A map of direct SARS-CoV-2 protein interactions implicates specific human host processes

Dae-Kyum Kim,[1,2,3,16] Benjamin Weller,[4,16] Chung-Wen Lin,[4,16] Dayag Sheykhkarimli,[1,2,3,16] Jennifer J. Knapp,[1,2,3,16] Nishka Kishore,[1,2,3] Mayra Sauer,[4] Ashyad Rayhan,[1,2,3] Veronika Young,[4] Nora Marin-de la Rosa,[4] Oxana Pogoutse,[1,2,3] Kerstin Spirohn,[5,6,7] Alexandra Strobel,[4] Florent Laval,[5,6,7,8,9,10] Patrick Schwehn,[4] Roujia Li,[1,2,3] Simin Rothballer,[4] Melina Altmann,[4] Patricia Cassonnet,[11,12,13] Guillaume Dugied,[11,12,13] Atina G. Cote,[1,2,3] Lena Elorduy Vergara,[4] Isaiah Hazelwood,[1,2,3] Bingruo B. Liu,[1,2,3] Maria Nguyen,[1,2,3] Ramakrishnan Pandiarajan,[4] Patricia A. Rodriguez Coloma,[4] Luc Willems,[8,10] Jean-Claude Twizere,[5,9,10] Caroline Demeret,[11,12,13] Yves Jacob,[11,12,13] Tong Hao,[5,6,7] Dave E. Hill,[5,6,7] Claudia Falter,[4] Marc Vidal,[5,6,*] Michael A. Calderwood,[5,6,7,*] Frederick P. Roth,[1,2,3,14,*] and Pascal Falter-Braun[4,15,17,*]

[1]Donnelly Centre, University of Toronto, Toronto, Ontario, Canada

[2]Department of Molecular Genetics, University of Toronto, Toronto, Ontario, Canada

[3]Lunenfeld-Tanenbaum Research Institute (LTRI), Sinai Health System, Toronto, Ontario, Canada

[4]Institute of Network Biology (INET), Helmholtz Center Munich, German Research Center for Environmental Health, Munich-Neuherberg, Germany

[5]Center for Cancer Systems Biology (CCSB), Dana-Farber Cancer Institute, Boston, MA, USA

[6]Department of Genetics, Blavatnik Institute, Harvard Medical School, Boston, MA, USA

[7]Department of Cancer Biology, Dana-Farber Cancer Institute, Boston, MA, USA

[8]Laboratory of Molecular and Cellular Epigenetics, GIGA Institute, University of Liège, Liège, Belgium

[9]Laboratory of Viral Interactomes, GIGA Institute, University of Liège, Liège, Belgium

[10]TERRA Teaching and Research Centre, University of Liège, Gembloux, Belgium

[11]Unité de Génétique Moléculaire des Virus à ARN, Département Virologie, Institut Pasteur, Paris, France

[12]UMR3569, Centre National de la Recherche Scientifique, Paris, France

[13]Université de Paris, Paris, France

[14]Department of Computer Science, University of Toronto, Toronto, Ontario, Canada

[15]Microbe-Host Interactions, Faculty of Biology, Ludwig Maximilians University (LMU) Munich, Planegg-Martinsried, Germany

[16]These authors made equal contributions.

[17]Lead contact

*Correspondence: Marc Vidal, Email: marc_vidal@dfci.harvard.edu; Michael A. Calderwood, Email: michael_calderwood@dfci.harvard.edu; Frederick P. Roth, Email: fritz.roth@utoronto.ca; Pascal Falter-Braun, Email: pascal.falter-braun@helmholtz-muenchen.de.

ABSTRACT

Key steps in viral propagation, immune suppression and pathology are mediated by direct, binary physical interactions between viral and host proteins. To understand the biology of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection, we generated an unbiased systematic map of binary physical interactions between viral and host interactions, complementing previous co-complex association maps by conveying more direct mechanistic understanding and enabling targeted disruption of direct interactions. To this end, we deployed two parallel strategies, identifying 205 virus-host and 27 intraviral binary interactions amongst 171 host and 19 viral proteins, with orthogonal validation by an internally benchmarked NanoLuc two-hybrid system to ensure high data quality. Host proteins interacting with SARS-CoV-2 proteins were enriched in various cellular processes, including immune signaling and inflammation, protein ubiquitination, and membrane trafficking. Specific subnetworks provide new hypotheses related to viral modulation of host protein homeostasis and T-cell regulation. The direct virus-host protein interactions we identified can now be prioritized as targets for therapeutic intervention. More generally, we provide a resource of systematic maps describing which SARS-CoV-2 and human proteins interact directly.

Keywords: COVID-19, SARS-CoV-2, physical protein interaction, immune signaling, protein ubiquitination, membrane trafficking

INTRODUCTION

Coronavirus disease of 2019 (COVID-19), a severe respiratory disease that emerged in December 2019, is caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) (Huang et al., 2020; Wu et al., 2020c). As of March 8th, 2021, over 116 million confirmed cases and ~2.6 million deaths have been reported globally (WHO report of coronavirus). Despite intensive research by the scientific community, many important open questions regarding molecular viral mechanisms and COVID-19 etiology remain. Several vaccines against SARS-CoV-2 have been approved, and deployment has begun in various countries. However, no vaccine offers 100% protection against infection, vaccination of the global population will take significant time, and multiple SARS-CoV-2 variants have emerged with varying potential to escape vaccine-derived immunity (Grubaugh et al., 2021). Moreover, therapies for infected patients who are acutely ill or suffering from 'Long COVID' have had limited efficacy (Carfì et al., 2020). SARS-CoV-2 is likely to be a public health risk into the foreseeable future, thus necessitating a deeper molecular understanding, with the potential to inform effective treatments of infected patients and preventive strategies.

Viruses exploit host processes throughout their replication cycle and must evade or block host immune responses. Several virus-host interactome ('virhostome') studies for other viruses have demonstrated the value of systematic network maps for understanding key viral processes (Calderwood et al., 2007; Muller et al., 2012; Neveu et al., 2012; Pfefferle et al., 2011). We have previously shown that systematic virus-host network maps can identify viral pathway perturbations underlying clinical disease manifestations (Gulbahce et al., 2012). For SARS-CoV-2, several studies have identified physical associations between viral and host proteins using affinity purification followed by mass spectrometry (AP-MS) (Gordon et al., 2020a, 2020b; Li et al., 2020; Nabeel-Shah et al., 2020; Stukalov et al., 2020), and identified spatially proximal host proteins using biotin identification (BioID) (Laurent et al., 2020; Samavarchi-Tehrani et al., 2020; St-Germain et al., 2020). Different technologies provide complementary views of the interactome, and protein relationships identified by distinct methods have different biological properties (Yu et al., 2008). Importantly, to this date no study sought to identify direct interactions between the proteins of SARS-CoV-2 and its human host, as opposed to protein pairs that are indirectly-associated or spatially proximal. Knowledge of direct protein-protein interactions (PPIs) is important for a mechanistic understanding of the virus-host relationship and enables efficient discovery of drugs that can disrupt these interactions (Lu et al., 2020), to obstruct the viral life cycle.

To sensitively identify direct interactions between viral and human host proteins, we carried out

protein interaction screens using a yeast two-hybrid (Y2H) strategy (Fields and Song, 1989), for which the vast majority of identified interactions are known to be direct (or 'binary') (Luck et al., 2020; Rolland et al., 2014). In total, we tested >400,000 virus-host protein pairs by each of two binary assays to reveal a PPI network comprising 205 direct virus-host and 27 intraviral interactions amongst 171 host and 19 viral proteins. Validation of the identified PPIs by an orthogonal split-luciferase based assay, the NanoLuc two-hybrid (N2H) system (Choi et al., 2019), demonstrated high data quality. The resulting binary virhostome interaction map suggested extensive viral protein crosstalk with: (i) regulation of immune signaling and inflammation; (ii) ubiquitin-mediated protein modulation and degradation; and (iii) the membrane trafficking mechanism. These data comprise an important resource for SARS-CoV-2 research, identifying a largely novel set of direct protein interactions that furthers our mechanistic understanding of SARS-CoV-2 and offers potential entry points for development of targeted COVID-19 therapeutics.

RESULTS AND DISCUSSION

Overview of screening methods

Because protein interaction screening methods differ in the subset of interactions they are able to detect (Braun et al., 2009), we employed two independent Y2H versions (Fig. 1A): (i) an auxotrophic imidazoleglycerol-phosphate dehydratase ("His3") selection with a growth-based readout ($Y2H_{HIS3}$) enabled by interaction-dependent transcriptional activation of the *HIS3* reporter gene, a system which has been used previously to generate high-quality interactome network maps (Altmann et al., 2018, 2020; Rual et al., 2005; Yu et al., 2008); and (ii) a green fluorescent protein (GFP) system ($Y2H_{GFP}$) based on the Barcode Fusion Genetics (BFG)-Y2H technology (Yachie et al., 2016), in which interaction-dependent activation of the GFP reporter gene is measured by fluorescence-activated cell sorting (FACS), enabling sorting of GFP-positive cells from a pooled culture (Kim et al., 2021; Yachie et al., 2016). The complementary detection profiles of $Y2H_{HIS3}$ and $Y2H_{GFP}$ (see below) are achieved by different system configurations: $Y2H_{HIS3}$ is based on N-terminally tagged viral and human open reading frame (ORF) constructs expressed from low-copy plasmids, while $Y2H_{GFP}$ employs C-terminal activation domain (AD) fusions of human ORFs expressed from high-copy plasmids, and also differs in terms of the length and sequence of the linker between the ORF and AD sequences.

Clone assembly and construction

Clones encoding SARS-CoV-2 proteins were independently synthesized and assembled (top panel of Fig. 1A). For $Y2H_{HIS3}$, Gateway Entry clones were generated based on SARS-CoV-2 reference (non-codon-optimized) ORFs from the sequenced genomes of three viral isolates (Wu et al., 2020a) (see Methods). Viral ORFs were transferred by recombinational cloning to the appropriate Gateway Destination vectors, which respectively express Gal4 DNA binding domain (DB) fused at the N-terminus of the 'bait' protein (DB-X) construct, and Gal4 activation domain (AD) fused at the N-terminus of the 'prey' protein (AD-Y) construct (see Methods). Human ORFs from two previously reported collections of haploid yeast clones were used, expressing ORFs N-terminally fused to Gal4 DB and AD domains (Luck et al., 2020).

For $Y2H_{GFP}$, our codon-optimized SARS-CoV-2 ORF collection was cloned into Gateway Entry plasmids (Kim et al., 2020), and then transferred into low-copy, barcoded bait and prey plasmids as N-terminal fusions. To interrogate human ORFs, a collection of high-copy prey plasmids with C-terminal fusions (see Methods) was assembled that collectively covered ~14K human ORFs, each generally

having two uniquely barcoded plasmids. We previously reported a subset of ~14K of these plasmids, with one uniquely-barcoded plasmid per ORF, (Luck et al., 2020), and an additional set of plasmids for ~14K ORFs was generated for this study to provide independently-barcoded replication.

For each of the two screens, the independently cloned viral ORFs included: non-structural proteins (NSPs) 1-16 (NSP11 omitted), derived from ORF1AB; the four structural proteins spike (S), envelope (E), membrane (M) and nucleocapsid (N); and 9 accessory genes, ORF3A, 3B, 6, 7A, 7B, 8, 9B, 9C and 10, distributed among the structural genes (Supplementary Table 1A; see Methods). Because the genome sequence and annotation of viral ORFs used for the two cloning efforts (Chan et al., 2020; Wu et al., 2020a) diverged for NSP12, ORF3B and ORF10, the two clone sets together cover 30 viral ORFs. Hereafter, we use a recently proposed nomenclature (Jungreis et al., 2020), so that NSP12 (Chan et al., 2020) and NSP12 (Wu et al., 2020a) will hereafter be referred to as NSP12-1 and NSP12-2, respectively, and ORF3B (Chan et al., 2020) and ORF3B (Wu et al., 2020a) will be referred to as ORF3D and ORF3B, respectively (Supplementary Table 1A).

<u>Screening for binary interactions between viral and human host proteins and amongst viral proteins</u>

Using $Y2H_{HIS3}$, 27 viral ORFs were screened against 17,472 human ORFs (83% 'completeness', i.e., covering 83% of the potential search space) in both orientations, i.e., with each protein screened as both bait and prey (Fig. 1B). Screens proceeded by mating each bait strain to 'mini-pools' of prey strains, followed by secondary phenotyping and sequencing of plasmids from isolated colonies to confirm interacting proteins (see Methods). Each identified human interaction candidate was tested three times in pairwise Y2H experiments against all 27 viral ORFs to verify the candidates. Only pairs scoring positive at least two times and not exhibiting an `auto-activation´ phenotype (growth in selective media in the absence of the prey plasmid) were considered *bona fide* Y2H interactions (see Methods). This screen yielded 119 interactions, collectively involving 14 viral and 93 human host proteins. We refer to this binary human SARS-CoV-2 interactome map as $HuSCI_{HIS3}$ (Human-SARS-CoV-2 interactome via the $Y2H_{HIS3}$ system).

For the $Y2H_{GFP}$ screen, we examined 14,627 prey human ORFs (70% completeness), most of which were represented by two uniquely-barcoded plasmids, collectively corresponding to 27,671 uniquely-barcoded yeast strains. Each of these strains was screened against bait strains for 28 viral ORFs (each represented by 2-6 uniquely-barcoded strains, collectively corresponding to 82 uniquely barcoded bait strains). Thus, we screened 409,556 bait-prey combinations represented by 2,269,022 uniquely barcoded diploid strains (Fig. 1B). Barcode sequencing of GFP-positive cells allowed

quantitative assessment of auto-activity levels and therefore, unlike $Y2H_{HIS3}$, enabled identification of genuine PPIs even for high-background baits via barcode enrichment analyses. After stringent filtering (based on effect size, significance, and replicate agreement) and retesting using the *HIS3* marker, the $Y2H_{GFP}$ screen identified 93 interactions among 13 viral and 84 human host proteins. We refer to this binary human SARS-CoV-2 interactome map as $HuSCI_{GFP}$ (Human-SARS-CoV-2 interactome via the $Y2H_{GFP}$ system), and refer to the union of $HuSCI_{HIS3}$ and $HuSCI_{GFP}$ as HuSCI (Supplementary Table 1B). The $Y2H_{GFP}$ screen also yielded 27 intraviral interactions amongst 19 viral proteins (here termed intraSCI; intraviral SARS-CoV-2 interactome) (Supplementary Table 1C).

