

A BioID-derived proximity interactome for SARS-CoV-2 proteins

Danielle G. May¹, Laura Martin-Sancho², Valesca Anschau², Sophie Liu³, Rachel J. Chrisopoulos¹, Kelsey L. Scott¹, Charles T. Halfmann¹, Ramon Díaz Peña², Dexter Pratt³, Alexandre R. Campos^{2*}, and Kyle J. Roux^{1,4*}

¹Enabling Technologies Group, Sanford Research, Sioux Falls, SD; ²Sanford Burnham Prebys Medical Discovery Institute, La Jolla, CA; ³Division of Genetics, Department of Medicine, University of California, San Diego, CA; ⁴Department of Pediatrics, Sanford School of Medicine, University of South Dakota, Sioux Falls, SD

*Co-corresponding authors

Abstract

The novel coronavirus SARS-CoV-2 is responsible for the ongoing COVID-19 pandemic and has caused a major health and economic burden worldwide. Understanding how SARS-CoV-2 viral proteins behave in host cells can reveal underlying mechanisms of pathogenesis and assist in development of antiviral therapies. Here we use BioID to map the SARS-CoV-2 virus-host interactome using human lung cancer derived A549 cells expressing individual SARS-CoV-2 viral proteins. Functional enrichment analyses revealed previously reported and unreported cellular pathways that are in association with SARS-CoV-2 proteins. We have also established a website to host the proteomic data to allow for public access and continued analysis of host-viral protein associations and whole-cell proteomes of cells expressing the viral-BioID fusion proteins. Collectively, these studies provide a valuable resource to potentially uncover novel SARS-CoV-2 biology and inform development of antivirals.

Introduction

The 2019 novel coronavirus, SARS-CoV-2, is the causative agent of Coronavirus Disease 2019 (COVID-19) responsible for a global pandemic. COVID-19 most often presents as a respiratory illness, yet also can cause gastrointestinal and/or neurological symptoms, and acute cardiac injury [1-3]. Presently, hundreds of millions of people have been infected with SARS-CoV-2 worldwide, and several million people have died as a result. Long-term effects of COVID-19 infection are reported by 10-30% of patients, and as millions of people recover from COVID-19, questions remain about vertical transmission of COVID-19 infection during pregnancy, and post-COVID syndrome symptoms including pulmonary fibrosis, neurological defects, and vascular dysfunction [4-10]. While wide-spread vaccination is likely to slow the spread of COVID-19, developing treatment strategies for new infections and long-term post-COVID symptoms will require a thorough understanding of the SARS-CoV-2 virus and how it affects patient cell biology.

A crucial component of the effort to study COVID-19 is the application of technologies that reveal how viral proteins behave in host cells. Current efforts to map the SARS-CoV-2 virus-host interactome have offered great insight into possible pathways directly affected by various viral proteins, yet differences in experimental approaches and data analysis methods inevitably lead to discrepancies when comparing reported interactomes [11-19]. As with any large-scale approach to identifying gene- or protein-networks, false-positives due to background contamination can hinder accurate data interpretation; therefore, the use of several approaches with multiple replicates by multiple independent studies will be required to ultimately map the full SARS-CoV-2 interactome.

Proximity-dependent labeling of host proteins via BioID or similar promiscuous biotin ligases fused to viral proteins has been used to study host-viral protein associations for a

number of viruses including herpes simplex virus type 1, Epstein-Barr virus, Zika virus, Ebola virus, and coronaviruses [20-24]. Here, we generated A549 human lung cancer cells stably expressing BioID-tagged SARS-CoV-2 viral proteins to identify whole-cell proteomic changes due to viral protein expression, and to identify specific protein-protein interactions (PPIs) between individual SARS-CoV-2 viral proteins and host-cell proteins. We also compared our BioID datasets with similar available proximity-based proteome datasets to develop a list of high-confidence candidate protein interactors. All of our data is available at (<https://alexproteomics.shinyapps.io/covid19proteomics/>) for further in-depth analysis by the scientific community and will serve to further our collective understanding of infection mechanisms and guide higher-confidence follow-up studies investigating specific PPIs and pathway alterations.

Results

Development of stable BioID cell lines expressing individual SARS-CoV-2 viral proteins

The SARS-CoV-2 virus generates two long polypeptides that are cleaved into 16 non-structural proteins (NSPs) as well as several downstream ORFs encoding 4 structural proteins (Spike, Envelope, Membrane, and Nucleocapsid – S, E, M, and N) and 9 accessory proteins. In order to identify global cellular changes associated with viral protein expression, as well as identify specific viral-host PPIs, the promiscuous biotin ligase BioID2 was fused to either the N- or C- terminus of individual SARS-CoV-2 proteins, and stably expressed by retroviral transduction in human lung cancer A549 cells (Table S1, Figure 1A). For each construct, we included a GGGS linker to alleviate steric hindrance between the BioID ligase and viral protein. Each cell line was validated by immunofluorescence (IF) and western blot (WB) for fusion-protein expression and biotinylation, revealing a wide-range of permissible expression levels and overall biotinylation (Figure 1B, Figure S1, S2). Three proteins (Spike, Nsp1, and ORF3b) were excluded from this study due to an inability to generate cells stably expressing BioID2-fusion proteins, leaving us with 26 viral-BioID2 fusion proteins. ORF8, a predicted luminal protein, was tagged with the TurboID ligase that was previously shown to be substantially more active in the ER lumen compared to BioID [25]. BioID2-alone was used as a control for these viral protein fusions, with the exception of ORF8, for which we utilized a signal sequence-TurboID-KDEL (TurboID-KDEL) to target and retain the ligase in the ER-lumen. Each cell line was processed in triplicate, and subjected to whole-cell lysis for global proteome analysis and affinity-purification of biotinylated proteins for identification of PPIs via mass spectrometry.

