mRNA export factor Rae1. Proc Natl Acad Sci U S A. 2010; 107 (23):10406–11. https://doi.org/10.1073/pnas.1005389107 PMID: 20498086

85. Anderson J, Schiffer C, Lee SK, Swanstrom R. Viral protease inhibitors. Handb Exp Pharmacol. 2009 (189):85–110. https://doi.org/10.1007/978-3-540-79086-0_4 PMID: 19048198

86. Arribas JR, Girard PM, Paton N, Winston A, Marcelin AG, Elbirt D, et al. Efficacy of protease inhibitor monotherapy vs. triple therapy: meta-analysis of data from 2303 patients in 13 randomized trials. HIV Med. 2016; 17(5):358–67. https://doi.org/10.1111/hiv.12348 PMID: 26709605

87. Kim Y, Lovell S, Tiew KC, Mandadapu SR, Alliston KR, Battaile KP, et al. Broad-spectrum antivirals against 3C or 3C-like proteases of picornaviruses, noroviruses, and coronaviruses. J Virol. 2012; 86 (21):11754–62. https://doi.org/10.1128/JVI.01348-12 PMID: 22915796

88. Jagdeo JM, Dufour A, Klein T, Solis N, Kleifeld O, Kizhakkedathu J, et al. N-Terminomics TAILS Identifies Host Cell Substrates of Poliovirus and Coxsackievirus B3 3C Proteinases That Modulate Virus Infection. J Virol. 2018; 92(8). https://doi.org/10.1128/JVI.02211-17 PMID: 29437971

89. Vuong W, Khan MB, Fischer C, Arutyunova E, Lamer T, Shields J, et al. Feline coronavirus drug inhibits the main protease of SARS-CoV-2 and blocks virus replication. Nat Commun. 2020; 11(1):4282. https://doi.org/10.1038/s41467-020-18096-2 PMID: 32855413

90. Ong S-E, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, et al. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Mol Cell Proteomics. 2002; 1:376–86. https://doi.org/10.1074/mcp.m200025-mcp200 PMID: 12118079

91. Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol. 2008; 26:1367–72. https://doi.org/10.1038/nbt.1511 PMID: 19029910

92. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS. 2012; 16(5):284–7. https://doi.org/10.1089/omi.2011.0118 PMID: 22455463

93. Wickham H. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York. 2016(ISBN 978-3-319-24277-4).

94. Kolde R. pheatmap: Pretty heatmaps [Software]. URL https://CRAN.R-project.org/package=pheatmap. 2015.

95. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003; 13(11):2498–504. https://doi.org/10.1101/gr.1239303 PMID: 14597658

96. Holden P, Horton WA. Crude subcellular fractionation of cultured mammalian cell lines. BMC Res Notes. 2009; 2:243. https://doi.org/10.1186/1756-0500-2-243 PMID: 20003239

97. Myers TA, Chanock SJ, Machiela MJ. LDlinkR: An R Package for Rapidly Calculating Linkage Disequilibrium Statistics in Diverse Populations. Front Genet. 2020; 11:157. https://doi.org/10.3389/fgene.2020.00157 PMID: 32180801

98. Ward LD, Kellis M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. Nucleic Acids Res. 2012; 40(Database issue):D930–4. https://doi.org/10.1093/nar/gkr917 PMID: 22064851

99. Consortium GT. The GTEx Consortium atlas of genetic regulatory effects across human tissues. Science. 2020; 369(6509):1318–30. https://doi.org/10.1126/science.aaz1776 PMID: 32913098

100. Schmiedel BJ, Singh D, Madrigal A, Valdovino-Gonzalez AG, White BM, Zapardiel-Gonzalo J, et al. Impact of Genetic Polymorphisms on Human Immune Cell Gene Expression. Cell. 2018; 175 (6):1701–15 e16. https://doi.org/10.1016/j.cell.2018.10.022 PMID: 30449622

101. Mumbach MR, Satpathy AT, Boyle EA, Dai C, Gowen BG, Cho SW, et al. Enhancer connectome in primary human cells identifies target genes of disease-associated DNA elements. Nat Genet. 2017; 49 (11):1602–12. https://doi.org/10.1038/ng.3963 PMID: 28945252

102. Bhattacharyya S, Chandra V, Vijayanand P, Ay F. Identification of significant chromatin contacts from HiChIP data by FitHiChIP. Nat Commun. 2019; 10(1):4221. https://doi.org/10.1038/s41467-019-11950-y PMID: 31530818

103. Gray DC, Mahrus S, Wells JA. Activation of specific apoptotic caspases with an engineered small-molecule-activated protease. Cell. 2010; 142(4):637–46. https://doi.org/10.1016/j.cell.2010.07.014 PMID: 20723762

104. Emmott E, Sweeney TR, Goodfellow I. A Cell-based Fluorescence Resonance Energy Transfer (FRET) Sensor Reveals Inter- and Intragenogroup Variations in Norovirus Protease Activity and Polyprotein Cleavage. J Biol Chem. 2015; 290(46):27841–53. https://doi.org/10.1074/jbc.M115.688234 PMID: 26363064