Having collectively identified 205 direct virus-host and 27 intraviral interactions amongst 171 host and 19 viral proteins, we next wished to ensure that these candidate interactions were of sufficiently high quality.

Assessment of interaction data quality

As a first level of quality control, we examined the overlap of interactions among our datasets. Considering that ~60% of all viral/human protein pairs were assayed in both screens, and that each assay has an assay sensitivity (fraction of true PPIs that an assay is capable of detecting in a fully-saturated screen) of 20-25%, and given an estimated sampling sensitivity (the extent to which each screen is saturated) of 50% and 60% for the two screens, we could have expected ~3% overlap between the screens *a priori*, which is close to the observed 3.6% overlap (see Methods). Given these screening parameters, we estimate that our merged dataset has detected roughly 20% of all binary SARS-CoV-2 virus host interactions (see Methods), which is comparable to previous high quality binary PPI datasets (Luck et al., 2020; Rolland et al., 2014; Yu et al., 2008).

To experimentally assess interactome map quality, we deployed an established empirical framework based on validation with a well-calibrated orthogonal biochemical interaction assay (Choi et al., 2019; Luck et al., 2020; Rolland et al., 2014; Yu et al., 2008). Because any interaction assay can detect only a fraction of *bona fide* interactions (Braun et al., 2009), the fraction of pairs validated in an orthogonal assay (the 'validation rate') must be calibrated against a positive control set of well-documented interactions (Positive Reference Set, PRS) and a negative control set of randomly-selected protein pairs (Random Reference Set, RRS). We used the previously established human PRS/RRS version 2 (hsPRS-v2 and hsRRS-v2) sets as human positive and random reference sets for calibration (Braun et al., 2009; Choi et al., 2019). As another benchmark, we derived a collection of 55 human coronavirus/host protein interactions from the literature (Cusick et al., 2009) using the

criteria of being supported by multiple sources, of which at least one indicates a binary interaction. We refer to this benchmark as the virus-host literature binary multiple reference set (vhLit-BM; Supplementary Table 2A). Further, we established a virus-host Random Reference Set (vhRRS) of SARS-CoV-1 and 2 viral and host protein pairs by randomly selecting 180 protein pairs not previously reported as interactions.

We subjected HuSCI, IntraSCI, and each benchmark set of protein pairs to the orthogonal yeast-based N2H validation assay (yN2H; Fig. 1C and Supplementary Fig. 1) (Choi et al., 2019). We set stringent thresholds for the N2H assay such that only one vhRRS and no hsRRS-v2 pair scored positive. At this threshold, HuSCI alone and HuSCI merged with IntraSCI exhibited validation rates of 11 ± 2% and 10 ± 2%, respectively, which were statistically indistinguishable from the two positive control sets (hsPRS-v2 and vhLit-BM, 9 ± 2%; p = 0.88 and 5 ± 2%, p = 0.06; Fisher's exact tests relative to the union of HuSCI and IntraSCI), and significantly better than negative control set validation rates (hsRRS-v2, 0.4 ± 0.4%, $p = 4 \times 10^{-7}$ and and vhRRS 1.1 ± 0.6%; $p = 1 \times 10^{-7}$; Fisher's exact tests relative to the union of HuSCI and IntraSCI). IntraSCI contained too few interactions for a separate statistical evaluation. Thus, our virus-host interaction map (HuSCI, Supplementary Table 1B together with IntraSCI, Supplementary Table 1C) exhibited biophysical quality at least on par with human-human and virus-host benchmark sets of interactions that have been supported by multiple experiments in the manually curated literature.

Thus, we present our validated, high-quality network of 205 virus-host and 27 intraviral direct, binary interactions amongst 171 host and 19 viral proteins (Fig. 2A). The high quality of our dataset was further supported by the finding that host proteins detected as viral interaction partners in our screen were enriched with Gene Ontology terms related to viral infection (Fig. 2B), and by the observation that 62 of the here-identified viral interactors are known targets of other viruses ($p = 1 \times 10^{-4}$ by empirical testing)(Orchard et al., 2014) (Fig. 2C). Moreover, we find that a significant number of physically targeted proteins in HuSCI have been found to change phosphorylation status upon SARS-CoV-2 infection (p = 0.002 by empirical testing; Supplementary Fig. 3) (Bouhaddou et al., 2020; Stukalov et al., 2020). In addition, a previous set of 19 intraviral interactions (Li et al., 2020) showed a significant overlap of 5 interactions with our intraSCI network of 27 interactions ($p = 2 \times 10^{-4}$ by empirical testing; Fig. 2D). These observations together suggested that our map is of high biophysical quality, and enriched for host proteins that are relevant to the biology of SARS-CoV-2.

<u>Complementarity of Y2H and AP-MS datasets</u>

All previous maps of association between SARS-CoV2 and host proteins (Gordon et al., 2020a, 2020b; Li et al., 2020; Nabeel-Shah et al., 2020; Stukalov et al., 2020) have been generated using methods that cannot differentiate between direct and indirect interactions. By contrast, binary assays such as Y2H detect protein interactions that are predominantly direct (Luck et al., 2020; Rolland et al., 2014). This fundamental difference necessarily reduces the expected overlap between protein pairs identified in these assays (Yu et al., 2008). More surprising is the overall low agreement (~2%) observed between different pairings of these four AP-MS based association datasets, although this may be partially explained by differences in experimental setup, i.e. distinct cell lines and protocols (Braun, 2013).

Non-overlap might also be expected if the SARS-CoV-2 virus-host interactome contains many weak interactions that are less reliably detected. Interestingly, the validation rates of known virus-host interactions (vhLit-BM) and the random reference set of virus-host pairs (vhRRS) are less clearly separated than the corresponding positive and random human-human protein interaction benchmarks. Although it is possible that this phenomenon is due to imperfect quality of our literature-derived benchmark for coronaviral-host interactions, the yN2H data points to a higher prevalence of interactions amongst randomly chosen virus-host pairs (vhRRS), which may in turn stem from limited evolutionary selection against non-specific interactions. Previous systematic studies of interactions between yeast and human proteins observed an overall density of between-species interactions that was comparable to that of within-species interactions, despite the billion-year absence of selection to maintain yeast-human interactions (Zhong et al., 2016). This suggests that, in the evolution of protein interactions, selection against non-specific protein interactions may be as important as selection for specific protein interactions. A similar conclusion was reached by an earlier study on a single yeast SH3 binding peptide, which interacted very specifically with a single domain in its native proteome, but showed a broad interaction profile with SH3 domains from other organisms (Zarrinpar et al., 2003). Thus, the fact that SARS-CoV-2 has only recently jumped a species barrier is consistent with an increased abundance of less-reliably detected (presumably weaker) virus-host protein interactions.

Despite fundamental assay differences, low overlap between AP-MS datasets, and a potential tendency for SARS-CoV-2 to have weaker interactions with human host proteins, there were many overlaps between our network and protein pairs reported in other studies. Among our 205 HuSCI virus-host interactions, 11 (~5.4%) had been previously reported by at least one SARS-CoV-2 AP-MS study (Gordon et al., 2020a, 2020b; Li et al., 2020; Nabeel-Shah et al., 2020; Stukalov et al., 2020) (Supplementary Table 1B). Additionally, we found 29 human proteins in common with four previous AP-MS studies (Gordon et al., 2020a, 2020b; Li et al., 2020; Nabeel-Shah et al., 2020; Stukalov et al.,

2020), albeit targeted by other viral proteins than the previously-reported (potentially indirect) associations. The interactions most consistently identified in other datasets are between the viral N protein and host stress granule proteins G3BP1 and G3BP2. Previously, it has been shown that overexpression of the viral N protein stimulates stress granule formation (Gupta et al., 2017; Reineke and Lloyd, 2015). The fact that all interactome datasets identify the interaction of N with G3BP1 and G3BP2 (Nabeel-Shah et al., 2020; Samavarchi-Tehrani et al., 2020), strengthen the likelihood that this pathway is targeted. While differences were expected between the specific interactions and associations of different studies, we do find overall agreement at the functional level (see next section), in keeping with previous reports that Y2H and AP-MS reveal complementary but mutually consistent perspectives on the interactome.

Functions enriched in a SARS-CoV-2/human virhostome network

To identify which host functions are targeted by SARS-CoV-2, we performed GO enrichment analysis (see Methods). The terms most enriched amongst host proteins can be categorized more broadly into three functional groups: (i) immune regulation, (ii) ubiquitination and (iii) viral trafficking (Fig. 2A and B). Comparison with the four AP-MS based networks revealed substantial functional overlap, especially for terms related to viral trafficking, including vesicle-mediated transport to the plasma membrane and Golgi network (Supplementary Fig. 2A - D). Despite functional consistency between the different datasets, the divergence between the interactions we identified and the previously reported associations illustrates the complementarity of these methods and emphasizes the need for both approaches to understand SARS-CoV-2 biology. Moreover, in addition to confirming previously identified cellular processes, our network revealed novel functional enrichment that may shed light on viral modulation of immune regulation and virally-targeted ubiquitination processes.

The most obviously clinically relevant functions of viral targets relate to immune regulation. Infection with SARS-CoV-2 is often accompanied by dysregulation of interferon (IFN)-mediated antiviral response, as well as elevated inflammatory cytokine signaling (Blanco-Melo et al., 2020; Hadjadj et al., 2020; Karki et al., 2021). Dysregulation of IFN responses facilitates viral replication in the early stages of infection, while the so-called cytokine storm leads to acute respiratory distress syndrome in severe cases of COVID-19 (Fajgenbaum and June, 2020). Recent studies suggest that SARS-CoV-2 inhibits type I IFN induction by blocking host mRNA translation (Thoms et al., 2020), possibly by targeting TBK1 (Gordon et al., 2020a, 2020b). Additionally, the transcriptional activation of IFN-stimulated genes is inhibited by preventing STAT1 nuclear import by an ORF6-NUP98 interaction (Miorin et al., 2020).

Members of the tripartite motif (TRIM) protein family of E3 ligases, which act to ubiquitinate target proteins, have been implicated in regulating anti-viral host defenses and innate immune signaling (van Gent et al., 2018). We found several members of the TRIM protein family, namely TRIM2, 3, 27, 32, 50 and 54, to be interaction targets of viral NSP16 and NSP14. TRIM27 has been shown to modify TBK1 and IKBKG, thereby modulating the activation of IRF3/IRF7 (Zheng et al., 2015) and NF-κB (Zha et al., 2006), which, in turn, activate transcription of type I IFN and proinflammatory cytokines, respectively. Interestingly, TRIM27 was also found to be a degradation target of ICP0 during HSV-1 infection (Conwell et al., 2015). TRIM32 modulates innate immune responses in several ways. It negatively regulates TRIF and therefore TRL3/4 responses (Yang et al., 2017), as well as type I IFN production through ubiquitination of STING (Zhang et al., 2012). TRIM32 was further shown to inhibit influenza A replication by targeting its polymerase for proteasomal degradation (Fu et al., 2015). The direct NSP16-TRIM32 interaction we found was previously reported as an association (Gordon et al., 2020b). Additionally, we find an interaction between NSP14 and IKBKG, as well as multiple interactions between the NF-kB family member REL and NSP14, NSP16, and NSP9. Taken together, the identified interactions reveal direct routes by which SARS-CoV-2 may be targeting both the type I IFN pathway (and thereby antiviral host innate immune signaling) and inflammatory cytokine signaling.

A phenomenon observed in COVID-19 patients, for which the mechanism is unclear, is the functional exhaustion of cytotoxic lymphocytes during the adaptive immune response (Zheng et al., 2020). Here, we find that host proteins related to immune response are enriched amongst the partners of viral NSP6 ($p < 10^{-2}$ by empirical testing), as are membrane proteins ($p < 10^{-2}$ by empirical testing). Supporting the latter enrichment, NSP6 is itself a membrane protein that induces formation of viral double-membrane vesicles (Angelini et al., 2013). We find that NSP6 interacts with membrane regulators of cytotoxic lymphocytes responding to virus infection, including CD40 (Bennett et al., 1998), CD27 (Ochsenbein et al., 2004) and IL27RA (Wehrens et al., 2018). CD40's interaction with its ligand CD40L is important for T cell priming (Toes et al., 1998) and CD40 deficiency is known to inhibit T cell development against influenza virus (Lee et al., 2003). Use of CD27 agonists is exploited in cancer immunotherapy to co-stimulate the T-cell response (van de Ven and Borst, 2015) and CD27 expression is up-regulated in HIV-infected patients (Ochsenbein et al., 2004). IL27, the ligand of IL27RA, suppresses T cell cytotoxicity and viral control during cytomegalovirus infection (Wehrens et al., 2016). Thus, our map points to NSP6 as a potential regulator of T cell development and related COVID-19 symptoms. More broadly, our map is enriched for host proteins relevant to immune regulation and provides numerous mechanistic hypotheses.

The second functional category we identified is related to ubiquitination and highlights functional cross-talk between different targeted pathways. In addition to the ubiquitination-dependent modulation of host immune responses discussed above, we find ubiquitin-dependent degradation to be targeted by SARS-CoV-2. Hijacking the ubiquitin proteasome system is a common trait of almost all viruses (Banks et al., 2003; Fanunza et al., 2019; Tang et al., 2018). Supporting the idea that this is occurring for SARS-CoV-2, we found SARS-CoV-2 viral proteins to target host proteins involved in ubiquitin-mediated degradation. More specifically, we identified viral ORF3D, ORF6 and ORF9C as interacting with the host proteins UBQLN1/2. This is particularly interesting in light of the recent discovery that ORF9C attenuates antiviral response in lung epithelial cells in a proteasome-dependent manner (Andres et al., 2020). Human UBQLN1/2 mediates protein degradation of its interaction partners, e.g. hnRNPA1 (Gilpin et al., 2015) and TDP43 (Cassel and Reitz, 2013), by recruiting them to the 26S protease or autophagosome (Renaud et al., 2019).