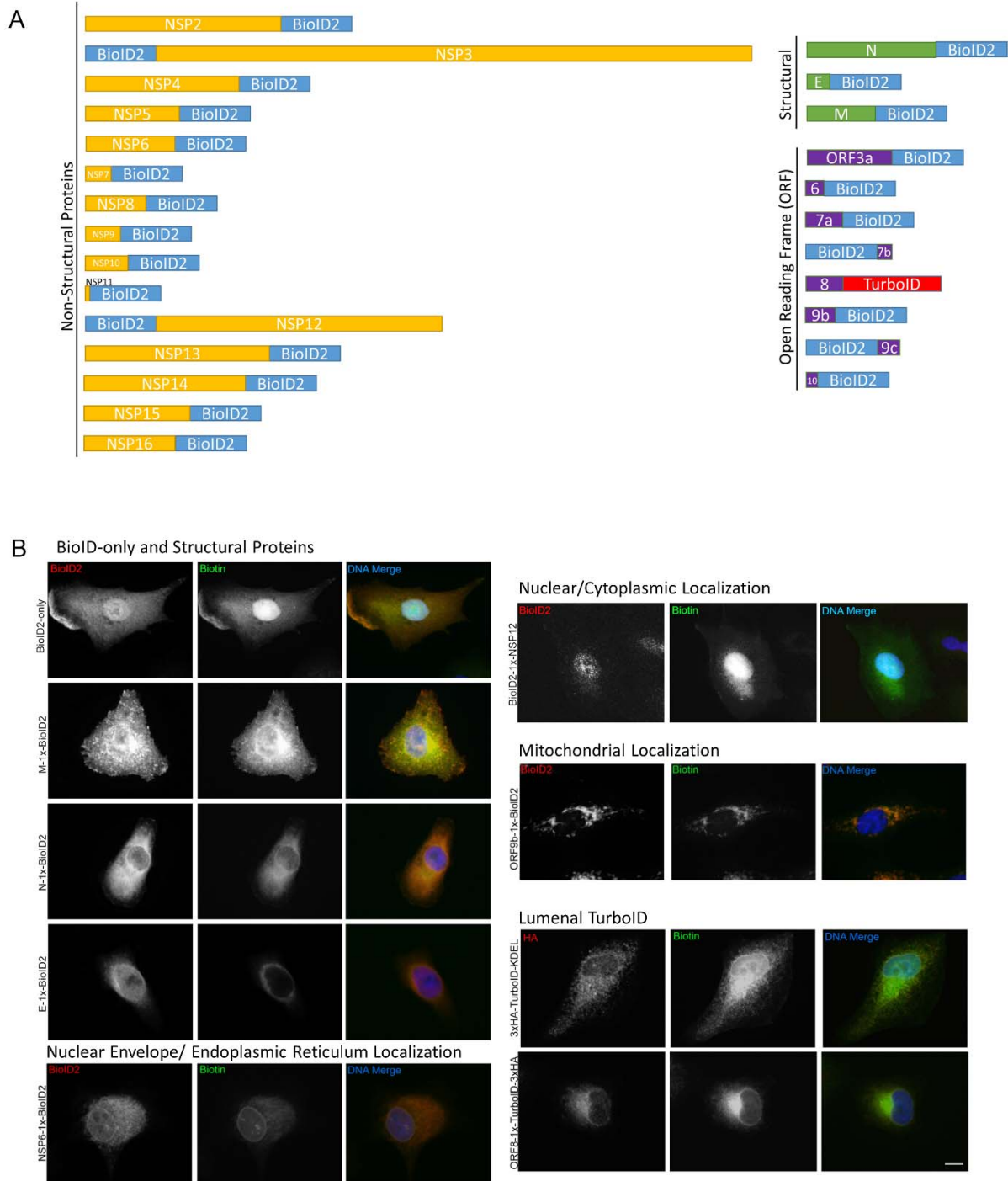


Figure 1. BioID2-viral fusion protein expression in A549 cells. A. Viral proteins were fused to either the N- or C-terminus of the BioID2 promiscuous biotin ligase. Schematic shows the orientation of NSPs (yellow), structural proteins (green), and ORF proteins (purple) fused to BioID2 to scale. B. A549 human lung cells stably expressing BioID2-fusion proteins were assessed for fusion-protein expression and localization (red) and promiscuous biotinylation (green) following the addition of exogenous biotin. Scale bar 10 μ m.

surprisingly, we saw a marked increase in proteins involved in cytokine signaling in immune system (HSA-1280215) in response to viral protein expression, including CD70, IRF9, and TNFSF9. The most significantly upregulated protein we identified was ITGB3 (logFC = +3.32 to +4.78), which has recently been shown to be upregulated in COVID-19 patient lung samples and has been hypothesized to be an alternative receptor for the SARS-CoV-2 virus [26, 27]. We found this ITGB3 upregulation in cells expressing ORF9c, ORF3a, ORF7b, E-protein, and NSP2. Interestingly, the most significantly downregulated proteins (logFC = -2.60 to -9.94) were MUC5AC and MUC5B, with dramatic reduction in cells expressing viral NSP12, NSP15, ORF7b, NSP3, NSP2, E, ORF9c, and ORF3a. Levels of these proteins were also significantly reduced in cells expressing NSP5, NSP6, NSP14, and N-protein, although to a lesser extent. MUC5AC/B are proteins involved in mucus secretion in the respiratory tract, and this data suggests that several of the SARS-CoV-2 proteins are capable of globally reducing cellular MUC5AC/B proteins, even when expressed individually. Proteins involved in DNA replication processes were also significantly suppressed, especially by NSP2 expression, including BRCA1 (NSP2, logFC = -2.58), PRIM2 (NSP2, logFC = -2.56), and CDCA2 (NSP2, logFC = -2.40), suggesting that NSP2 may play a key role in directly and/or indirectly disrupting cell cycle progression and apoptosis pathways.

Network analysis of SARS-CoV-2 host interactors reveals novel biology

In addition to whole-cell proteomic analysis, each stable cell line was subjected to BioID proximity-labeling to identify specific viral-host PPIs. Following statistical test, we identified 3,011 significant viral-host PPIs, with a log2FC \geq 2.3, p-value \leq 0.01, at least 2 quantitative peptide features and detected in less than 75% of the proximity-labeling CRAPome contaminant database experiments. This list of significant PPIs is available in our Covid-19 Proteomics website allowing users to interactively explore networks and functions of the detected PPIs. To understand the functional and biochemical relationships between the identified SARS-CoV-2 interactors, we conducted hierarchy and pathway enrichment analyses (*see Methods*) on a subset of 876 proteins uniquely associated with one of the 26 SARS-CoV-2 proteins (Table S3) with CrapomeScore \leq 0.5 (i.e., detected in 50% or less of the CRAPome proximity-labeling experiments, see Methods). These analyses revealed that most identified SARS-CoV-2 interactors were associated with seven clusters that included host translation machinery, endocytosis and vesicle transport, metabolism, glycosylation, cell junctions and ion transport, maintenance of homeostasis, and mitochondria function (Figure 3A). Subclusters within host translation were processing of mRNAs and non-sense mediated decay (NMD) ($p = 1.13^{-36}$), which is involved in degradation of aberrant self and non-self mRNAs including those of coronaviruses [28]. Consistent with previous systems-level studies of SARS-CoV-2, a significant number of the interactors were associated with endocytosis and vesicle trafficking pathways, including members of the SNARE complex, which are important for membrane fusion of vesicles and exocytosis [29], as well as GTPases that regulate vesicle docking and likely support SARS-CoV-2 trafficking and egress. Notably we found a highly enriched cluster of SARS-CoV-2 interactors involved in cholesterol biosynthesis ($p = 1.13^{-24}$) (Figure 3B), providing further evidence of the importance of this pathway for SARS-CoV-2 replication and highlighting potential targets for therapeutic efforts [30]. SARS-CoV-2 interactors were also associated with mitochondria function, including proteins of the TIM/TOM complex that mediate mitochondrial import ($p = 1.44^{-11}$), and proteins involved in electron transport chain ($p = 2.93^{-14}$) and oxidative phosphorylation ($p = 3.32^{-7}$), which could reflect SARS-CoV-2 energetic requirements for replication [31]. Notably, our analysis revealed several pathways involved in cell junctions and

ion transport. These included members of the SWELL complex (Figure 3C), which are involved in transport of cGAMP generated upon activation of the immune sensor cGAS by DNA viruses or mtDNA release [32], as well as proteins involved in cell adherens junctions ($p = 2.01^{-5}$), previously shown to be targeted by viruses to alter the environment of bystander cells and suggested as therapeutic targets to prevent viral spread [33]. In addition, amongst SARS-CoV-2 interactors were several members of the ABC-transporter family ($p = 7.58^{-12}$) (Figure 3D), involved in translocation of substrates across membranes, and previously linked to development of multidrug resistance (MDR) and oxidative stress response to viral and bacterial infection [34]. Proteins interacting with SARS-CoV-2 were also associated with the peroxisome ($p = 1.88^{-6}$). SARS-CoV-2 infection has been shown to recruit peroxisomes to viral replication organelles, and the association of SARS-CoV-2 with members of the peroxisome could reflect a requirement to reduce oxidative stress resulting from the extensive remodeling of cellular endomembranes or as a lipid source for viral replication [35, 36].

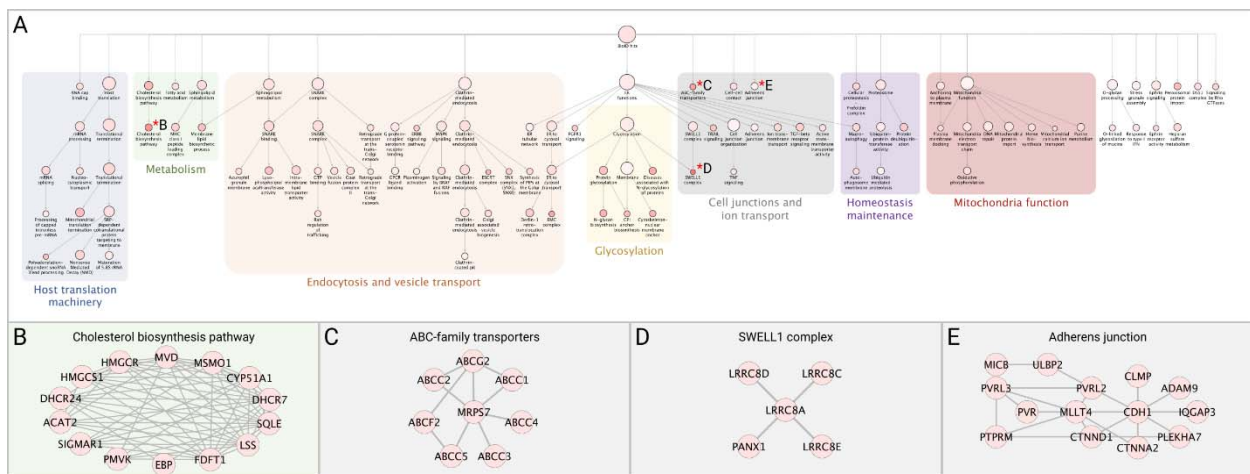


Figure 3. Network analysis of SARS-CoV-2 interactors. The network containing the 876 identified SARS-CoV-2 interactors was subjected to supervised community detection and the resultant hierarchy is shown. Each node represents a cluster of interconnected proteins and each edge (marked by an arrow) represents containment of one community (target) by another (source). Indicated are enriched biological processes as determined by gProfiler. (B-E) Asterisks (*) denote selected zoom-in insets from the hierarchy. Nodes represent human proteins, and edges are interactions from STRING.

Focused analysis of individual viral-host protein interactions

To identify the relationships between discrete SARS-CoV-2 proteins and cellular functions, we conducted pathway analyses on the cellular PPI candidates for each viral protein (see Methods). In the interest of brevity, we report here on the relationships identified for four SARS-CoV-2 viral proteins; however, all BioID data and several tools for pathway analysis have been made available on the SARS-CoV-2 proteomics website described above.

ORF3a

We identified 68 unique interactors for protein ORF3a, 37 of which were transmembrane proteins, including endosomal, lysosomal, and other vesicular proteins. SARS-CoV-2 utilizes deacidified lysosomes for egress and, consistent with this process, Orf3a interactors revealed enrichment in lysosomal transport proteins, regulators of endosome and lysosome fusion, and regulators of pH and ion homeostasis (Figure 4A) [37]. As previously reported, we identified HOPS endosomal tethering complex proteins VPS11 and VPS39, as well as WWP1, a HECT

ubiquitin ligase that has been previously associated with viral budding via the VPS pathway [11, 38], but we also identified ORF3a interactors to be involved in cell adhesion and adherens junctions, which could be exploited by SARS-CoV-2 to control cell-to-cell communication and promote cell spread. These results suggest that ORF3a plays a multifaceted role during viral infection, including a major role in membrane reorganization and trafficking, perhaps specifically utilizing the HECT/VPS viral budding pathway to enhance viral release.

ORF6

We identified 50 candidate interactors for ORF6, a membrane-associated protein reported to localize to the ER [39]. In line with previous studies, we identified SEC24A/B, proteins associated with COPII-coated vesicle transport, and other SEC complex proteins involved in ER homeostasis as associated with ORF6 (Figure 4B) [16, 17]. Other ORF6-associated proteins included cell cycle regulators of G1 to S phase transition GSPT1 and GSPT2, as well as PYCR1, PYCR2, and RRM2B, previously linked to cell cycle arrest at G1 phase [40]. Several RNA viruses manipulate critical cell cycle regulators or induce cell cycle arrest to favor viral replication, including inhibition of early apoptosis in infected cells, evasion of immune defenses, or to promote assembly of viral particles [41]. Additionally, several ORF6 interactors involved in deubiquitylation were identified, potentially suggesting a mechanism for deubiquitylation of viral proteins to evade degradation at the proteasome or by autophagy. Alternatively, ORF6 could influence deubiquitylation pathways to stabilize cellular factors that are supportive of viral replication, including USP5, which acts as a negative regulator of type I IFN signaling and has been found to increase in abundance during SARS-CoV-2 replication [42].

ORF8

The ORF8 protein has been implicated in modulating innate and adaptive immune response, specifically via downregulation of MHC-I [43, 44]. Furthermore, deletion of SARS-CoV-2 ORF8 lead to a decrease in proinflammatory cytokine release and increases efficacy of immune response in COVID-19 patients [45]. In line with this, we identified 64 unique PPIs, including 6 proteins involved in type I IFN signaling and 6 proteins involved in O-linked glycosylation (Figure 4C). This data supports previous work linking O-linked glycosylation process with ORF8 which could serve to evade the immune system using molecular mimicry and glycan shielding [11, 18, 46, 47]. Furthermore, the ORF8 dataset was enriched in several factors of the innate immune response, including OAS1, OASL, MX1, and PLSCR1, all of which are implicated in negative regulation of viral genome replication (GO:0045071), potentially supporting ORF8 as a key regulator of host immune response during SARS-CoV-2 infection. Additionally, we observed ORF8 associations with proteins implicated in MAPK signaling (KRAS, LGALS3, LGALS8 and ARRB2), and dephosphorylation process (MTMR1, MTMR2 and MTMR10; $p=4.31^{-4}$). These proteins play a role in intracellular membrane trafficking, and vesicle transport [48] and thus may serve to establish a mechanism for viral spread controlling cell signaling, replication and survival.

NSP4

Coronavirus NSP4 is part of the viral replication complex and rearranges host cell membranes to induce double-membrane vesicles for viral replication [49]. Our BioID analysis

identified 112 protein candidates uniquely associated with NSP4, including proteins involved in membrane lipid biosynthesis pathways, glycerophospholipid metabolism, and members of the N-glycan precursor biosynthesis machinery (Figure 4D). NSP4 was also associated with proteins involved in ubiquitination and proteasome degradation (CUL1, HERC2 and ANAPC2; $p=4.27^{-4}$), as well as members of the ER-associated protein degradation (ERAD) pathway (SEC61B, SEC62, ANAPCP2 and MARCH6; $p=8.10^{-8}$), suggesting a potential mechanism by which viral proteins can evade host-degradation machinery. Association of NSP4 with ERAD proteins could suggest antagonism of ERAD-mediated degradation of viral proteins by for instance autophagy, or an attempt to manipulate ERAD pathway to degrade immune regulators with antiviral properties to facilitate viral trafficking and release [50].