A third functional category that was strongly represented among our interactors is related to viral trafficking via the ER-Golgi membrane network. Several similar terms were described by the physical association studies including: endomembrane system organization (Gordon et al., 2020a, 2020b), Golgi membrane (Stukalov et al., 2020), and Golgi vesicle transport (Li et al., 2020). The robust identification of these processes by all experimental approaches emphasize their importance in the viral life cycle. Of particular interest was NSP16, which is known to methylate mRNA to facilitate viral replication and escape from innate immune recognition (Wang et al., 2015), but which exhibits many interactors related to viral transport. In addition, we found NSP16 interactors RAB3IL1 (Wandinger-Ness and Zerial, 2014), VPS52 (Conibear et al., 2003), COG6 (Blackburn et al., 2019), and EXO6 (Boehm et al., 2017), which are important factors of transport into the extracellular space via ER-Golgi network, and are closely related to viral trafficking (Sicari et al., 2020). Therefore, in addition to its functions in viral replication, analysis of interactions suggests a role for NSP16 in virion production and release.

Pairwise test of published SARS-CoV-1 interactions supports conserved immune regulation mechanism

Because interactions are often conserved between species (the ‘interolog’ phenomenon (Matthews et al., 2001; Walhout et al., 2000; Yu et al., 2004)), we sought to cast a wider net by carrying out pairwise assessment of SARS-CoV-2 viral-human pairs corresponding to interactions previously reported for SARS-CoV-1. Of the 76 known SARS-CoV-1 binary virus-host PPIs that had been reported by Y2H, 10% (6 out of 62 PPIs, excluding 14 which are not amenable to testing given due to viral bait autoactivation) were recapitulated in our pairwise yN2H assay when substituting SARS-CoV-2 proteins

for the originally reported SARS-CoV-1 proteins. The identified interologs included NSP3-MKRN3, NSP8-H2AFY2, NSP8-TERF1, NSP9-FAHD1, NSP13-N4BP2L2, and ORF7B-CAMLG. Interestingly, two of these interactions – NSP8-TERF1 and ORF7B-CAMLG, were also detected in our HuSCI network. Furthermore, the majority of the host proteins in the identified interologs are associated with the functional pathways that were found to be enriched in our network analysis (see Methods). MKRN3 (Kanber et al., 2009) and CAMLG (Peng et al., 2010) are proteins involved in protein ubiquitination. TERF1 (de Lange, 2005), H2AFY2 (Zhang et al., 2005), and N4BP2L2 (Salipante et al., 2009) are involved in various pathways regulating transcriptional mechanisms. Of particular note is the interaction between viral NSP13 and human N4BP2L2: NSP13 is a highly conserved helicase (Chan et al., 2020) that has been associated with suppression of IFN production and signalling (Yuen et al., 2020). Also, human N4BP2L2 is known to interact with both neutrophil elastase and the transcriptional repressor GFI1 to modulate the production of neutrophils - a type of white blood cells essential to host innate immunity (Salipante et al., 2009). Interestingly, NSP13 has been suggested to act as a transcriptional regulator, adding to its importance in viral replication (Gordon et al., 2020a, 2020b). Taken together, a closer investigation of the direct interaction between viral NSP13 and host N4BP2L2 proteins may be of importance to understanding the mechanisms behind an 'under siege' host immune response following SARS-CoV-2 viral infection.

Limitations and further directions

All protein interaction assays have limitations intrinsic to each method. Y2H assays are limited by the fact that proteins are exogenously expressed with functional assay tags and targeted to the nucleus. The heterologous nature of the assays and circumvention of physiological transcriptional regulation are a limitation, but also a benefit relative to screens in a given cell line which will miss interactions with proteins not expressed in that cell, but which may be critical targets in other tissues (Hikmet et al., 2020). Although it has been demonstrated repeatedly that Y2H systems, when quality controlled by orthogonal validation with empirically benchmarked assays as done here, yield high-quality interactions that enable important and robust biological insights (Braun et al., 2009; Choi et al., 2019; Luck et al., 2020; Rolland et al., 2014; Venkatesan et al., 2009; Yu et al., 2008), there is no guarantee that every detected biophysical interaction will be physiologically relevant.

A known issue with every carefully conducted interaction assay is that true interactions can be missed, with only 20-40% of reference interactions being detectable by any single assay (Braun et al., 2009; Choi et al., 2019; Luck et al., 2020; Yu et al., 2008). This was a major motivation for applying

complementary parallel approaches in this study (Choi et al., 2019). Future efforts might expand the barcoded ORFeome approach to distinct assay versions, including alternative linkers and orientations for fused bait and prey tags. In addition, assay sensitivity can be increased by implementing the same assay version in different conditions (Kim et al., 2021; Liu et al., 2020). More physiologically relevant protein interactions might also be discovered by conducting our screen in the presence of a third viral protein which might be necessary for mediating other direct interactions. Our genome-wide viral and human barcoded ORFeomes are readily transferable to new genetic and environmental backgrounds.

One of the main advantages of Y2H, unlike protein associations identified by AP-MS, is that it reports direct interactions (Luck et al., 2020; Rolland et al., 2014). None of the viral-host protein interaction maps generated so far can robustly distinguish between direct and indirect interactions. Comprehensive knowledge of direct interactions is necessary for an accurate mechanistic understanding of viral infection and progression for potential specific and efficient interventions. Another future extension of the $Y2H_{GFP}$ method could exploit its ability to measure many protein pairs in a single pool for rapid screening of potential interaction-disrupting drugs against hundreds of virus-host protein interactions, such as those identified here. Taken together, we expect that combining our map of direct SARS-CoV-2 protein interactions with previously established protein physical association and proximity networks will broaden our mechanistic understanding of viral proliferation and enable rapid development of therapeutic approaches to combat current and future pandemics.

FIGURE LEGENDS

Figure 1. Generation and quality assessment of proteome-wide direct interaction maps amongst SARS-CoV-2 and human host proteins. (A) Schematic description of the experimental pipeline. **(B)** Screening space for each of the two parallel $Y2H_{HIS}$ and $Y2H_{GFP}$ screens. "Percent overlap" is relative to the union of protein pairs tested by both methods. **(C)** Orthogonal N2H assay validation of HuSCI and IntraSCI along with positive (hsPRS-v2 and vhLit-BM) and negative (hsRRS-v2 and vhRRS) benchmarking sets. Asterisks indicate significant differences from viral negative benchmarking set ($p < 0.05$, fisher's exact test).

Figure 2. Network representation and functional assessment of proteome-wide direct interaction map between SARS-CoV-2 and human host proteins (A) Host proteins grouped by the most enriched broad functional terms, including immune regulation, ubiquitination and viral trafficking. **(B)** Functions enriched amongst the host proteins found in HuSCI. A great majority of the functional annotations are novel, with only one term (cytoplasmic stress granule) enriched amongst the host proteins of previous AP-MS based maps. **(C)** Comparison of previously known intraviral interactions and IntraSCI. **(D)** Comparison of previously known SARS-CoV-2 targets with those identified in HuSCI.

Supplementary Figure 1. Rate of detection in an orthogonal N2H assay at different scoring thresholds. Rare at which interactions are detected by yN2H for HuSCI and IntraSCI, as well as positive (hsPRS-v2 and vhLit-BM) and negative (hsRRS-v2 and vhRRS) benchmark sets, across stringency thresholds.

Supplementary Figure 2. Functional enrichment in the host proteins of four AP-MS based networks. (A) Gordon *et al*. (Gordon et al., 2020a, 2020b) **(B)** Stukalov, *et al*. (Stukalov et al., 2020) **(C)** Li, *et al. (Li et al., 2020)* **(D)** Nabeel, *et al*. (Nabeel-Shah et al., 2020)

Supplementary Figure 3 - related to Figure 2. Comparison of phosphoregulated proteins by viral infection and host targets in SCI.

STAR METHODS

- **KEY RESOURCES TABLE**
- **LEAD CONTACT AND MATERIALS AVAILABILITY**
- **EXPERIMENTAL MODEL AND SUBJECT DETAILS**
- **METHOD DETAILS**
    - Selection & clone preparation of functional coding regions from SARS-CoV-2
    - Screening based on His3 to generate the $HuSCI_{HIS3}$ map
    - Screening based on GFP to generate the $HuSCI_{GFP}$ map
    - Validation by an orthogonal yN2H assay
    - Bioinformatic analysis of binary interactome map
    - Identification of interologs via pairwise Y2H testing
- **DATA AND CODE AVAILABILITY**
    - The read counting based on the expected barcode pipeline is available at https://github.com/RyogaLi/BFG_Y2H/.
- **QUANTIFICATION AND STATISTICAL ANALYSIS**

TABLE LEGENDS

Supplementary Table 1A. Total list of viral proteins used in both Y2H screens. Annotations are given for the putative function of individual proteins, as well as amino acid sequences.

Supplementary Table 1B. Total list of PPIs between viral and host proteins in HuSCI. The screen in which the interaction was found ($HuSCI_{HIS3}$ and/or $HuSCI_{GFP}$) is specified, as well as interactions also found by the four AP-MS studies (Gordon et al., 2020a, 2020b; Li et al., 2020; Nabeel-Shah et al., 2020; Stukalov et al., 2020)

Supplementary Table 1C. Total list of PPIs amongst viral proteins (IntraSCI) identified by $Y2H_{GFP}$. The overlap with a previous intra-viral PPI study is also indicated (Li et al., 2020)

Supplementary Table 2A. List of PPIs in virus-host literature binary multiple reference set (vhLit-BM). The number of methods by which the interaction was identified is indicated.

Supplementary Table 2B. List of protein pairs in viral Random Reference Set (vhRRS).

Supplementary Table 3A. Total list of viral clones in AD-Nterm-Cen Y2H destination vector for $Y2H_{GFP}$.

Supplementary Table 3B. Total list of viral clones in DB-Nterm-Cen Y2H destination vector for $Y2H_{GFP}$.

Supplementary Table 4. A curated list of previously identified binary interactions between SARS-CoV-1 and human host proteins.

Supplementary Table 5. Genotypes of toolkit strains used in $Y2H_{GFP}$.

ACKNOWLEDGEMENTS

This work was supported by Canadian Institutes for Health Research Foundation Grant (F.P.R.); the Canada Excellence Research Chairs Program (F.P.R.); the Thistledown Foundation (F.P.R.); the LabEx Integrative Biology of Emerging Infectious Diseases (10-LABX-0062; Y.J.) and Platform for European Preparedness Against (Re-)emerging Epidemics, EU (602525; Y.J.); the European Union's Horizon 2020 Research and Innovation Programme (Project ID 101003633, RiPCoN; P.FB, C.B., P.A.). F.L. was supported by a Belgian American Educational Foundation Doctoral Research Fellowship and a Wallonia-Brussels International (WBI)-World Excellence Fellowship. M.V. is a Chercheur Qualifié Honoraire from the Fonds de la Recherche Scientifique (FRS-FNRS, Wallonia-Brussels Federation, Belgium).

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

Altmann, M., Altmann, S., Falter, C., and Falter-Braun, P. (2018). High-quality yeast-2-hybrid interaction network mapping. Curr. Protoc. Plant Biol. *3*, e20067.

Altmann, M., Altmann, S., Rodriguez, P.A., Weller, B., Elorduy Vergara, L., Palme, J., Marín-de la Rosa, N., Sauer, M., Wenig, M., Villaécija-Aguilar, J.A., et al. (2020). Extensive signal integration by the phytohormone protein network. Nature *583*, 271–276.

Andres, A.D., Feng, Y., Campos, A.R., Yin, J., Yang, C.-C., James, B., Murad, R., Kim, H., Deshpande, A.J., Gordon, D.E., et al. (2020). SARS-CoV-2 ORF9c Is a membrane-associated protein that suppresses antiviral responses in cells. bioRxiv.

Angelini, M.M., Akhlaghpour, M., Neuman, B.W., and Buchmeier, M.J. (2013). Severe acute respiratory syndrome coronavirus nonstructural proteins 3, 4, and 6 induce double-membrane vesicles. mBio *4*. e00524-13

Banks, L., Pim, D., and Thomas, M. (2003). Viruses and the 26S proteasome: hacking into destruction. Trends Biochem. Sci. *28*, 452–459.

Bennett, S.R., Carbone, F.R., Karamalis, F., Flavell, R.A., Miller, J.F., and Heath, W.R. (1998). Help for cytotoxic-T-cell responses is mediated by CD40 signalling. Nature *393*, 478–480.

Blackburn, J.B., D'Souza, Z., and Lupashin, V.V. (2019). Maintaining order: COG complex controls golgi trafficking, processing, and sorting. FEBS Lett. *593*, 2466–2487.

Blanco-Melo, D., Nilsson-Payant, B.E., Liu, W.-C., Uhl, S., Hoagland, D., Møller, R., Jordan, T.X., Oishi, K., Panis, M., Sachs, D., et al. (2020). Imbalanced host response to SARS-CoV-2 drives development of COVID-19. Cell *181*, 1036–1045.e9.

Boehm, C.M., Obado, S., Gadelha, C., Kaupisch, A., Manna, P.T., Gould, G.W., Munson, M., Chait, B.T., Rout, M.P., and Field, M.C. (2017). The trypanosome exocyst: a conserved structure revealing a new role in endocytosis. PLoS Pathog. *13*, e1006063.

Bouhaddou, M., Memon, D., Meyer, B., White, K.M., Rezelj, V.V., Correa Marrero, M., Polacco, B.J., Melnyk, J.E., Ulferts, S., Kaake, R.M., et al. (2020). The global phosphorylation landscape of SARS-CoV-2 infection. Cell *182*, 685–712.e19.

Braun, P. (2013). Reproducibility restored--on toward the human interactome. Nat. Methods *10*, 301, 303.

Braun, P., Tasan, M., Dreze, M., Barrios-Rodiles, M., Lemmens, I., Yu, H., Sahalie, J.M., Murray, R.R., Roncari, L., de Smet, A.-S., et al. (2009). An experimentally derived confidence score for binary protein-protein interactions. Nat. Methods *6*, 91–97.

Calderwood, M.A., Venkatesan, K., Xing, L., Chase, M.R., Vazquez, A., Holthaus, A.M., Ewence, A.E., Li, N., Hirozane-Kishikawa, T., Hill, D.E., et al. (2007). Epstein-Barr virus and virus human protein interaction maps. Proc. Natl. Acad. Sci. U. S. A. *104*, 7606–7611.

Carfì, A., Bernabei, R., Landi, F., and Gemelli Against COVID-19 Post-Acute Care Study Group (2020).

Persistent symptoms in patients after acute COVID-19. JAMA *324*, 603–605.