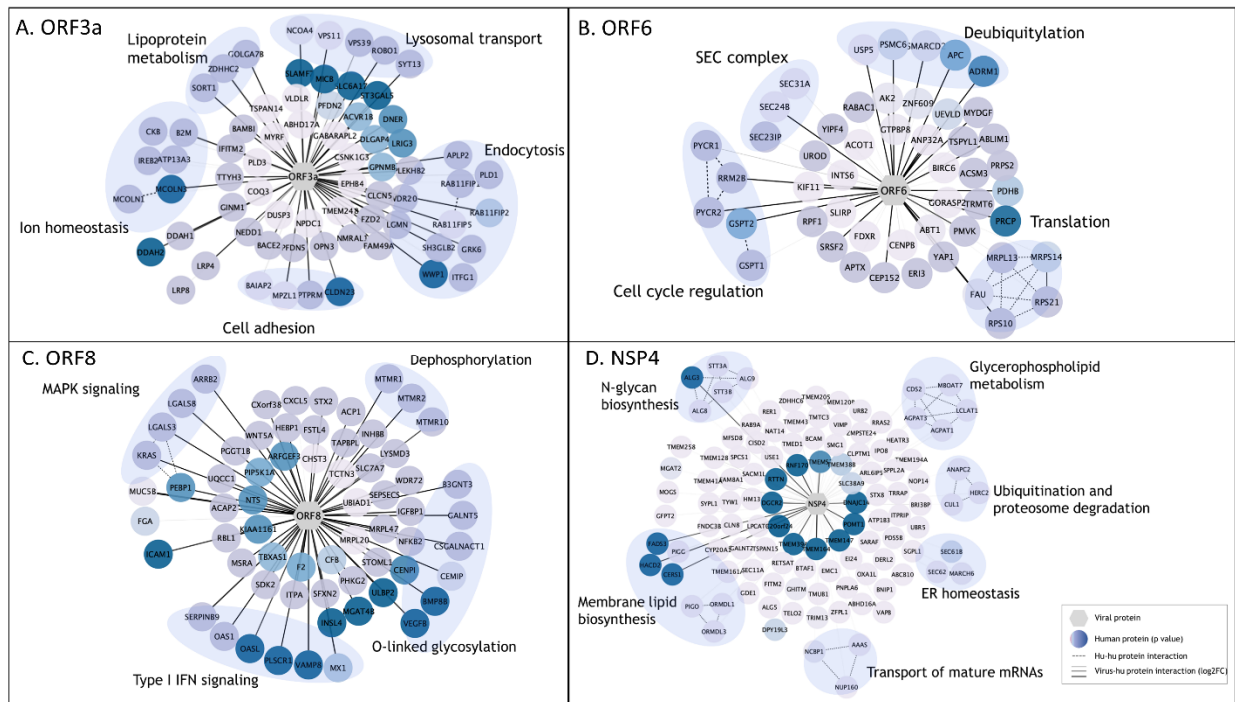


Figure 4. Enriched pathway analysis of PPIs for selected SARS-CoV-2 viral baits. High-confidence associations between indicated SARS-CoV-2 proteins (hexagons) and human proteins (circles/nodes). Node color is proportional to the p value (the darkest, the lowest the p value). Human-human interactions as determined by STRING are represented by dashed edges. Human-viral interactions are indicated with solid edges and their thickness are proportional to the \log_2FC (the thickest, the highest \log_2FC).

Integrated analysis with previously published datasets

SARS-CoV-2 interaction with the cellular restrictome

To further explore the interplay of the identified SARS-CoV-2 interactors and the innate immune response, we leveraged a recent gain-of-function screen that identified 65 interferon stimulated genes (ISG) that act to inhibit SARS-CoV-2 replication [51]. Cross-comparison between these two datasets revealed that 7 of these ISGs were found in association with one or more SARS-CoV-2 proteins, including ISG15, IFIT1, IFIT5, IFITM2, IFITM3, MLKL, and SPATS2L (Table S4). Our study revealed an association between viral N and SPAT2SL, an ISG that was found to inhibit SARS-CoV-2 RNA replication and is involved in formation of stress granules [51, 52]. SARS-CoV-2 has been suggested to antagonize stress granules to evade immune responses and these data suggest that N and SPAT2SL interaction could be important for this mechanism

[53, 54]. SARS-CoV-2 Orf9c has been recently associated with evasion of immune responses though the molecular regulators are yet to be defined [55]. Our study elucidated Or9c in association with IFIT5, an ISG that targets *non-self* RNA for degradation and was found to inhibit SARS-CoV-2 replication [51], thus suggesting that this factor could be targeted by Orf9c for immune evasion. Finally, Orf8b and Orf9b have been shown to trigger mechanisms of cell death [56]. Consistent with those findings, Orf9c was found to associate with the activator of necroptosis MLKL [57]. More work will be required to characterize these factors and investigate their role in SARS-CoV-2 pathogenesis.

Utilizing previous SARS-CoV-2 BioID interactome datasets to develop a list of high-confidence interactions

In order to develop a high-confidence list of viral-host PPIs, we compared our BioID dataset to two previously published proximity-labeling datasets [16, 17]. We were able to corroborate 154 unique viral-host PPIs across 13 SARS-CoV-2 viral proteins identified by at least one of these reports, with 25 total interactions identified by all three datasets (Table S5). Nine ORF9b PPIs were identified by all three reports, including the antiviral signaling protein MAVS and mitochondrial fusion/fission proteins MFF, MTFR1L, and USP30. We also further substantiated ORF6 involvement in ER to Golgi vesicle transport via manipulation of COPII complex (SEC31A, SEC23IP), with a possible role in mediating RABAC1 for vesicle release from the Golgi. Recently a fourth epidemic wave of COVID-19 in Hong Kong was attributed to a mutation in ORF3a, and another ORF3a mutation was previously associated with higher mortality rate [58, 59]. We identified 9 high-confidence ORF3a interactions, including late endosome membrane proteins VPS39 and PLD1, as well as endosome recycling proteins RAB11FIP1/2/5, which could serve as potential therapeutic targets. Altogether, this comparison of similarly produced datasets will allow for higher-confidence follow-up studies based on interaction evidence put forth by three separate interactome studies.

Discussion

Previous attempts to map the SARS-CoV-2 viral interactome have varied in experimental approach, data analysis parameters, cell lines used, and specific viral baits [11-18]. We report here both global proteome analysis and BioID-based proximity interactome analysis in human A549 lung cells for all but three SARS-CoV-2 viral proteins and compare this data to previously reported COVID-BioID datasets to identify consistently reported candidates. We developed a website to host this data to allow for more in-depth analysis of global proteomic changes in response to individual viral proteins, analyze enriched interaction candidates, pathway enrichment bioinformatics analysis, and comparative analysis to other reported datasets including the ISG SARS-CoV-2 inhibitors and previous COVID-BioID datasets. While this report primarily discusses unique identifications to highlight the highest-confidence interactors, the COVID-19 Proteomics website (<https://alexproteomics.shinyapps.io/covid19proteomics>) will allow for multiplexed and variable in-depth analysis of the MS data presented here. This study complements previous proximity interactome studies by strengthening confidence in reported interactors, as well as identifying new interactors and potentially affected pathways.

In this study, we profiled 26 of the 29 known SARS-CoV-2 viral proteins. Consistent with previous studies, we noted substantial suppression of NSP1 translation in our NSP1-BioID2 stable cell line (data not shown) and therefore chose to exclude NSP1 from this study [17, 60]. Surprisingly, we were unable to successfully express the Spike protein, which is one of the most widely-studied SARS-CoV-2 proteins, and chose not to pursue Spike-BioID due to the extensive characterization of Spike already underway (see [61] for review). Finally, the hypothetical ORF3b protein was also excluded due to our inability to generate cell lines stably expressing this protein. The use of human A549 lung-cancer derived cells for these studies is both a strength and a limitation. These cells do retain some fundamental traits of alveolar type-II pulmonary epithelial cells; however, A549 cells are not clearly representative of normal human pulmonary epithelial cells.

Our global profiles of human lung cells overexpressing individual SARS-CoV-2 viral proteins produced a large dataset of significantly upregulated or downregulated cellular proteins, enabling the ability to identify specific viral proteins influencing specific changes in cell biology. This data supports previous reports of ITGB3 overexpression in SARS-CoV-2 infected cells and tissues, and further identifies the specific viral proteins that could be influencing the overexpression. If ITGB3 is indeed working as an alternate receptor for SARS-CoV-2 viral uptake, it may be that targeting ITGB3 or the specific viral proteins that upregulate ITGB3 levels could have therapeutic benefit to slow cell-to-cell spread of the virus. Additionally, our findings that several of the SARS-CoV-2 proteins can reduce cellular levels of MUC5AC/B, possibly via increased secretion [62, 63], gives insight into one of the mechanisms by which the virus causes devastation of the respiratory system in the most severe COVID-19 cases.