Cassel, J.A., and Reitz, A.B. (2013). Ubiquilin-2 (UBQLN2) binds with high affinity to the C-terminal region of TDP-43 and modulates TDP-43 levels in H4 cells: Characterization of inhibition by nucleic acids and 4-aminoquinolines. Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics *1834*, 964–971.

Chan, J.F.-W., Kok, K.-H., Zhu, Z., Chu, H., To, K.K.-W., Yuan, S., and Yuen, K.-Y. (2020). Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan. Emerging Microbes & Infections *9*, 221–236.

Choi, S.G., Olivet, J., Cassonnet, P., Vidalain, P.-O., Luck, K., Lambourne, L., Spirohn, K., Lemmens, I., Dos Santos, M., Demeret, C., et al. (2019). Maximizing binary interactome mapping with a minimal number of assays. Nature Communications *10*, 3907.

Conibear, E., Cleck, J.N., and Stevens, T.H. (2003). Vps51p mediates the association of the GARP (Vps52/53/54) complex with the late Golgi t-SNARE Tlg1p. Mol. Biol. Cell *14*, 1610–1623.

Conwell, S.E., White, A.E., Harper, J.W., and Knipe, D.M. (2015). Identification of TRIM27 as a novel degradation target of herpes simplex virus 1 ICP0. J. Virol. *89*, 220–229.

Cusick, M.E., Yu, H., Smolyar, A., Venkatesan, K., Carvunis, A.-R., Simonis, N., Rual, J.-F., Borick, H., Braun, P., Dreze, M., et al. (2009). Literature-curated protein interaction datasets. Nat. Methods *6*, 39–46.

Fajgenbaum, D.C., and June, C.H. (2020). Cytokine storm. N. Engl. J. Med. *383*, 2255–2273.

Fanunza, E., Frau, A., Corona, A., and Tramontano, E. (2019). Insights into ebola virus VP35 and VP24 interferon inhibitory functions and their initial exploitation as drug targets. Infect. Disord. Drug Targets *19*, 362–374.

Fields, S., and Song, O. (1989). A novel genetic system to detect protein-protein interactions. Nature *340*, 245–246.

Fu, B., Wang, L., Ding, H., Schwamborn, J.C., Li, S., and Dorf, M.E. (2015). TRIM32 Senses and restricts influenza a virus by ubiquitination of PB1 polymerase. PLoS Pathog. *11*, e1004960.

van Gent, M., Sparrer, K.M.J., and Gack, M.U. (2018). TRIM proteins and their roles in antiviral host defenses. Annu Rev Virol *5*, 385–405.

Gibson, D.G., Young, L., Chuang, R.-Y., Venter, J.C., Hutchison, C.A., 3rd, and Smith, H.O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat. Methods *6*, 343–345.

Gilpin, K.M., Chang, L., and Monteiro, M.J. (2015). ALS-linked mutations in ubiquilin-2 or hnRNPA1 reduce interaction between ubiquilin-2 and hnRNPA1. Hum. Mol. Genet. *24*, 2565–2577.

Gordon, D.E., Hiatt, J., Bouhaddou, M., Rezelj, V.V., Ulferts, S., Braberg, H., Jureka, A.S., Obernier, K., Guo, J.Z., Batra, J., et al. (2020b). Comparative host-coronavirus protein interaction networks reveal pan-viral disease mechanisms. Science *370*, 1181.

Gordon, D.E., Jang, G.M., Bouhaddou, M., Xu, J., Obernier, K., White, K.M., O'Meara, M.J., Rezelj, V.V.,

Guo, J.Z., Swaney, D.L., et al. (2020a). A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. Nature *583*, 459–468.

Grubaugh, N.D., Hodcroft, E.B., Fauver, J.R., Phelan, A.L., and Cevik, M. (2021). Public health actions to control new SARS-CoV-2 variants. Cell 184, 1127-1132.

Gulbahce, N., Yan, H., Dricot, A., Padi, M., Byrdsong, D., Franchi, R., Lee, D.-S., Rozenblatt-Rosen, O., Mar, J.C., Calderwood, M.A., et al. (2012). Viral perturbations of host networks reflect disease etiology. PLoS Comput. Biol. *8*, e1002531.

Gupta, N., Badeaux, M., Liu, Y., Naxerova, K., Sgroi, D., Munn, L.L., Jain, R.K., and Garkavtsev, I. (2017). Stress granule-associated protein G3BP2 regulates breast tumor initiation. Proc. Natl. Acad. Sci. U. S. A. *114*, 1033–1038.

Hadjadj, J., Yatim, N., Barnabei, L., Corneau, A., Boussier, J., Smith, N., Péré, H., Charbit, B., Bondet, V., Chenevier-Gobeaux, C., et al. (2020). Impaired type I interferon activity and inflammatory responses in severe COVID-19 patients. Science *369*, 718–724.

Hikmet, F., Méar, L., Edvinsson, Å., Micke, P., Uhlén, M., and Lindskog, C. (2020). The protein expression profile of ACE2 in human tissues. Mol. Syst. Biol. *16*, e9610.

Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., Zhang, L., Fan, G., Xu, J., Gu, X., et al. (2020). Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet *395*, 497–506.

Jungreis, I., Nelson, C.W., Ardern, Z., Finkel, Y., Krogan, N.J., Sato, K., Ziebuhr, J., Stern-Ginossar, N., Pavesi, A., Firth, A.E., et al. (2020). Conflicting and ambiguous names of overlapping ORFs in SARS-CoV-2: A homology-based resolution. Preprint

Kanber, D., Giltay, J., Wieczorek, D., Zogel, C., Hochstenbach, R., Caliebe, A., Kuechler, A., Horsthemke, B., and Buiting, K. (2009). A paternal deletion of MKRN3, MAGEL2 and NDN does not result in Prader–Willi syndrome. Eur. J. Hum. Genet. *17*, 582–590.

Karki, R., Sharma, B.R., Tuladhar, S., Williams, E.P., Zalduondo, L., Samir, P., Zheng, M., Sundaram, B., Banoth, B., Malireddi, R.K.S., et al. (2021). Synergism of TNF-α and IFN-γ triggers inflammatory cell death, tissue damage, and mortality in SARS-CoV-2 infection and cytokine shock syndromes. Cell *184*, 149–168.e17.

Kim, D.-K., Knapp, J.J., Kuang, D., Chawla, A., Cassonnet, P., Lee, H., Sheykhkarimli, D., Samavarchi-Tehrani, P., Abdouni, H., Rayhan, A., et al. (2020). A comprehensive, flexible collection of SARS-CoV-2 coding regions. G3 *10*, 3399–3402.

Kim, D.-K., Sheykhkarimli, D., Kishore, N., Rayhan, A., Kuang, D., Li, R., Li, S., Škalič, M., Colobella, C., Cote, A.G., et al. (2021). Global dynamics of environment-dependent interactomes. Unpublished.

Kingma, D.P., and Ba, J. (2014). Adam: A method for stochastic optimization. arXiv.

de Lange, T. (2005). Shelterin: the protein complex that shapes and safeguards human telomeres. Genes Dev. *19*, 2100–2110.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*,

357–359.

Laurent, E.M.N., Sofianatos, Y., Komarova, A., Gimeno, J.-P., Tehrani, P.S., Kim, D.-K., Abdouni, H., Duhamel, M., Cassonnet, P., Knapp, J.J., et al. (2020). 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.

Lee, B.O., Hartson, L., and Randall, T.D. (2003). CD40-deficient, influenza-specific CD8 memory T cells develop and function normally in a CD40-sufficient environment. J. Exp. Med. *198*, 1759–1764.

Li, J., Guo, M., Tian, X., Wang, X., Yang, X., Wu, P., Liu, C., Xiao, Z., Qu, Y., Yin, Y., et al. (2020). Virus-host interactome and proteomic survey reveal potential virulence factors influencing SARS-CoV-2 pathogenesis. Med (N Y).

Liu, Z., Miller, D., Li, F., Liu, X., and Levy, S.F. (2020). A large accessory protein interactome is rewired across environments. Elife *9*.

Lu, H., Zhou, Q., He, J., Jiang, Z., Peng, C., Tong, R., and Shi, J. (2020). Recent advances in the development of protein-protein interactions modulators: mechanisms and clinical trials. Signal Transduct Target Ther *5*, 213.

Luck, K., Kim, D.-K., Lambourne, L., Spirohn, K., Begg, B.E., Bian, W., Brignall, R., Cafarelli, T., Campos-Laborie, F.J., Charloteaux, B., et al. (2020). A reference map of the human binary protein interactome. Nature *580*, 402–408.

Maas, A.L., Hannun, A.Y., and Ng, A.Y. (2013). Rectifier nonlinearities improve neural network acoustic models. In Proc. Icml, (Citeseer), p. 3.

Matthews, L.R., Vaglio, P., Reboul, J., Ge, H., Davis, B.P., Garrels, J., Vincent, S., and Vidal, M. (2001). Identification of potential interaction networks using sequence-based searches for conserved protein-protein interactions or “interologs.” Genome Res. *11*, 2120–2126.

Miorin, L., Kehrer, T., Sanchez-Aparicio, M.T., Zhang, K., Cohen, P., Patel, R.S., Cupic, A., Makio, T., Mei, M., Moreno, E., et al. (2020). SARS-CoV-2 Orf6 hijacks Nup98 to block STAT nuclear import and antagonize interferon signaling. Proc. Natl. Acad. Sci. U. S. A. *117*, 28344–28354.

Muller, M., Jacob, Y., Jones, L., Weiss, A., Brino, L., Chantier, T., Lotteau, V., Favre, M., and Demeret, C. (2012). Large scale genotype comparison of human papillomavirus E2-host interaction networks provides new insights for e2 molecular functions. PLoS Pathog. *8*, e1002761.

Nabeel-Shah, S., Lee, H., Ahmed, N., and Marcon, E. (2020). SARS-CoV-2 Nucleocapsid protein attenuates stress granule formation and alters gene expression via direct interaction with host mRNAs. bioRxiv.

Neveu, G., Cassonnet, P., Vidalain, P.-O., Rolloy, C., Mendoza, J., Jones, L., Tangy, F., Muller, M., Demeret, C., Tafforeau, L., et al. (2012). Comparative analysis of virus-host interactomes with a mammalian high-throughput protein complementation assay based on Gaussia princeps luciferase. Methods *58*, 349–359.

Ochsenbein, A.F., Riddell, S.R., Brown, M., Corey, L., Baerlocher, G.M., Lansdorp, P.M., and Greenberg, P.D. (2004). CD27 expression promotes long-term survival of functional effector–memory CD8+cytotoxic

T lymphocytes in HIV-infected patients. J. Exp. Med. *200*, 1407–1417.

Orchard, S., Ammari, M., Aranda, B., Breuza, L., Briganti, L., Broackes-Carter, F., Campbell, N.H., Chavali, G., Chen, C., del-Toro, N., et al. (2014). The MIntAct project--IntAct as a common curation platform for 11 molecular interaction databases. Nucleic Acids Res. *42*, D358–D363.

Peng, Z., Shi, T., and Ma, D. (2010). RNF122: a novel ubiquitin ligase associated with calcium-modulating cyclophilin ligand. BMC Cell Biol. *11*, 41.

Pfefferle, S., Schöpf, J., Kögl, M., Friedel, C.C., Müller, M.A., Carbajo-Lozoya, J., Stellberger, T., von Dall'Armi, E., Herzog, P., Kallies, S., et al. (2011). The SARS-coronavirus-host interactome: identification of cyclophilins as target for pan-coronavirus inhibitors. PLoS Pathog. *7*, e1002331.

Reineke, L.C., and Lloyd, R.E. (2015). The stress granule protein G3BP1 recruits protein kinase R to promote multiple innate immune antiviral responses. J. Virol. *89*, 2575–2589.

Renaud, L., Picher-Martel, V., Codron, P., and Julien, J.-P. (2019). Key role of UBQLN2 in pathogenesis of amyotrophic lateral sclerosis and frontotemporal dementia. Acta Neuropathol Commun *7*, 103.

Rolland, T., Taşan, M., Charloteaux, B., Pevzner, S.J., Zhong, Q., Sahni, N., Yi, S., Lemmens, I., Fontanillo, C., Mosca, R., et al. (2014). A proteome-scale map of the human interactome network. Cell *159*, 1212–1226.

Rual, J.-F., Venkatesan, K., Hao, T., Hirozane-Kishikawa, T., Dricot, A., Li, N., Berriz, G.F., Gibbons, F.D., Dreze, M., Ayivi-Guedehoussou, N., et al. (2005). Towards a proteome-scale map of the human protein–protein interaction network. Nature *437*, 1173–1178.

Salipante, S.J., Rojas, M.E.B., Korkmaz, B., Duan, Z., Wechsler, J., Benson, K.F., Person, R.E., Grimes, H.L., and Horwitz, M.S. (2009). Contributions to neutropenia from PFAAP5 (N4BP2L2), a novel protein mediating transcriptional repressor cooperation between Gfi1 and neutrophil elastase. Mol. Cell. Biol. *29*, 4394–4405.

Samavarchi-Tehrani, P., Abdouni, H., Knight, J.D.R., Astori, A., Samson, R., Lin, Z.-Y., Kim, D.-K., Knapp, J.J., St-Germain, J., Go, C.D., et al. (2020). A SARS-CoV-2 – host proximity interactome. bioRxiv.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. *13*, 2498–2504.

Sicari, D., Chatziioannou, A., Koutsandreas, T., Sitia, R., and Chevet, E. (2020). Role of the early secretory pathway in SARS-CoV-2 infection. J. Cell Biol. *219*.

St-Germain, J.R., Astori, A., Samavarchi-Tehrani, P., Abdouni, H., Macwan, V., Kim, D.-K., Knapp, J.J., Roth, F.P., Gingras, A.-C., and Raught, B. (2020). A SARS-CoV-2 BioID-based virus-host membrane protein interactome and virus peptide compendium: new proteomics resources for COVID-19 research. bioRxiv.

Stukalov, A., Girault, V., Grass, V., Bergant, V., Karayel, O., Urban, C., Haas, D.A., Huang, Y., Oubraham, L., Wang, A., et al. (2020). Multi-level proteomics reveals host-perturbation strategies of

SARS-CoV-2 and SARS-CoV. bioRxiv.

Tang, Q., Wu, P., Chen, H., and Li, G. (2018). Pleiotropic roles of the ubiquitin-proteasome system during viral propagation. Life Sci. *207*, 350–354.