While previous interactome studies have reported PPI candidates even when identified in up to 6 viral protein interactomes [16, 17], we highlighted here only unique protein candidates for each viral bait, as to reduce the possibility of capturing promiscuous interactors and keeping in mind that due to its limited coding capacity RNA viruses have likely very little functional redundancy within their genomes [64, 65]. For this reason, and for brevity, we chose to pursue a strict analysis for this report, predominantly focusing on PPIs uniquely identified for each bait. Unfortunately, the strict thresholds used for this report returned no significant interactors for NSP5 or NSP10; however, further analysis of the data utilizing the website could yield true interactors. While this approach should allow for the identification of high-confidence interactors, it is important to note that many of the viral proteins reside in the same subcellular compartments and would thus likely be proximate to many of the same proteins thus leading to their detection. However, it may be that proteins identified by more than one viral bait do have unique biological relevance to at least one viral bait, thus we have enabled iBAQ intensity analysis on the website to allow for more in-depth comparative analysis of the MS data to reveal those more substantial associations.

Our data further supports previously published studies including the role of ORF3a in extensive membrane remodeling and viral budding via interaction with VPS39 and VPS11, and suggest potential novel interactions between Orf3a and cell adhesion factors, which are important for cell-to-cell communication. In line with previous studies, our ORF6 data supports interaction with SEC-complex proteins and suggests novel roles in cell cycle regulation and viral immune evasion via deubiquitination mechanisms [16, 17]. ORF8 is known to play a role in immune evasion, and our data supports a possible role in viral immune evasion via O-linked glycosylation and suggests immune signaling disruption via interaction with effectors and

regulators of the type I IFN response. Our data has demonstrated clear support of previously published reports, and our novel findings implicating new roles for SARS-CoV-2 viral proteins will allow for an even more comprehensive understanding of how SARS-CoV-2 interacts with the host cells.

As variants arise and COVID-19 infections continue to threaten lives and cause lingering effects through post-COVID syndrome, the need for a clear SARS-CoV-2 viral-host interactome has never been more evident. The ability to identify crucial viral-host interactions and potentially disrupt those interactions with appropriate therapeutics could allow for fast-tracked treatments to be made available to those suffering from COVID-19 and long-term symptoms. More studies will be needed to verify the viral-host interactions presented here, and drug studies will be necessary to assess PPI relevance and effects on SARS-CoV-2 viability, but the resources provided here will serve to help guide researchers and provide high-confidence directions for future studies.

Materials and Methods

Plasmids

SARS-CoV-2 viral proteins were amplified via PCR from Addgene constructs with a 1x (GGGS) linker incorporated into each primer set (see Table S1). Amplified PCR products were fused to biotin ligases via In-Fusion Recombination into myc-BioID2 pBabe (Addgene #80900; XhoI/PmeI), BioID2-HA pBabe (Addgene #120308; BamHI/EcoRI), or TurboID-3xHA pBabe (BamHI/EcoRI) [25]. mycBioID2 (Addgene #80900) was used as a control for BioID2 cell lines. Human albumin signal sequence-3xHA-TurboID-KDEL pBabe control construct was made by two-step In-Fusion Recombination. Human albumin signal sequence and 3xHA-TurboID [25] were PCR-amplified with KDEL built into the reverse primer. Fragments were inserted into myc-BioID2 pBabe, replacing mycBioID2 (Addgene #80900; EcoRI/PmeI). All fusion-protein plasmids will be made available on Addgene.

Cell Culture

A549 cells were obtained from the American Type Culture Collection (ATCC; CCL-185™). Stable cell lines for all constructs were generated using retroviral transduction. HEK293 Phoenix cells (National Gene Vector Biorepository, Indianapolis, IN) were transfected with each construct using Lipofectamine 3000 (Thermo Fisher Scientific) per manufacturer's recommendation. The transfected cells were incubated at 37°C for 6 hours. After 6 hours incubation, the transfected cells were replenished with fresh medium and further incubated at 32°C for 72 h. The culture media was filtered through a 0.45-µm filter and added to A549 cells along with Polybrene (4 µg/ml; Santa Cruz Biotechnology, Dallas, TX). At 72 h after transduction, puromycin (0.5 µg/ml; Thermo Fisher Scientific) was added to the target cells. The expression of fusion proteins and functional biotinylation following addition of 50µM biotin was further verified using IF and WB. The stable cell lines were maintained in 5.0% CO₂ at 37°C in DMEM (HyClone, Logan, UT) supplemented with 10% fetal bovine serum (FBS). All cells were tested monthly for mycoplasma contamination.

Immunofluorescence

Cells grown on glass coverslips were fixed in 3% (wt/vol) paraformaldehyde/phosphate-buffered saline (PBS) for 10 min and permeabilized by 0.4% (wt/vol) Triton X-100/PBS for 15 min. For labeling fusion proteins, chicken anti-BioID2 (1:5000; BID2-CP-100; BioFront Technologies) or mouse anti-hemagglutinin primary antibody was used (HA; 1:1000; 12CA5; Covance). The primary antibody was detected using Alexa Fluor 568–conjugated goat anti-chicken (1:1000; A11041, Invitrogen) or Alexa Fluor 568–conjugated goat anti-mouse (1:1000; A11004; Thermo Fisher Scientific). Alexa Fluor 488–conjugated streptavidin (S32354; Thermo Fisher Scientific) was used to detect biotinylated proteins. DNA was detected with Hoechst dye 33342. Coverslips were mounted using 10% (wt/vol) Mowiol 4-88 (Polysciences). Epifluorescence images were captured using a Nikon Eclipse NiE (40 ×/0.75 Plan Apo Nikon objective) microscope.

Western Blot Analysis

To analyze total cell lysates by immunoblot, 1.2×10^6 cells were lysed in SDS–PAGE sample buffer, boiled for 5 min, and sonicated to shear DNA. Proteins were separated on 4–20% gradient gels (Mini-PROTEAN TGX; Bio-Rad, Hercules, CA) and transferred to nitrocellulose membrane (Bio-Rad). After blocking with 10% (vol/vol) adult bovine serum and 0.2% Triton X-100 in PBS for 30 min, the membrane was incubated with appropriate primary antibodies: chicken anti-BioID2 (1:5000; BID2-CP-100; BioFront Technologies) or rabbit polyclonal anti-hemagglutinin (1:2000; Ab9110; Abcam). The primary antibodies were detected using horseradish peroxidase (HRP)–conjugated anti-chicken (1:40,000; A9046; Sigma-Aldrich) or anti-rabbit (1:40,000; G21234; Thermo Fisher Scientific). The signals from antibodies were detected using enhanced chemiluminescence via a Bio-Rad ChemiDoc MP System (Bio-Rad, Hercules, CA). Following detection of each antibody, the membrane was quenched with 30% H_2O_2 for 30 minutes. To detect biotinylated proteins, the membrane was incubated with HRP-conjugated streptavidin (1:40,000; ab7403; Abcam) in 0.2% Triton X-100 in PBS for 45 min.

Sample Preparation

For each large-scale BioID2 pulldown sample, two 10cm dishes at 80% confluency were incubated with 50 μ m biotin for 18 hours, washed twice with PBS, and cell pellets collected for automated BioID pulldown at the Proteomics Facility at Sanford Burnham Prebys Medical Institute. TurboID samples were prepared similarly, but were treated with 50 μ m biotin for only 4 hours. Briefly, cells were lysed in 8M urea, 50 mM ammonium bicarbonate (ABC) and benzonase, and the lysate was centrifuged at 14,000 x g for 15 minutes to remove cellular debris. Supernatant protein concentration was determined using a bicinchoninic acid (BCA) protein assay (Thermo Scientific). Disulfide bridges were reduced with 5 mM tris(2-carboxyethyl)phosphine (TCEP) at 30°C for 60 min, and cysteines were subsequently alkylated with 15 mM iodoacetamide (IAA) in the dark at room temperature for 30 min. Each sample was separated into two aliquots, one for whole cell proteome profiling and the other for proximity-dependent labeling analysis. Whole cell protein lysate was digested overnight with mass spec grade Trypsin/Lys-C mix (1:25 enzyme/substrate ratio). Following digestion, samples were acidified with formic acid (FA) and subsequently desalted using AssayMap C18 cartridges mounted on an Agilent AssayMap BRAVO liquid handling system. Cartridges were sequentially conditioned with 100% acetonitrile (ACN) and 0.1% FA, samples were then loaded, washed with 0.1% FA, and peptides eluted with 60% ACN, 0.1% FA. Finally, the organic solvent was removed in a SpeedVac concentrator prior to LC-MS/MS analysis.