The ORFeome Collaboration (2016). The ORFeome collaboration: a genome-scale human ORF-clone resource. Nature Methods *13*, 191–192.

Thoms, M., Buschauer, R., Ameismeier, M., Koepke, L., Denk, T., Hirschenberger, M., Kratzat, H., Hayn, M., Mackens-Kiani, T., Cheng, J., et al. (2020). Structural basis for translational shutdown and immune evasion by the Nsp1 protein of SARS-CoV-2. Science *369*, 1249–1255.

Toes, R.E.M., Schoenberger, S.P., van der Voort, E.I.H., Offringa, R., and Melief, C.J.M. (1998). CD40–CD40 ligand interactions and their role in cytotoxic T lymphocyte priming and anti-tumor immunity. Semin. Immunol. *10*, 443–448.

Venkatesan, K., Rual, J.-F., Vazquez, A., Stelzl, U., Lemmens, I., Hirozane-Kishikawa, T., Hao, T., Zenkner, M., Xin, X., Goh, K.-I., et al. (2009). An empirical framework for binary interactome mapping. Nat. Methods *6*, 83–90.

van de Ven, K., and Borst, J. (2015). Targeting the T-cell co-stimulatory CD27/CD70 pathway in cancer immunotherapy: rationale and potential. Immunotherapy *7*, 655–667.

Walhout, A.J., Sordella, R., Lu, X., Hartley, J.L., Temple, G.F., Brasch, M.A., Thierry-Mieg, N., and Vidal, M. (2000). Protein interaction mapping in C. elegans using proteins involved in vulval development. Science *287*, 116–122.

Wandinger-Ness, A., and Zerial, M. (2014). Rab proteins and the compartmentalization of the endosomal system. Cold Spring Harbor Perspectives in Biology *6*, a022616–a022616.

Wang, Y., Sun, Y., Wu, A., Xu, S., Pan, R., Zeng, C., Jin, X., Ge, X., Shi, Z., Ahola, T., et al. (2015). Coronavirus nsp10/nsp16 methyltransferase can be targeted by nsp10-derived peptide in vitro and in vivo to reduce replication and pathogenesis. J. Virol. *89*, 8416–8427.

Wehrens, E.J., Wong, K., Gupta, A., Benedict, C., and Zuniga, E. (2016). IL-27 suppresses CD4 and CD8 T cell cytotoxicity and viral control during cytomegalovirus infection. The Journal of Immunology *196*, 217.4–217.4.

Wehrens, E.J., Wong, K.A., Gupta, A., Khan, A., Benedict, C.A., and Zuniga, E.I. (2018). IL-27 regulates the number, function and cytotoxic program of antiviral CD4 T cells and promotes cytomegalovirus persistence. PLoS One *13*, e0201249.

Weile, J., Sun, S., Cote, A.G., Knapp, J., Verby, M., Mellor, J.C., Wu, Y., Pons, C., Wong, C., van Lieshout, N., et al. (2017). A framework for exhaustively mapping functional missense variants. Mol. Syst. Biol. *13*, 957.

Wu, A., Peng, Y., Huang, B., Ding, X., Wang, X., Niu, P., Meng, J., Zhu, Z., Zhang, Z., Wang, J., et al. (2020a). Genome composition and divergence of the novel coronavirus (2019-nCoV) originating in china. Cell Host & Microbe *27*, 325–328.

Wu, A., Peng, Y., Huang, B., Ding, X., Wang, X., Niu, P., Meng, J., Zhu, Z., Zhang, Z., Wang, J., et al.

(2020b). Genome composition and divergence of the novel coronavirus (2019-nCoV) originating in China. Cell Host & Microbe *27*, 325–328.

Wu, F., Zhao, S., Yu, B., Chen, Y.-M., Wang, W., Song, Z.-G., Hu, Y., Tao, Z.-W., Tian, J.-H., Pei, Y.-Y., et al. (2020c). A new coronavirus associated with human respiratory disease in China. Nature *579*, 265–269.

Yachie, N., Petsalaki, E., Mellor, J.C., Weile, J., Jacob, Y., Verby, M., Ozturk, S.B., Li, S., Cote, A.G., Mosca, R., et al. (2016). Pooled-matrix protein interaction screens using Barcode Fusion Genetics. Mol. Syst. Biol. *12*, 863.

Yang, Q., Liu, T.-T., Lin, H., Zhang, M., Wei, J., Luo, W.-W., Hu, Y.-H., Zhong, B., Hu, M.-M., and Shu, H.-B. (2017). TRIM32-TAX1BP1-dependent selective autophagic degradation of TRIF negatively regulates TLR3/4-mediated innate immune responses. PLoS Pathog. *13*, e1006600.

Yu, F., and Koltun, V. (2015). Multi-scale context aggregation by dilated convolutions. arXiv.

Yu, H., Luscombe, N.M., Lu, H.X., Zhu, X., Xia, Y., Han, J.-D.J., Bertin, N., Chung, S., Vidal, M., and Gerstein, M. (2004). Annotation transfer between genomes: protein-protein interologs and protein-DNA regulogs. Genome Res. *14*, 1107–1118.

Yu, H., Braun, P., Yildirim, M.A., Lemmens, I., Venkatesan, K., Sahalie, J., Hirozane-Kishikawa, T., Gebreab, F., Li, N., Simonis, N., et al. (2008). High-quality binary protein interaction map of the yeast interactome network. Science *322*, 104–110.

Yuen, C.-K., Lam, J.-Y., Wong, W.-M., Mak, L.-F., Wang, X., Chu, H., Cai, J.-P., Jin, D.-Y., To, K.K.-W., Chan, J.F.-W., et al. (2020). SARS-CoV-2 nsp13, nsp14, nsp15 and orf6 function as potent interferon antagonists. Emerg. Microbes Infect. *9*, 1418–1428.

Zarrinpar, A., Park, S.-H., and Lim, W.A. (2003). Optimization of specificity in a cellular protein interaction network by negative selection. Nature *426*, 676–680.

Zha, J., Han, K.-J., Xu, L.-G., He, W., Zhou, Q., Chen, D., Zhai, Z., and Shu, H.-B. (2006). The Ret finger protein inhibits signaling mediated by the noncanonical and canonical IkappaB kinase family members. J. Immunol. *176*, 1072–1080.

Zhang, J., Hu, M.-M., Wang, Y.-Y., and Shu, H.-B. (2012). TRIM32 protein modulates type I interferon induction and cellular antiviral response by targeting MITA/STING protein for K63-linked ubiquitination. J. Biol. Chem. *287*, 28646–28655.

Zhang, R., Poustovoitov, M.V., Ye, X., Santos, H.A., Chen, W., Daganzo, S.M., Erzberger, J.P., Serebriiskii, I.G., Canutescu, A.A., Dunbrack, R.L., et al. (2005). Formation of MacroH2A-containing senescence-associated heterochromatin foci and senescence driven by ASF1a and HIRA. Dev. Cell *8*, 19–30.

Zheng, M., Gao, Y., Wang, G., Song, G., Liu, S., Sun, D., Xu, Y., and Tian, Z. (2020). Functional exhaustion of antiviral lymphocytes in COVID-19 patients. Cell. Mol. Immunol. *17*, 533–535.

Zheng, Q., Hou, J., Zhou, Y., Yang, Y., Xie, B., and Cao, X. (2015). Siglec1 suppresses antiviral innate immune response by inducing TBK1 degradation via the ubiquitin ligase TRIM27. Cell Res. *25*,

1121–1136.

Zhong, Q., Pevzner, S.J., Hao, T., Wang, Y., Mosca, R., Menche, J., Taipale, M., Taşan, M., Fan, C., Yang, X., et al. (2016). An inter-species protein-protein interaction network across vast evolutionary distance. Mol. Syst. Biol. *12*, 865.

Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A.H., Tanaseichuk, O., Benner, C., and Chanda, S.K. (2019). Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nat. Commun. *10*, 1523.

STAR METHOD

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals		
carbenicillin	Bioshop	CAR544.10
chloramphenicol	Bioshop	CLR201.10
Omni plates	ThermoFisher Scientific	12-565-296
3-amino-1,2,4-triazole (3-AT)	Sigma Aldrich	A8056
cycloheximide (CYH)	Sigma Aldrich	C7698
Phosphate-Buffered Saline (PBS)	Wisent	311-425-CL
Propidium iodide (PI)	Bioship	PPI888.10
Critical Commercial Assays		
HiPure Plasmid Filter Maxiprep Kit	Invitrogen	K210016
Gateway LR reactions	ThermoFisher Scientific	11791100
Yeast Plasmid Miniprep Kits	Zymoprep	D2004
NucleoSpin Gel and PCR Clean-up kit	Macherey-Nagel	740609.250
Qubit	Invitrogen	Q32851
Agencourt AMPure XP	Beckman Coulter	A63881
KAPA library quantification kit	KAPA Biosystems	KK4824
KAPA SYBR FAST qPCR Master Mix	KAPA Biosystems	KM4101
96-well DNA extraction kit	Qiagen	27193
Deposited Data		
SARS-CoV-2 Genomes (Wuhan/IVDC-HB-01/2019 (HB01), Wuhan/IVDCHB-04/2019 (HB04) and Wuhan/IVDC-HB-05/2019 (HB05))	(Wu et al., 2020c)	NC_045512
Human ORFeome collection v9.1	(Luck et al., 2020; The	N/A

	ORFeome Collaboration, 2016)	
barcoded human ORFeome (bhORFeome)	(Luck et al., 2020)	N/A
hsPRS-v2 (Human Positive Reference Set, v2)	(Braun et al., 2009; Choi et al., 2019)	N/A
hsRRS-v2 (Human Random Reference Set, v2)	(Braun et al., 2009; Choi et al., 2019)	N/A
virus-host literature binary multiple reference set (vH-Lit-BM)	This manuscript	N/A
viral Random Reference Set (vRRS)	This manuscript	N/A
Experimental Models: Cell lines		
ccdB-resistant electrocompetent cells	(Yachie et al., 2016)	N/A
One Shot ccdB Survival 2 T1R Competent Cells	Invitrogen	A10460
Yeast (Y8800, MATa, Prey strains)	(Altmann et al., 2018, 2020)	N/A
Yeast (Y8930, MATα, Bait strains)	(Altmann et al., 2018, 2020)	N/A
Yeast (Y8930, MATα, Bait strains)	(Choi et al., 2019)	N/A
Yeast (RY3011, MATa, Prey strains)	(Kim et al., 2021)	N/A
Yeast (RY3031, MATα, Bait strains)	(Kim et al., 2021)	N/A
Experimental Models: Organisms/Strains		
pENTR223.1*Sfi	Invitrogen	N/A
pAR068 (C-terminal AD fusion, 2μ origin)	(Yachie et al., 2016)	N/A
pHiDEST-AD (N-terminal AD fusion, CEN origin)	(Yachie et al., 2016)	N/A
pHiDEST-DB (N-terminal DB fusion, CEN origin)	(Yachie et al., 2016)	N/A
pPC86 (N-terminal AD fusion, CEN origin)	(Altmann et al., 2018, 2020)	N/A
pDEST-N2H (-N1, -N2, -C1 and -C2)	(Choi et al., 2019)	N/A
Softwares and Algorithms		
bowtie2 (v2.2.3)	(Langmead and Salzberg, 2012)	RRID:SCR_016368 (http://bowtie-bio.sourcefo

		rge.net/bowtie2/index.shtml)
bcl2fastq2 (v2.20.0.422)	Illumina	N/A
Cytoscape (v3.8.2)	(Shannon et al., 2003)	N/A
Metascape web-platform	(Zhou et al., 2019)	N/A
Others		
BioMatrix Robot	S&P Robotics	BM3-BC

LEAD CONTACT AND MATERIALS AVAILABILITY

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Pascal Falter-Braun (pascal.falter-braun@helmholtz-muenchen.de).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Yeast strains are listed in the Key Resources Table.

METHOD DETAILS

Selection & clone preparation of SARS-CoV-2 functional ORFs

Gene synthesis of SARS-CoV-2 ORFs into Gateway compatible vectors

Two independent collections of Gateway Entry plasmids containing the ORFs of SARS-CoV-2 were constructed, one for $HuSCI_{HIS3}$ and another for $HuSCI_{GFP}$.

The $HuSCI_{GFP}$ collection is described in our recent paper (Kim et al., 2020). Briefly, it includes all but one (NSP11, which is too short for Gateway cloning) of the codon-optimized ORFs of SARS-CoV-2, synthesized based on a published genome (Wu et al., 2020c), and then transferred to Gateway-compatible Entry plasmids, in both 'open' (without a stop codon, to enable C-terminal fusions) and 'closed' (with a stop codon) configurations. In a similar fashion, the $HuSCI_{HIS3}$ collection of Gateway-compatible Entry plasmids was based on SARS-CoV-2 ORFs from published viral genome sequences Wuhan/IVDC-HB-01/2019 (HB01), Wuhan/IVDCHB-04/2019 (HB04) and Wuhan/IVDC-HB-05/2019 (HB05) (Wu et al., 2020b). All viral ORFs for $HuSCI_{HIS3}$ were synthesized by

Twist Bioscience, and include two linkers at the 5´ and 3´ends, respectively, which contain SfiI restriction sites for cloning into pENTR223.1*SfiI. Additionally, the 5´ linker contains an alternative translational start ATG flanked by BamHI sites, and the 3´ linker contains a stop codon flanked by PacI and AsiSI restriction sites, to replace the native translational stop. This design allows for the removal of either the start or stop codon, to accomodate in-frame fusion to N-terminal or C-terminal protein tags, respectively. For $HuSCI_{HIS3}$, synthesized viral ORFs were cloned into pENTR223.1 by SfiI restriction digest. The start ATG was then removed by BamHI digest and subsequent re-ligation.

In total, 28 viral ORFs were synthesized for $HuSCI_{GFP}$ and 27 for $HuSCI_{HIS3}$. Together, these included: NSP1-16 (except NSP11), all derived from ORF1ab; the four structural proteins S, E, M and N; and the 10 accessory ORFs ORF3A, 3B, 3D, 6, 7A, 7B, 8, 9B, 9C and 10 (Wu et al., 2020b).

Preparation of barcoded Y2H destination plasmids

Barcoded 'prey' (pAR068: C-terminal AD fusion, 2μ origin / pHiDEST-AD: N-terminal AD fusion, CEN origin), and 'bait' (pHiDEST-DB: N-terminal DB fusion, CEN origin) Destination plasmid vectors were prepared using our published protocols (Yachie et al., 2016), with the integration of the barcode locus at the SacI restriction site. Following Gibson assembly (Gibson et al., 2009), DNA was transformed (BioRad MicroPulser electroporator, 1652100) into *ccdB*-resistant electrocompetent cells (made from One Shot *ccdB* Survival 2 T1R Competent Cells (Invitrogen, A10460)) . After selection on Luria-Bertani (LB) agar plates with 100 μg/mL carbenicillin (Bioshop, CAR544.10) and 35 μg/mL chloramphenicol (Bioshop, CLR201.10), bacterial colonies, which are representative of clonal barcoded plasmids, were pooled together and mixed thoroughly to generate a high complexity pool. Pooled barcoded 'bait' and 'prey' destination plasmid DNA were then isolated by maxiprep (Invitrogen, K210016).

Isolation and sequencing of clonal barcoded null plasmids

To isolate clonal barcoded plasmids, the destination plasmid pool was transformed into One Shot *ccdB* Survival 2 T1R Competent Cells (Invitrogen, A10460). Cells were selected on omni plates (Fisher Scientific, 12-565-296) containing LB agar supplemented with 100 μg/mL carbenicillin and 35 μg/mL chloramphenicol, and incubated overnight at 37°C such that each plate yielded 500–1,500 isolated single colonies. Single colonies were picked using a Biomatrix robot (S&P Robotics) and arrayed into 384-well cell culture plates containing 80 μL of LB with carbenicillin (100 μg/mL) and chloramphenicol (35 μg/mL) in each well. The cell culture plates were sealed (Foil Seal, VWR, 60941-126) and incubated at 37°C for 16h.

To determine the sequence of each plasmid barcode, the Kiloseq procedure was performed as described (Weile et al., 2017) with modifications reflecting that we only used this approach to sequence barcodes for these empty vectors. In the first step, we performed colony PCR in 384-well plate format, in which well-specific indexing primers were attached to the barcode sequence. Amplified DNA was pooled plate-wise and DNA pools were purified with a PCR clean-up kit (Macherey-Nagel). Each plate-pool of DNA was then re-amplified with primers designed to add Illumina adapters in order to generate a library of amplicons such that each amplicon carried plate- and well- specific information, allowing multiple plates of amplicons to be sequenced together. Indexed barcoded libraries were assessed with the KAPA library quantification kit (Kapa Biosystems, KK4824) and sequenced using an Illumina NextSeq 500. Sequence analysis was carried out as described previously (Luck et al., 2020; Weile et al., 2017; Yachie et al., 2016).

Subcloning of SARS-CoV-2 ORFs into Y2H Destination plasmids

In preparation for the $HuSCI_{HIS3}$ Y2H screen, each viral ORF was moved to the Destination vectors pPC86 (N-terminal AD fusion, CEN origin) (Altmann et al., 2018, 2020) and pHiDEST-DB (N-terminal DB fusion, CEN origin) (Yachie et al., 2016) by Gateway recombinational cloning to produce AD-Y and DB-X N-terminal fusion constructs, respectively. Since viral ORF sequences had been synthesized and their sequences verified prior to cloning, successful recombination into destination plasmids was confirmed by PCR evaluation of insert size. The viral ORF sequences for all interactions identified in $HuSCI_{HIS3}$ were re-confirmed by Sanger sequencing post-screen (see below).

Gateway LR reactions (Gateway™ LR Clonase™ II Enzyme mix, 11791100) were performed for the entire $HuSCI_{GFP}$ library of SARS-CoV-2 ORFs, and the human gene ACE2 into three different configurations of barcoded destination plasmids (Luck et al., 2020; Yachie et al., 2016): pAR068 (C-terminal AD fusion, 2μ origin (high copy number)), pHiDEST-AD (N-terminal AD fusion, CEN origin (low copy number)), pHiDEST-DB (N-terminal DB fusion, CEN origin (low copy number)) such that each ORF was linked to two to six barcodes in every configuration. Gateway LR reactions were performed individually to ensure that each ORF was linked with known barcode sequences in the constructed destination plasmid collection. These collections of barcoded destination vectors were then sequence-verified to confirm ORF-barcode pairs using Sanger sequencing (TCAG, SickKids Hospital). ORFs and associated barcode sequences are described in Supplementary Table 3.

HIS3 based screening to generate the HuSCI$_{HIS3}$ map

The Y2H$_{HIS3}$ screening pipeline is essentially as described (Rual et al., 2005; Yu et al., 2008). Briefly, AD-Y and DB-X plasmids containing 27 viral ORFs were transformed into yeast strains Y8800 (MATa) and Y8930 (MATα), respectively (Altmann et al., 2018). Individual DB-X viral fusion constructs were first mated with yeast containing AD-empty plasmid, in order to identify 'auto-activating' 'baits'. Only NSP1 showed a high level of activation of the *LYS2::GAL1-HIS3* reporter on selective plates containing synthetic complete media lacking leucine, tryptophan and histidine (SC-Leu-Trp-His), supplemented with 20 mM of 3-amino-1,2,4-triazole (3-AT) (Sigma Aldrich, A8056) and was therefore excluded from screening in this orientation. All 26 remaining DB-X viral ORFs were individually mated with 99 pools, each containing ~188 AD-Y human ORF constructs, obtained from the human ORFeome collection v9.1 from the Center for Cancer Systems biology (CCSB), comprising a total of 17,472 ORFs (Luck et al., 2020; The ORFeome Collaboration, 2016). For an alternative screening orientation, yeast cultures containing 27 AD-Y viral fusion constructs were pooled and the pool was mated with DB-X human fusion constructs individually, also derived from the human ORFeome collection v9.1. DB-X auto-activating constructs had been removed from the collection prior to screening, as previously described (Luck et al., 2020). Screening in both configurations was performed two times in order to increase coverage.

Mating was performed by spotting 5 µL of each of the saturated haploid yeast cultures onto plates containing yeast extract peptone dextrose (YEPD) medium and incubating for 24 hours at 30°C. Yeast spots were then transferred onto SC-Leu-Trp-His + 1mM 3-AT plates using a replica plating block (Altmann et al., 2018, 2020) and incubated for 72 hours at 30°C.

Y2H positive colonies were identified by growth on plates containing SC-Leu-Trp-His + 1mM 3-AT and 3 colonies per spot were picked and individually cultured in SC-Leu-Trp liquid medium and grown for 2 days at 30°C. This culture was spotted (5 µL) on SC-Leu-Trp plates and grown for another 2 days at 30°C. Yeast spots were then transferred onto SC-Leu-Trp-His + 1mM 3-AT and SC-Leu-His + 1mM 3-AT + 1mg/L (~3.4µM) cycloheximide (CYH) (Sigma Aldrich, C7698) plates using a replica plating block to test for cycloheximide-sensitive expression of the *LYS2::GAL1-HIS3* reporter, which identifies spontaneous DB-X auto-activators that can arise during screening (Yu et al., 2008). Positive scoring colonies (growth on SC-Leu-Trp-His + 3-AT, but not on SC-Leu-His + 3-AT + CYH) were picked and interaction candidates of AD-Y or DB-X human ORFs, respectively, were identified using end-read Sanger sequencing of PCR products (Altmann et al., 2018). Identified AD-Y and DB-X human interaction candidates were picked from archival glycerol stocks, cultured in liquid medium and mated (as described above) one-by-one with all 27 DB-X or AD-Y viral ORFs, respectively. This binary verification mating was

performed four times for the orientation of viral DB-X versus human AD-Y and Y2H interactions were scored positive if at least three of four repeats showed robust growth on SC-Leu-Trp-His + 1mM 3-AT plates. For viral AD-Y versus human DB-X, binary verification mating was done in triplicate and positive interactions scored if growth was observed in at least two out of three repeats.

In order to eliminate any operator bias when scoring yeast growth for positive interactions, yeast growth was scored automatically with a custom dilated convolutional neural network (Yu and Koltun, 2015). For this, previous datasets of more than 1,500 images of biochemically and functionally validated binary Y2H studies were used (Altmann et al., 2020). Each image was scaled to achieve equal pixel-distance between the yeast spots of different images. The images were then cropped and sliced, so that each yeast spot became a separate file. In addition, the mean grayscale image of all 96 yeast spots of the same plate was calculated. With this dataset, a simple front-end prediction module was trained. The front-end module consisted of six dilated convolutional layers with exponential increasing dilation rate and two dense layers at the end. After each layer except the last, a Leaky-ReLU activation was added (Maas et al., 2013). The model was optimized with a combination of Softmax and Cross entropy and an Adam Optimizer (Kingma and Ba, 2014). The model achieved an accuracy of >0.9 during all folds of a ten-fold cross-fold validation.

All Y2H interacting pairs were picked from selective plates and the identities of all ORFs (DB-X and AD-Y) were confirmed by Sanger sequencing of PCR products (Altmann et al., 2018).

GFP based screening to generate the HuSCI$_{GFP}$ map

Generation of Y2H$_{GFP}$ strains for HuSCI$_{GFP}$

Development of GFP based Y2H technology, as well as generation of toolkit strains is as described in our forthcoming publication (Kim et al., 2021). Briefly, Y2H strains carrying multiple copies of GFP based Y2H reporters were engineered and empirically benchmarked with previously established PRS and RRS (hsPRS-v2 and hsRRS-v2) sets before proceeding with the experiments described in this study. Genotypes of toolkit strains are described in Supplementary Table 5.

Generation of the barcoded human ORFeome (bhORFeome) for screening

A Gateway Entry clone collection of the human ORFeome was available from CCSB (Luck et al., 2020; The ORFeome Collaboration, 2016). Based on this collection, *en masse* Gateway LR reactions using a pool of barcoded destination plasmids, bacterial transformations, colony picking and kiloseq were performed as described (Weile et al., 2017; Yachie et al., 2016). After obtaining the raw sequencing

results, bowtie2 (v2.2.3) (Langmead and Salzberg, 2012) was used to align the reads in each well to the reference sequences. To optimize this collection of barcoded ORFs, this process was repeated three additional times, using only the missing ORFs from the previous rounds in subsequent barcoding attempts. The final clones were arrayed using an automated robot system (S&P Robotics Inc., BM3-BC). Approximately 1% of the clones were validated by Sanger sequencing with >95% success rate for both ORF and barcode sequences. Our final union of barcoded 'bait' (DB) and 'prey' (AD) plasmids, called 'barcoded human ORFeome (bhORFeome)', includes 16,747 fully sequence verified human ORFs (96% of the entire human ORFeome collection) with most (~95%) ORFs represented by 2 unique barcodes. Prior to yeast transformation of the bhORFeome, we cataloged the barcoded 'bait' and 'prey' destination plasmid collection into a 10-by-10 screening matrix consisting of ten DB and ten AD groups, respectively. Each haploid AD and DB group contains ~1,400 ORFs with two distinct sets of unique barcodes, and ~200 ORFs with a single unique barcode set.

Transformation of SARS-CoV-2 destination plasmids into GFP based Y2H strains

Each uniquely barcoded destination plasmid containing a sequence-verified SARS-CoV-2 ORF was transformed to the corresponding yeast strain (RY3011 for AD fusions, and RY3031 for DB fusions), individually in 96-well plates. After the transformants were selected by growth on SC media lacking specific amino acids (leucine for DB, and tryptophan for AD), 10 replica plates were made for each of the destination vector version collections. These plates were then scraped and pooled together to make glycerol stocks that are pools of all the barcoded SARS-CoV-2 ORFs plus the human ORF ACE2 in each plasmid configuration, to be screened using $Y2H_{GFP}$.

Pooled Glycerol stocks of haploid yeast:

Yeast Strain	**Plasmid backbone**	**ORFs**	**Barcodes**
RY3031	pHiDEST-DB (N-term, DB, CEN)	SARS-CoV-2 + ACE2	>= 2 each
RY3011	pHiDEST-AD (N-term, AD, CEN)	SARS-CoV-2 + ACE2	>= 2 each

Mating of pooled haploid yeast

For $HuSCI_{GFP}$, multiple pooled matings were performed using the frozen haploid glycerol stock pools. Each of the 10 pools of human ORFs (in C-terminal AD fusion plasmids with 2μ origin; pAR068) were separately mixed with the pool of SARS-CoV-2 ORFs plus human ACE2 (in N-terminal DB fusion plasmids with CEN origins; pHiDEST-DB) for a total of 10 pooled matings to cover virus-host interactions. Another separate mating was done between the SARS-CoV-2 pools in both AD and DB fusion, CEN origin plasmids (pHiDEST-AD, pHiDEST-DB). These 11 pool-pool matings were set up to achieve >100x coverage with respect to the number of possible barcode combinations, while taking into account the number of viable colony-forming units (vCFUs) of the haploid glycerol stocks, and the expected mating efficiency. Negative controls consisting of empty destination vectors with matching configurations and known barcodes were also added to each mating. Matings were achieved by mixing together equal amounts of each haploid strain (plus added negative controls), spreading the cell mixture onto a double-strength YEPD plus adenine (2x YPAD) agar plate, and incubating at 30°C for 24h. Mating plates were then scraped with distilled water, and each mated culture was spread across 20 large (15 cm) SC-Leu-Trp petri plates, supplemented with additional histidine (80 mM) to ensure more equal representation of all strains (SC-Leu-Trp+10xHis), which are then incubated at 30°C for a further 72h. These plates were then scraped and assay-ready pooled diploid glycerol stocks were made for each of the 11 groups.

Selection of yeast with interacting pair of DB-X and AD-Y by FACS

Diploid pooled glycerol stocks were thawed on ice and inoculated into 1L flasks with a total starting viable CFU of 30M (to achieve ~100x coverage with respect to ORF pairs), and incubated at 30°C at 200 rpm for 24h. Smaller cultures (10 mL) of negative controls were also prepared. Before FACS, 'presort' cultures were prepared for each sample (2 x 10 mL of cultures with OD_{600} 10) with doxycycline added (10 μg/mL) to these cultures to induce barcode swapping while these cultures incubate at 30°C for 24h (Yachie et al., 2016). To prepare for FACS, cells were concentrated by centrifugation (500 x g, 5min) and resuspended in Phosphate-Buffered Saline (PBS, Wisent, 311-425-CL) to a final OD_{600} value of 10 to limit further growth. Negative controls were also centrifuged and resuspended in a final volume of 1mL PBS. Propidium iodide (PI, Bioshop, PPI888.10) was added to all samples and controls to a final concentration of 4 mg/L to discriminate dead yeast cells during FACS.