Affinity-purification of biotinylated proteins

Affinity purification and digestion of biotinylated proteins were carried out in an automated fashion in a Bravo AssayMap platform (Agilent) using AssayMap streptavidin cartridges (Agilent). Briefly, cartridges were first primed with 50 mM ammonium bicarbonate, and then proteins were slowly loaded onto the streptavidin cartridge. Background contamination was removed with 8M urea, 50 mM ammonium bicarbonate. Finally, cartridges were washed with Rapid digestion buffer (Promega, Rapid digestion buffer kit) and proteins were subjected to on-cartridge digestion with mass spec grade Trypsin/Lys-C Rapid digestion enzyme (Promega, Madison, WI) at 70°C for 1h. Digested peptides were then desalted in the Bravo platform using AssayMap C18 cartridges, and dried down in a SpeedVac concentrator.

Mass Spectrometry

Prior to LC-MS/MS analysis, dried peptides were reconstituted with 2% ACN, 0.1% FA and concentration was determined using a NanoDropTM spectrophotometer (ThermoFisher). Samples were then analyzed by LC-MS/MS using a Proxeon EASY-nanoLC system (ThermoFisher) coupled to an Orbitrap Fusion Lumos mass spectrometer (Thermo Fisher Scientific). Peptides were separated using an analytical C18 Aurora column (75 μ m x 250 mm, 1.6 μ m particles; IonOpticks) at a flow rate of 300 nL/min (60°C) using a 75-min gradient: 1% to 5% B in 1 min, 6% to 23% B in 44 min, 23% to 34% B in 28 min, and 34% to 48% B in 2 min (A= FA 0.1%; B=80% ACN: 0.1% FA). The mass spectrometer was operated in positive data-dependent acquisition mode. MS1 spectra were measured in the Orbitrap in a mass-to-charge (m/z) of 375 – 1500 with a resolution of 60,000 at m/z 200. Automatic gain control target was set to 4×10^5 with a maximum injection time of 50 ms. The instrument was set to run in top speed mode with 2-second cycles for the survey and the MS/MS scans. After a survey scan, the most abundant precursors (with charge state between +2 and +7) were isolated in the quadrupole with an isolation window of 0.7 m/z and fragmented with HCD at 30% normalized collision energy. Fragmented precursors were detected in the ion trap as rapid scan mode with automatic gain control target set to 1×10^4 and a maximum injection time set at 35 ms. The dynamic exclusion was set to 20 seconds with a 10 ppm mass tolerance around the precursor.

Data Analysis

All raw files were processed with MaxQuant (version 1.5.5.1) using the integrated Andromeda Search engine against a target/decoy version of the curated human Uniprot proteome without isoforms (downloaded in January of 2020) and the GPM cRAP sequences (commonly known protein contaminants). First search peptide tolerance was set to 20 ppm, main search peptide tolerance was set to 4.5 ppm. Fragment mass tolerance was set to 20 ppm. Trypsin was set as enzyme in specific mode and up to two missed cleavages was allowed. Carbamidomethylation of cysteine was specified as fixed modification and protein N-terminal acetylation and oxidation of methionine were considered variable modifications. In addition, the phosphopeptide-enriched samples were also searched with phosphorylation of serine, threonine and tyrosine was considered as variable modification. The target-decoy-based false discovery rate (FDR) filter for spectrum and protein identification was set to 1%.

Statistical analysis of interactome data was carried out using in-house R script (version 3.5.1, 64-bit), including R Bioconductor packages such as limma and MSstats. First, peptide intensities were log₂-transformed and loess-normalized (limma package) across replicates of each bait or control batch to account for systematic errors. Note that normalization was not carried out across all samples due to significant differences in pulldowns of different baits and/or their controls. Testing for differential abundance was performed using MSstats bioconductor package based on a linear mixed-effects model. Importantly, the log₂FC and pvalue of proteins missing completely in one condition (i.e., ORF or control) was imputed as follows. The imputed Log₂FC of each bait vs control comparison was calculated as the average of the protein

intensity (i.e., sum of peptide intensities of a given protein within a given sample) across the triplicate of the same bait, divided by 3.3. On the other hand, the imputed pvalue was computed by dividing 0.05 by the number of replicates of a given bait the protein was confidently identified. Therefore, the imputed log₂FC gives a notion of the average protein intensity in a pulldown, while the imputed pvalue reports the confidence of identification in the sense of reproducibility of detection. In addition, we generated a CrapomeScore for each identified protein in the experiment. The CrapomeScore is the fraction of all streptavidin-based experiments in the Crapome database (reprint-apms.org) that the prey protein is identified. The CrapomeScore ranges from 0 to 1, and a protein with a score of 1 means that it was identified in all streptavidin-based experiments in the Crapome database.

Network analysis of SARS-CoV-2 interactors

A hierarchical model of cellular processes and structures predicted to interact with SARS-CoV-2 was derived via multi-scale community detection performed on a large protein interaction network. We selected a network derived from the STRING database as our starting network: the subset of the STRING interactions with a combined confidence score greater than 0.7 (available in the Network Data Exchange (NDEx) at <https://www.ndexbio.org/viewer/networks/275bd84e-3d18-11e8-a935-0ac135e8bacf> [66, 67]). Then, human proteins interacting with SARS-CoV-2 proteins were filtered by log₂FC≥2.32, p value≤0.01, and n≥2. The specific high-confidence interactions also were filtered based on the CRAPome contaminant database with a score ≤0.5 [68]. A subnetwork "proximal" to those proteins then was identified by network propagation using the Cytoscape Diffusion tool [69].

Multi-scale community detection analysis was performed on this subnetwork using the community detection algorithm HiDeF via the Community Detection APplication and Service (CDAPS; app available at <http://apps.cytoscape.org/apps/cycommunitydetection>) [70, 71]. The resulting hierarchical model describes "communities" in the network at multiple scales, where communities are subnetworks of proteins interacting more with each other than with other proteins in the network. The analysis also infers a structure to the network, one in which communities are hypotheses for processes or structures that interact with SARS-CoV-2 proteins. The communities are organized into a hierarchy in which larger communities subsume smaller communities [72, 73]. Finally, the hierarchy network (<https://doi.org/10.18119/N9531R>) was styled, communities were subjected to enrichment analysis in GO biological processes using the g:Profiler package in CDAPS, p values were calculated based on the hypergeometric distribution, and a layout was applied.

Virus-centric analysis of SARS-CoV-2 interactors

To provide a visual model that displays high-confidence cellular factors that interact with individual SARS-CoV-2 proteins, we utilized the same network derived from the STRING database (confidence score >0.7), with protein groups that had degree of connection =1, log₂FC≥2.32, and p value≤ 0.01. In addition, we filtered for promiscuity using the CRAPome repository (Crapome≤0.5) and included only proteins that were found in 2 or more biological replicates [68]. The resultant high-confidence interactors were visualized using Cytoscape (v3.8.0) and tested for enrichment in GO biological process terms using the hypergeometric distribution [74].

Web-based Shiny App

An accompanying web-based Shiny application (<https://alexproteomics.shinyapps.io/covid19proteomics>) was created to allow visualization and further functional analysis of the BioID and whole cell proteome statistical analysis data. The application uses several applications, including clusterProfiler (v3.18.1) for functional analysis with the *enricher* function using the Broad Institute molecular signature databases (v7.4) including canonical pathways (Reactome, KEGG, WikiPathways), Immune collection, chemical and genetic perturbation signatures, regulatory transcription factor targets (TFT), oncogenic signatures, and Gene Ontology (Human Phenotype, Cellular Component, Biological Process and Molecular Function).