Using a mated negative control, the FACS gate for GFP positive cells was set to capture 0.1% of GFP-negative cells to yield a 0.01% false positive rate (FPR). FACS was performed on a Sony MA900

cell sorter with the aid of a core facility at Lunenfeld-Tanenbaum Research Institute (LTRI) at Mount Sinai Hospital in Toronto. We sorted over 100M cells per group for >100x coverage with respect to the number of ORF pairs.

After FACS, GFP positive cells for each sample were spread across 10 plates of selective agar media (SC-Leu-Trp+Ade+10xHis) and incubated at 30°C for 72h. These plates were then scraped, centrifuged, and resuspended into 2 x 10mL cultures at OD_{600} of 10. Doxycycline (10 µg/mL) was then added to these cultures to induce barcode swapping while they were incubated at 30°C for 24h.

After doxycycline induction of both presort and GFP-positive samples, cells were harvested and plasmid DNA was extracted via yeast miniprep (Zymo Research, D2004). Fused barcodes were then amplified with primers that attach modified Illumina i5 and i7 adapters in order to uniquely identify each different sample. PCR products were run on a gel and the bright band at ~350 bp was size-selected and purified using the NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel, 740609.250).

NGS library preparation and sequencing

DNA concentrations were measured for each sample using a Qubit (Invitrogen, Q32851) and, using the DNA concentrations as a guide, samples were pooled together to ensure relative sequencing depth of each library was proportional to the number of protein pairs tested. The pooled NGS library was then processed with Agencourt AMPure XP (Beckman Coulter, A63881) to remove primer dimers, and quantified by qPCR (KAPA SYBR FAST qPCR Master Mix, KAPA Biosystems, KM4101). The library was sequenced on an Illumina NextSeq using a mid- or high-output 150 cycle kit to achieve the appropriate read-depth necessary for subsequent analysis.

Read counting based on the expected barcodes

The sequencing data was demultiplexed using bcl2fastq2 (v2.20.0.422) provided by Illumina with the following command:

“bcl2fastq -r 10 -p 20 -w 10 --no-lane-splitting --barcode-mismatches 1 --adapter-stringency 0.7 --ignore-missing-bcls --ignore-missing-filter --ignore-missing-positions”

After demultiplexing, the fastq files were aligned to the group specific reference files using bowtie2 (Langmead and Salzberg, 2012) with the following parameters:

For read 1: -q --norc --local --very-sensitive-local -t -p 23 --reorder

For read 2: -q --nofw --local --very-sensitive-local -t -p 23 --reorder

The reference files contain the expected barcode sequences for the ORFs in each group. After

alignments, we count the number of reads mapped to each up/dn tag sequence. The reads were filtered out if the MAPQ score was lower than 20. Based on the BFG barcode recombination process (Yachie et al., 2016), paired end reads should map to up-up or dn-dn when an interaction is present. Hence, for each group, we build a matrix with DB-X and AD-Y as row names and virus ORFs as column names. We count the number of reads mapped to up-up and dn-dn separately and merge them together as our final read counts. The pipeline was implemented in Python2.7 and the source code can be found on Github (https://github.com/RyogaLi/BFG_Y2H/).

Calculation of interaction score and statistical test to identify confident PPI between SARS-CoV-2 and human ORFs

As described in (Yachie et al., 2016) we used marginal frequency for PreSort. Interaction score (IS) was defined by

$$IS_{ij} = \frac{f_{ij}^{GFP}}{f_{ij}^{PreSort}}$$

$$f_i^{PreSort} = \sum_i c_{ij}^{PreSort} / \sum_j [\sum_i c_{ij}^{PreSort}] \quad f_j^{PreSort} = \sum_j c_{ij}^{PreSort} / \sum_i [\sum_j c_{ij}^{PreSort}]$$

$$f_{ij}^{PreSort} = \max(f_i^{PreSort}, f_{AD}^{Floor}) \times \max(f_j^{PreSort}, f_{DB}^{Floor}) \qquad f_{AD}^{Floor} = 10^{-5} \quad f_{DB}^{Floor} = 10^{-4}$$

$$f_{ij}^{GFP} = c_{ij}^{GFP} / \sum_{ij} c_{ij}^{GFP}$$

with c, read count; i, AD barcode; j, DB barcode; f. frequency.

For each DB barcode, we identified the score threshold achieving a 1% false positive rate (FPR) using the 960 AD null barcodes included in this screen. An interaction was accepted as positive only if the ORF pair was above this IS score threshold for ≥ 2 barcode pairs.

Calculation of interaction score to identify confident PPI between SARS-CoV-2 ORFs

For intraviral screening, adding in the 100 AD-null controls would have enlarged the screening space by a factor of >100 over the actual protein pairs being screened. For this reason, we did not add the negative controls in this screen to use as a 1% false positive rate. Instead, we used a similar method to the one that was used for viral-human PPIs: we used an effect size cutoff using the fold change of the median and each DB clone with a threshold of 1,000 in 3 independent barcode pairs, similar to the scoring method previously used for BFG-Y2H with HIS3-based growth selection (Yachie et al., 2016).

Pairwise retesting using HIS3 selection

Candidate interaction pairs for $HuSCI_{GFP}$ were retested in a pairwise HIS3 growth-based Y2H assay to verify the initial interaction result, following the same protocol as outlined above for the $Y2H_{HIS3}$ screen. Briefly, identified AD-Y human and DB-X viral interaction candidates were picked from archival glycerol stocks, cultured in liquid medium and mated (as described above) one-by-one with barcode replicates pooled together prior to mating. DB-X viral interaction candidates were also individually mated with AD-empty controls to assess whether they are classed as 'auto-activators'. Viral ORFs NSP1 and NSP12 were subsequently omitted from this retesting due to DB auto-activation. Interactions were scored after growing these individually mated interacting pairs on selection medium lacking histidine (SC-Leu-Trp-His), and medium lacking histidine supplemented with 3-AT (SC-Leu-Trp-His+1mM 3-AT) for stronger interactions. After 72-96h of yeast growth these pairwise tests were scored according to the standardized scoring method used for the $Y2H_{HIS3}$ screen (Altmann et al., 2018, 2020). Any interacting pairs that scored above the threshold for positive interactions (score ≥ 3) were collected together and subsequently sent for validation.

Benchmarking $Y2H_{HIS3}$ and $Y2H_{GFP}$ against positive and negative control sets

Assay sensitivity can be defined as a fraction of true interactions that can be detected by a given assay. $Y2H_{HIS3}$ toolkit strains were benchmarked previously (Braun et al., 2009). Briefly, bait and prey constructs constituting 92 hsPRSv1 and 92 hsRRSv1 pairs were tested using the identical protocol to the one used to verify $HuSCI_{HIS3}$. $Y2H_{HIS3}$ recovered 23 out of 92 hsPRSv1 pairs when tested in both bait-prey configurations corresponding to an assay sensitivity of $S_{a\text{-}HIS3}$ = 25%. Only *HIS3* reporter based benchmarking results are used here to match screening conditions. A different version of $Y2H_{GFP}$ using low copy plasmids and N-terminally fused hybrid proteins ($lcnY2H_{GFP}$) was benchmarked using 84 pairs of hsPRSv1 and 92 pairs of hsRRSv1 sets and flow cytometry was used to score for interactions based on percentage of singlets in GFP-positive gate, which was set up using empty bait and prey constructs. In addition, $lcnY2H_{GFP}$ was benchmarked in a pooled setting using all possible combinations of proteins constituting 78 hsPRSv2 and 77 RRSv2 pairs supplemented with a 14 pairs of Y2H-positive controls defined as Calibration Set (CS, (Yachie et al., 2016)). The experiment was carried out and interactions were scored as described above, except empirical null distribution was not used.

Sampling sensitivity can be defined as a fraction of detectable (by a given assay) true interactions that can be detected in a single experiment. Let assay sensitivity be S_a and sampling sensitivity be S_s, such that sensitivity of a given screen S can be calculated as $S = S_a \times S_s$. When sampling is saturated by

repeated screens (S_s = 1), the screening sensitivity is equal to assay sensitivity, *i.e.* S = S_a. In pairwise settings S_s = 1, and therefore assay sensitivity is given by the fraction of hsPRSv1/v2 pairs that score positive. Sampling sensitivity of $Y2H_{HIS3}$ has previously shown to be $S_{s\text{-}HIS3}$ = 60% (Altmann et al., 2018), yielding screening sensitivity of S_{HIS3} = $S_{a\text{-}HIS3}$ x $S_{s\text{-}HIS3}$ = 0.25 x 0.6 x 100% = 15% . Given that $Y2H_{HIS3}$ screen had a search space completeness of 83% (T_{HIS3}=83%) of all possible virus-human protein combinations, completion level of the map generated by $Y2H_{HIS3}$ is C_{HIS3} = T_{HIS3} x S_{HIS3} = 0.83 x 0.15 x 100% = 12.45%. $lcnY2H_{GFP}$ recovered 12 out of 82 ($S_{a\text{-}lcnGFP}$ = 15%) hsPRSv1 pairs when tested in a pairwise single bait-prey configuration and 8 of 92 (9%, $S_{s\text{-}lcnGFP}$ = 9/15 x 100% = 60%) hsPRSv2+CS pairs when tested in a pooled single bait-prey configuration, yielding S_{lcnGFP} = $S_{a\text{-}lcnGFP}$ x $S_{s\text{-}lcnGFP}$ = 0.15 x 0.6 x 100% = 9%. It has previously shown that using high-copy C-terminal fusions increases sensitivity by ~50% without affecting the precision (Luck et al., 2020). Thus, we can derive screening sensitivity of $Y2H_{GFP}$ from that of $lcnY2H_{GFP}$ as S_{GFP} = S_{lcnGFP} x 1.5 = 9% * 1.5 = 13.5%. Given that $Y2H_{GFP}$ tested 70% (T_{GFP}=70%) of all possible virus-human protein combinations, completion level of the map generated by $Y2H_{GFP}$ is C_{GFP} = T_{GFP} x S_{GFP} = 0.70 x 0.135 x 100% = 9.45%. Only 4 out of 31 (12.9%) hsPRSv1 pairs detected by the union of $Y2H_{HIS3}$ and $lcnY2H_{GFP}$ were detected with both methods, indicating a high degree of orthogonality. In addition, $Y2H_{GFP}$ implemented in this study includes additional differences such as high-copy and C-terminal fusion constructs for human proteins. Therefore, we conservatively estimate 90% orthogonality between $Y2H_{HIS3}$ and $Y2H_{GFP}$ (*i.e.* ~90% of detected interactions are different, $O_{HIS3+GFP}$= 90%). Thus, we estimate that the fraction of all true interactions captured by our merged interactome maps is $C_{HIS3+GFP}$ = (C_{HIS3} + C_{GFP}) x $O_{HIS3+GFP}$ ≅ (0.125 + 0.095) x 0.9 x 100% = 19.7%.

yN2H Validation

To assess the quality of our dataset, we used the yN2H assay (Choi et al., 2019) for validation of PPIs detected in both $HuSCI_{HIS3}$ and $HuSCI_{GFP}$. This assay was performed as previously described (Choi et al., 2019). Briefly, viral ORFs and human interaction partners were transferred into pDEST-N2H plasmids (pDEST-N2H-N1, -N2, -C1, and -C2) containing a *LEU2* (N1/C1 vectors) or a *TRP1* (N2/C2 vectors) auxotrophy marker, via Gateway LR cloning. Plasmid DNA was extracted from bacteria using a 96-well DNA extraction kit (Qiagen, 27193) and transformed into haploid *S. cerevisiae* Y8800 (MATa) and Y8930 (MATα) strains. The resulting transformants carrying N1/C1/Fragment1 vectors (*LEU2* cassette, in Y8930 strain) or N2/C2/Fragment2 vectors (*TRP1* cassette, in Y8800 strain) were inoculated into a 96-well microplate (Fisher Scientific/Corning, 07200720A) containing 160 µL liquid SC-Leu or

SC-Trp media per well, respectively. In parallel, established PRS and RRS (hsPRS-v2/hsRRS-v2) (Choi et al., 2019) were inoculated. In addition, two protein pairs from the hsPRS-v2, with different N2H signal intensities, were included in duplicates on every plate (NCBP1/NCBP2 and SKP1/SKP2). We also included a set of ~100 randomly selected pairs of viral human protein pairs as a random reference set (vhRRS; Supplementary Table 2B), for which human proteins were randomly selected from hsRRS-v2. Viral human protein pairs in vhRRS and other viral human datasets were randomly distributed in the plates and tested together with hsPRS-v2/hsRRS-v2, which were in separate plates.

After overnight incubation at 30°C of the haploid cultures, mating was performed by transferring 5μL of each haploid strain into a 96-well microplate containing 160 μL YEPD medium per well. The combined cells were incubated overnight at 30°C. Yeast cells expressing viral or human ORFs were also mated with a yeast strain harbouring a plasmid containing only Fragment 1 or Fragment 2 in order to measure background signal (e.g., N1-X was mated with N2-Y, where X and Y are the proteins of a tested pair, as well as Fragment 2 and vice versa). Selection for diploid yeast cells was conducted by inoculating 10 μL of mated yeast into a 96-well microplate containing 160 μL medium per well (SC-Leu-Trp). After overnight incubation at 30°C, 50 μL was transferred from the first diploid selection culture into 1.2 mL medium (SC-Leu-Trp) in deep well plates (Qiagen, 19579), which were incubated overnight at 30°C with shaking at 900 rpm. These plates were then centrifuged (2,500 rpm, 15 min) and the supernatant discarded. Each yeast cell pellet was individually resuspended in 100 μL of the NanoLuc Assay solution by gently pipetting up-and-down several times until fully resuspended. Homogenized solutions were transferred into white flat-bottom 96-well plates (Greiner Bio-One, 655073) that were then incubated in the dark (1h, room temperature). Using a luminometer (TriStar² S LB 942, Berthold, 61457) set up for one-second orbital shaking before each measurement, luminescence was evaluated for each sample (integration time of the luminescence signal was set at 2s per sample). To score each protein pair X–Y, we calculated a normalized luminescence ratio (NLR) corresponding to the raw luminescence value of the tested pair (X–Y) divided by the maximum luminescence value from one of the two controls (X-Fragment 2 or Fragment 1-Y) (Choi et al., 2019). The 1% RRS threshold was based on the vhRRS and determined using the R quantile function.