Acknowledgements

This research was funded by Sanford Health, NIH grant R35GM126949, NIGMS grant U24 CA184427, NIAID grants U19AI135972 and U19 AI118610, and NCI Cancer Center Support Grant P30 CA030199. The Imaging Core and Biochemistry Core at Sanford Research, which facilitated these studies, are supported by Institutional Development Awards from the National Institute of General Medical Sciences and the National Institutes of Health under grant P20GM103620. This work was further supported by the Sanford Burnham Prebys Proteomics Core.

Conflicts of Interest

Sanford Research has licensed BioID reagents to BioFront Technologies. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Author Contributions

KJR and ARC conceived the project and secured funding. DGM, RJC, KLS, and CTH cloned COVID19-BioID plasmids, generated and validated stable cell lines, and harvested cells for BioID pulldown. RDP and ARC performed sample preparation, affinity purification, and mass spectrometry acquisition and data analysis. DGM, LMS, VA, DP, SL, and ARC performed bioinformatics analysis and data interpretation. ARC created the website. DGM, LMS, VA, and KJR wrote the manuscript, with input from other authors.

References

1. Villapol, S., *Gastrointestinal symptoms associated with COVID-19: impact on the gut microbiome*. *Transl Res*, 2020. **226**: p. 57-69.
2. De Lorenzo, A., et al., *Acute cardiac injury in patients with COVID-19*. *Am J Cardiovasc Dis*, 2020. **10**(2): p. 28-33.
3. Davies, D.A., A. Adlimoghaddam, and B.C. Albeni, *The Effect of COVID-19 on NF-kappaB and Neurological Manifestations of Disease*. *Mol Neurobiol*, 2021.
4. Maltezou, H.C., A. Pavli, and A. Tsakris, *Post-COVID Syndrome: An Insight on Its Pathogenesis*. *Vaccines (Basel)*, 2021. **9**(5).
5. Orru, G., et al., *Long-COVID Syndrome? A Study on the Persistence of Neurological, Psychological and Physiological Symptoms*. *Healthcare (Basel)*, 2021. **9**(5).
6. Saadaoui, M., M. Kumar, and S. Al Khodor, *COVID-19 Infection during Pregnancy: Risk of Vertical Transmission, Fetal, and Neonatal Outcomes*. *J Pers Med*, 2021. **11**(6).

7. Bari, E., et al., *Mesenchymal Stromal Cell Secretome for Post-COVID-19 Pulmonary Fibrosis: A New Therapy to Treat the Long-Term Lung Sequelae?* Cells, 2021. **10**(5).
8. Tenforde, M.W., et al., *Symptom Duration and Risk Factors for Delayed Return to Usual Health Among Outpatients with COVID-19 in a Multistate Health Care Systems Network - United States, March-June 2020.* MMWR Morb Mortal Wkly Rep, 2020. **69**(30): p. 993-998.
9. Logue, J.K., et al., *Sequelae in Adults at 6 Months After COVID-19 Infection.* JAMA Netw Open, 2021. **4**(2): p. e210830.
10. Pavli, A., M. Theodoridou, and H.C. Maltezos, *Post-COVID syndrome: Incidence, clinical spectrum, and challenges for primary healthcare professionals.* Arch Med Res, 2021.
11. Gordon, D.E., et al., *A SARS-CoV-2 protein interaction map reveals targets for drug repurposing.* Nature, 2020. **583**(7816): p. 459-468.
12. Hoffmann, H.H., et al., *Functional interrogation of a SARS-CoV-2 host protein interactome identifies unique and shared coronavirus host factors.* bioRxiv, 2020.
13. Kumar, N., et al., *Integrative Network Biology Framework Elucidates Molecular Mechanisms of SARS-CoV-2 Pathogenesis.* iScience, 2020. **23**(9): p. 101526.
14. Li, J., et al., *Virus-Host Interactome and Proteomic Survey Reveal Potential Virulence Factors Influencing SARS-CoV-2 Pathogenesis.* Med (N Y), 2020.
15. Li, J., et al., *Virus-host interactome and proteomic survey of PMBCs from COVID-19 patients reveal potential virulence factors influencing SARS-CoV-2 pathogenesis.* bioRxiv, 2020: p. 2020.03.31.019216.
16. Laurent, E.M.N., et al., *Global BioID-based SARS-CoV-2 proteins proximal interactome unveils novel ties between viral polypeptides and host factors involved in multiple COVID19-associated mechanisms.* bioRxiv, 2020: p. 2020.08.28.272955.
17. Samavarchi-Tehrani, P., et al., *A SARS-CoV-2 – host proximity interactome.* bioRxiv, 2020: p. 2020.09.03.282103.
18. Stukalov, A., et al., *Multilevel proteomics reveals host perturbations by SARS-CoV-2 and SARS-CoV.* Nature, 2021.
19. Davies, J.P., et al., *Comparative multiplexed interactomics of SARS-CoV-2 and homologous coronavirus non-structural proteins identifies unique and shared host-cell dependencies.* bioRxiv, 2020.
20. Boruchowicz, H., et al., *The XPO6 exportin mediates HSV-1 gM nuclear release late in infection.* J Virol, 2020.
21. Coyaud, E., et al., *Global Interactomics Uncovers Extensive Organellar Targeting by Zika Virus.* Mol Cell Proteomics, 2018. **17**(11): p. 2242-2255.
22. Fan, J., et al., *Proximity proteomics identifies novel function of Rab14 in trafficking of Ebola virus matrix protein VP40.* Biochem Biophys Res Commun, 2020. **527**(2): p. 387-392.
23. Rider, M.A., et al., *The interactome of EBV LMP1 evaluated by proximity-based BioID approach.* Virology, 2018. **516**: p. 55-70.
24. V'Kovski, P., et al., *Determination of host proteins composing the microenvironment of coronavirus replicase complexes by proximity-labeling.* Elife, 2019. **8**.
25. May, D.G., et al., *Comparative Application of BioID and TurboID for Protein-Proximity Biotinylation.* Cells, 2020. **9**(5).
26. Islam, A., et al., *Transcriptome of nasopharyngeal samples from COVID-19 patients and a comparative analysis with other SARS-CoV-2 infection models reveal disparate host responses against SARS-CoV-2.* J Transl Med, 2021. **19**(1): p. 32.
27. Sigrist, C.J., A. Bridge, and P. Le Mercier, *A potential role for integrins in host cell entry by SARS-CoV-2.* Antiviral Res, 2020. **177**: p. 104759.

28. Wada, M., et al., *Interplay between coronavirus, a cytoplasmic RNA virus, and nonsense-mediated mRNA decay pathway*. Proc Natl Acad Sci U S A, 2018. **115**(43): p. E10157-E10166.
29. Han, J., K. Pluhackova, and R.A. Bockmann, *The Multifaceted Role of SNARE Proteins in Membrane Fusion*. Front Physiol, 2017. **8**: p. 5.
30. Abu-Farha, M., et al., *The Role of Lipid Metabolism in COVID-19 Virus Infection and as a Drug Target*. Int J Mol Sci, 2020. **21**(10).
31. Tikou, V., M.W. Tan, and I. Dikic, *Mitochondrial Functions in Infection and Immunity*. Trends Cell Biol, 2020. **30**(4): p. 263-275.
32. Aguirre, S., et al., *Dengue virus NS2B protein targets cGAS for degradation and prevents mitochondrial DNA sensing during infection*. Nat Microbiol, 2017. **2**: p. 17037.
33. Dong, D., W. Xie, and M. Liu, *Alteration of cell junctions during viral infection*. Thorac Cancer, 2020. **11**(3): p. 519-525.
34. Roy, U., et al., *Role of MRP transporters in regulating antimicrobial drug inefficacy and oxidative stress-induced pathogenesis during HIV-1 and TB infections*. Front Microbiol, 2015. **6**: p. 948.
35. Cortese, M., et al., *Integrative Imaging Reveals SARS-CoV-2-Induced Reshaping of Subcellular Morphologies*. Cell Host Microbe, 2020. **28**(6): p. 853-866 e5.
36. York, A., *Exploiting peroxisomes*. Nat Rev Microbiol, 2018. **16**(11): p. 659.
37. Wang, X., G. Melino, and Y. Shi, *Actively or passively deacidified lysosomes push beta-coronavirus egress*. Cell Death Dis, 2021. **12**(3): p. 235.
38. Martin-Serrano, J., et al., *HECT ubiquitin ligases link viral and cellular PPXY motifs to the vacuolar protein-sorting pathway*. J Cell Biol, 2005. **168**(1): p. 89-101.
39. Lee, J.G., et al., *Characterization of SARS-CoV-2 proteins reveals Orf6 pathogenicity, subcellular localization, host interactions and attenuation by Selinexor*. Cell Biosci, 2021. **11**(1): p. 58.
40. Kuo, M.L., et al., *PYCR1 and PYCR2 Interact and Collaborate with RRM2B to Protect Cells from Overt Oxidative Stress*. Sci Rep, 2016. **6**: p. 18846.
41. Bagga, S. and M.J. Bouchard, *Cell cycle regulation during viral infection*. Methods Mol Biol, 2014. **1170**: p. 165-227.
42. Zhang, H., et al., *Ubiquitin-Modified Proteome of SARS-CoV-2-Infected Host Cells Reveals Insights into Virus-Host Interaction and Pathogenesis*. J Proteome Res, 2021. **20**(5): p. 2224-2239.
43. Flower, T.G., et al., *Structure of SARS-CoV-2 ORF8, a rapidly evolving immune evasion protein*. Proc Natl Acad Sci U S A, 2021. **118**(2).
44. Zhang, Y., et al., *The ORF8 protein of SARS-CoV-2 mediates immune evasion through down-regulating MHC-Iota*. Proc Natl Acad Sci U S A, 2021. **118**(23).
45. Young, B.E., et al., *Effects of a major deletion in the SARS-CoV-2 genome on the severity of infection and the inflammatory response: an observational cohort study*. Lancet, 2020. **396**(10251): p. 603-611.
46. Vigerust, D.J. and V.L. Shepherd, *Virus glycosylation: role in virulence and immune interactions*. Trends Microbiol, 2007. **15**(5): p. 211-8.
47. Watanabe, Y., et al., *Exploitation of glycosylation in enveloped virus pathobiology*. Biochim Biophys Acta Gen Subj, 2019. **1863**(10): p. 1480-1497.
48. Kim, S.A., et al., *Myotubularin and MTMR2, phosphatidylinositol 3-phosphatases mutated in myotubular myopathy and type 4B Charcot-Marie-Tooth disease*. J Biol Chem, 2002. **277**(6): p. 4526-31.
49. Angelini, M.M., et al., *Severe acute respiratory syndrome coronavirus nonstructural proteins 3, 4, and 6 induce double-membrane vesicles*. mBio, 2013. **4**(4).
50. Byun, H., et al., *ERAD and how viruses exploit it*. Front Microbiol, 2014. **5**: p. 330.
51. Martin-Sancho, L., et al., *Functional landscape of SARS-CoV-2 cellular restriction*. Mol Cell, 2021.
52. Miller, C.L., *Stress Granules and Virus Replication*. Future Virol, 2011. **6**(11): p. 1329-1338.

53. Cascarina, S.M. and E.D. Ross, *A proposed role for the SARS-CoV-2 nucleocapsid protein in the formation and regulation of biomolecular condensates*. *FASEB J*, 2020. **34**(8): p. 9832-9842.
54. Wang, J., et al., *SARS-CoV-2 nucleocapsid protein undergoes liquid-liquid phase separation into stress granules through its N-terminal intrinsically disordered region*. *Cell Discov*, 2021. **7**(1): p. 5.
55. Lu, F., *SARS-CoV-2 ORF9c: a mysterious membrane-anchored protein that regulates immune evasion?* *Nat Rev Immunol*, 2020. **20**(11): p. 648.
56. Shi, C.S., et al., *SARS-Coronavirus Open Reading Frame-8b triggers intracellular stress pathways and activates NLRP3 inflammasomes*. *Cell Death Discov*, 2019. **5**: p. 101.
57. Samson, A.L., et al., *MLKL trafficking and accumulation at the plasma membrane control the kinetics and threshold for necroptosis*. *Nat Commun*, 2020. **11**(1): p. 3151.
58. Chu, D.K.W., et al., *Introduction of ORF3a-Q57H SARS-CoV-2 Variant Causing Fourth Epidemic Wave of COVID-19, Hong Kong, China*. *Emerg Infect Dis*, 2021. **27**(5): p. 1492-1495.
59. Majumdar, P. and S. Niyogi, *ORF3a mutation associated with higher mortality rate in SARS-CoV-2 infection*. *Epidemiol Infect*, 2020. **148**: p. e262.
60. Kamitani, W., et al., *A two-pronged strategy to suppress host protein synthesis by SARS coronavirus Nsp1 protein*. *Nat Struct Mol Biol*, 2009. **16**(11): p. 1134-40.
61. Huang, Y., et al., *Structural and functional properties of SARS-CoV-2 spike protein: potential antiviral drug development for COVID-19*. *Acta Pharmacol Sin*, 2020. **41**(9): p. 1141-1149.
62. Edwards, D.A., et al., *Exhaled aerosol increases with COVID-19 infection, age, and obesity*. *Proc Natl Acad Sci U S A*, 2021. **118**(8).
63. Lu, W., et al., *Elevated MUC1 and MUC5AC mucin protein levels in airway mucus of critical ill COVID-19 patients*. *J Med Virol*, 2021. **93**(2): p. 582-584.
64. Elena, S.F., et al., *Mechanisms of genetic robustness in RNA viruses*. *EMBO Rep*, 2006. **7**(2): p. 168-73.
65. Krakauer, D.C. and J.B. Plotkin, *Redundancy, antiredundancy, and the robustness of genomes*. *Proc Natl Acad Sci U S A*, 2002. **99**(3): p. 1405-9.
66. Pratt, D., et al., *NDEx 2.0: A Clearinghouse for Research on Cancer Pathways*. *Cancer Res*, 2017. **77**(21): p. e58-e61.
67. Pratt, D., et al., *NDEx, the Network Data Exchange*. *Cell Syst*, 2015. **1**(4): p. 302-305.
68. Mellacheruvu, D., et al., *The CRAPome: a contaminant repository for affinity purification-mass spectrometry data*. *Nat Methods*, 2013. **10**(8): p. 730-6.
69. Carlin, D.E., et al., *Network propagation in the cytoscape cyberinfrastructure*. *PLoS Comput Biol*, 2017. **13**(10): p. e1005598.
70. Zheng, F., et al., *HiDeF: identifying persistent structures in multiscale 'omics data*. *Genome Biol*, 2021. **22**(1): p. 21.
71. Singhal, A., et al., *Multiscale community detection in Cytoscape*. *PLoS Comput Biol*, 2020. **16**(10): p. e1008239.
72. Yu, M.K., et al., *Translation of Genotype to Phenotype by a Hierarchy of Cell Subsystems*. *Cell Syst*, 2016. **2**(2): p. 77-88.
73. Kramer, M., et al., *Inferring gene ontologies from pairwise similarity data*. *Bioinformatics*, 2014. **30**(12): p. i34-42.
74. Shannon, P., et al., *Cytoscape: a software environment for integrated models of biomolecular interaction networks*. *Genome Res*, 2003. **13**(11): p. 2498-504.