Bioinformatic analysis of binary interactome map

Network construction and visualization

Starting from a list of protein interaction pairs detected by $Y2H_{HIS3}$ or $Y2H_{GFP}$ and validated by yN2H, we

constructed a graph (network) where viral proteins and human targets are represented by vertices (nodes), and interactions are represented by edges (links). Cytoscape (v3.8.2) was used to visualize the interaction network (Shannon et al., 2003).

Gene Ontology enrichment analysis

The Metascape web-platform was used to select GO terms enriched for our binary interaction human targets and four different AP-MS studies (Gordon et al., 2020a, 2020b; Li et al., 2020; Nabeel-Shah et al., 2020; Stukalov et al., 2020; Zhou et al., 2019). The human ORFeome collection used for the Y2H screen was used as the custom background for our binary interaction host proteins, and the background for four different AP-MS studies was according to their screening search space. GO terms with a hypergeometric $p < 0.01$, an observed gene count ≥ 3 and an enrichment factor >1.5 were collected and grouped into clusters based on their membership similarities. The adjusted p-values were calculated using the Benjamin-Hochberg procedure. The enrichment factor is the ratio between the observed and the expected counts. Kappa scores were used as the similarity metric when performing hierarchical clustering on the enriched terms, and sub-trees with a similarity > 0.3 were considered a cluster. The most statistically significant GO term within a cluster was chosen to represent the cluster.

Identification of interologs via pairwise Y2H testing

Curation of previously known binary interactions involving SARS-CoV-1 and host proteins

To identify interologs representative of conserved interactions, we examined known binary interactions between SARS-CoV-1 and human proteins. A curated list of these virus-host interactions (97 unique interactions) can be found in Supplementary Table 4. The barcoded collections described above (bhORFeome, and barcoded SARS-CoV-2 destination ORF collections) were used for these interaction tests, and the pairwise tests were performed in the same orientation as the large scale $Y2H_{GFP}$ screen (bhORFeome in pAR068 and viral ORFs in pHiDEST-DB). A total of 77 binary interactions involving barcoded human ORFs available in our collection were examined.

Pairwise testing of known SARS-CoV-1 binary interactions to identify interologs

Individual *E. coli* clones of barcoded human ORFs were cherry picked from our bhORFeome (in pAR068) using a BioMatrix Robot (S&P Robotics Inc.). Approximately 20% of these clones were then verified via Sanger sequencing (>90% success rate for both ORF and BC sequences). A total of 63

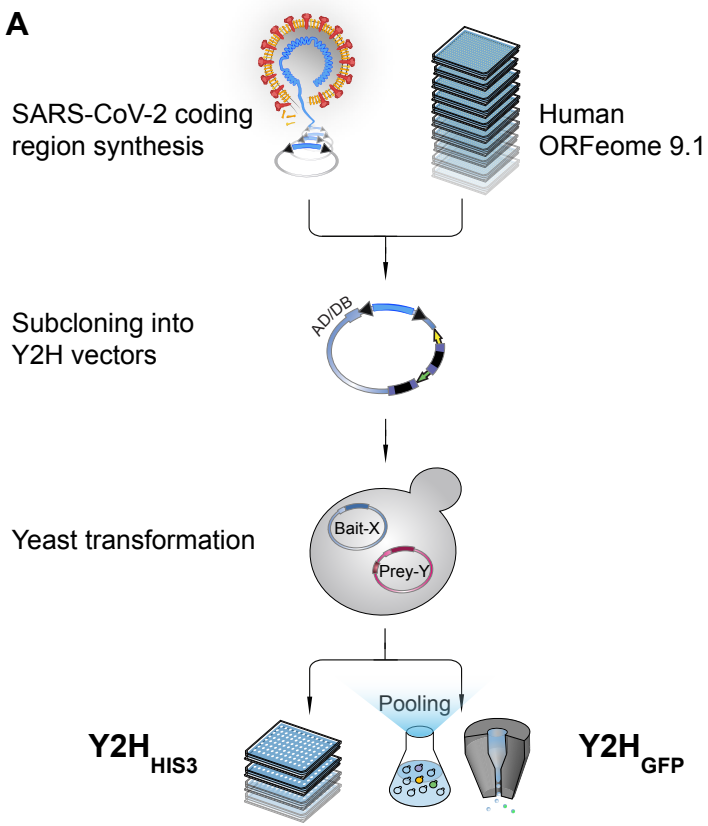
human destination plasmid clones (60 with 2 different unique barcode sets and 3 with a single unique barcode set) were identified for this assay. The *E. coli* clones bearing each barcode set were cultured in 96-deepwell plates with LB + 100 µg/mL carbenicillin media at 37°C for 48h. Cultures for the two barcode sets for the same ORF were merged in equal amounts before *E. coli* plasmid mini-prep (Macherey-Nagel, 740727.250). Plasmid DNA was then transformed into the RY3011 yeast strain in 96-well format to keep each ORF separate (Frozen-EZ yeast transformation II kit, Zymo Research, T2001). The resultant haploid prey (AD-Y) transformants were cultured to saturation at 30°C in SC media lacking tryptophan (SC-Trp). Glycerol stocks were made with a final concentration of 25% glycerol and then stored at -80°C. The haploid bait strains of individual virus ORFs were previously prepared (see above). Haploids were arranged in a 96-well format to accommodate the pairwise assay.

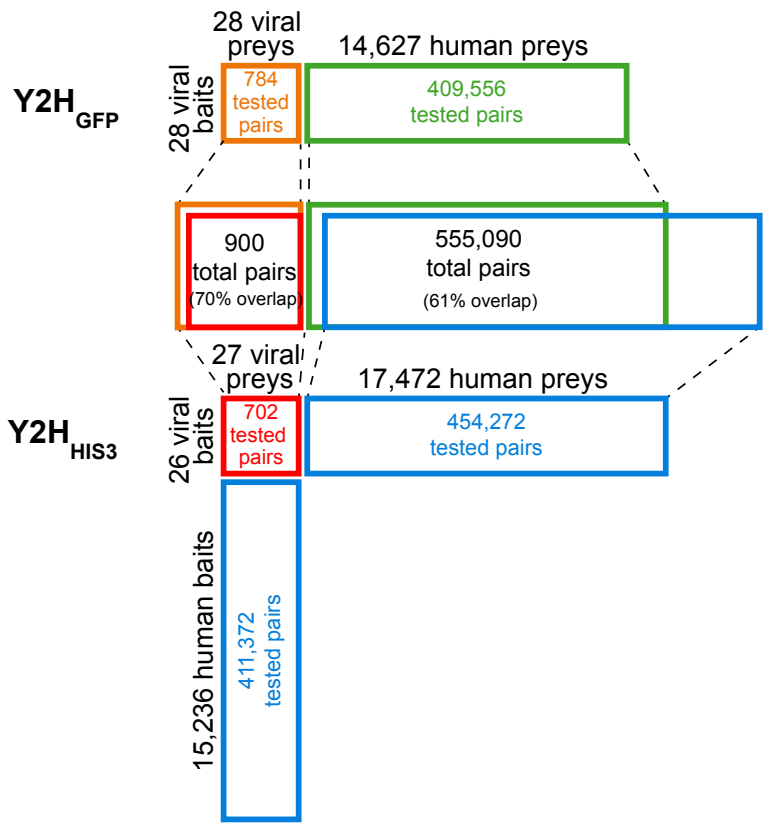
Frozen haploid glycerol stocks of both bait and prey strains were inoculated (5 uL) into 180 uL of haploid growth media (SC-Leu for bait, and SC-Trp for prey) and grown at 30°C for until saturation (48-72h). Mating was then performed individually by mixing 40 µL of saturated haploid yeast cultures in 96-well plates, followed by a 24h incubation in 2x YPAD media at room temperature. The mated cultures were then washed with ddH_2O, transferred to diploid selection media (SC-Leu-Trp+10xHis), and cultured at 30°C for 3 days. Next, 7 µL of saturated diploid cultures were spotted onto solid diploid selection media and incubated at 30°C for 4 days. Diploid yeast spots were then inoculated into fresh liquid diploid media and grown to saturation in deep-well plates. As a negative control, a pool of empty barcoded AD strains were also mated with the individual DB-X strains, to detect any auto-activation among the viral baits. Resultant diploids along with glycerol stocks of assay diploids were stored at -80°C.

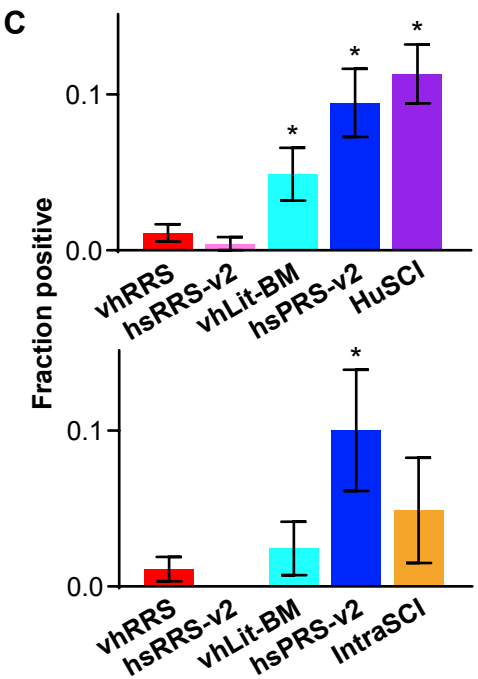
To perform the *HIS3*-based Y2H growth-selection assay, cultures were started by inoculating 20µL of glycerol stock into a total volume of 1,200 µL liquid diploid selection media (SC-Leu-Trp+10xHis). Following growth saturation at 30°C, 100 µL of these cultures were transferred into 96-well plates. The BioMatrix robot was then used to condense 96-well plates into 384-well plates by pinning individual wells onto plates bearing solid diploid media for outgrowth at 30°C for 3 overnights. We then used the BioMatrix robot to transfer yeast colonies (now in 384-well format) to plates containing ddH_2O, which were subsequently used to spot yeast onto the following Y2H selection media plates in duplicates: i) diploid media, SC-Leu-Trp+10xHis, ii) -His selection, SC-Leu-Trp-His, and iii) -His+3-AT selection, SC-Leu-Trp-His+1mM 3-AT. Plates were incubated at 30°C for 3 overnights before image capture and analysis. Successful interactions were indicated by the presence or absence of colony growth in both replicates, upon selection in -His and/or in the more stringent -His+3-AT conditions.

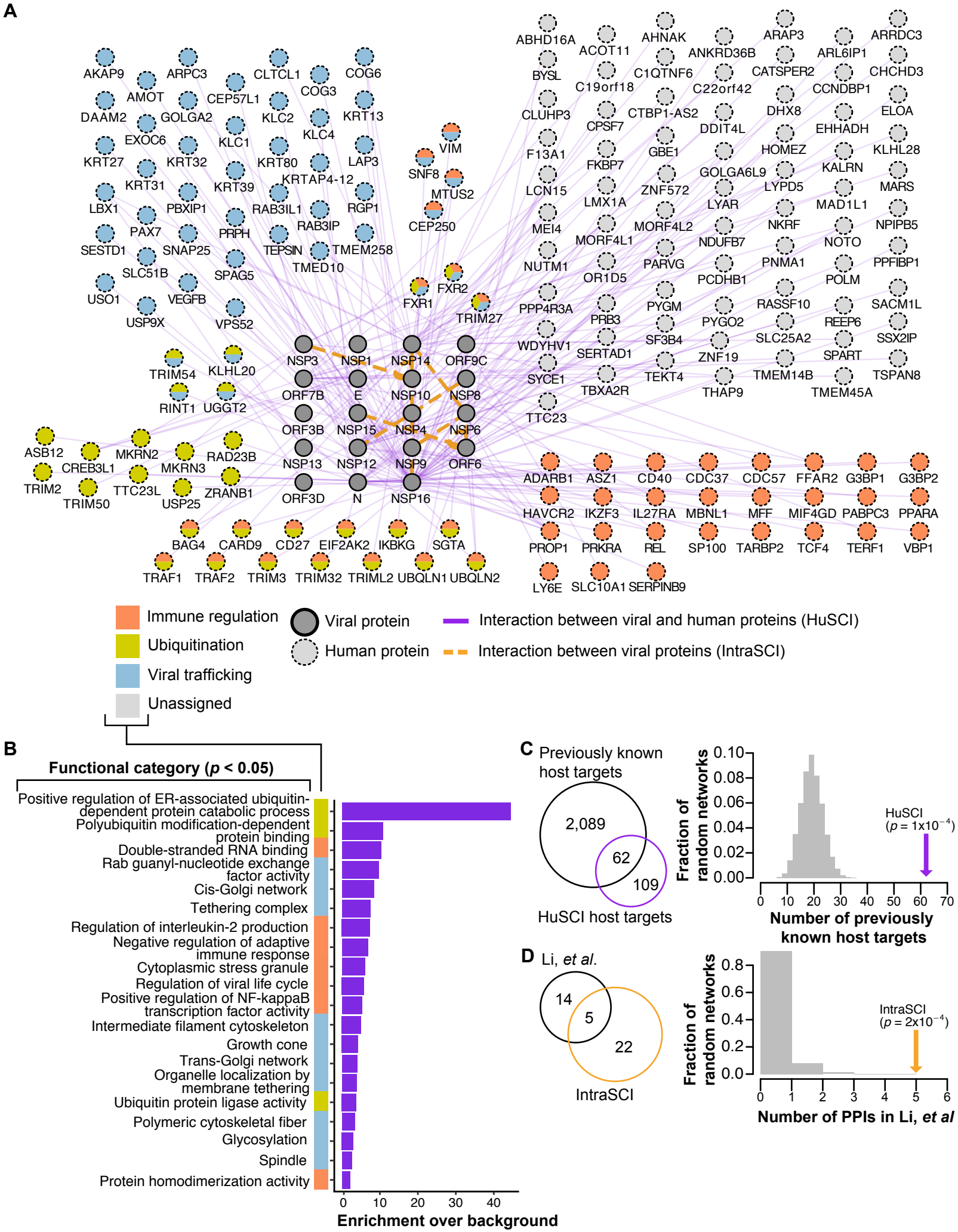
QUANTIFICATION AND STATISTICAL ANALYSIS

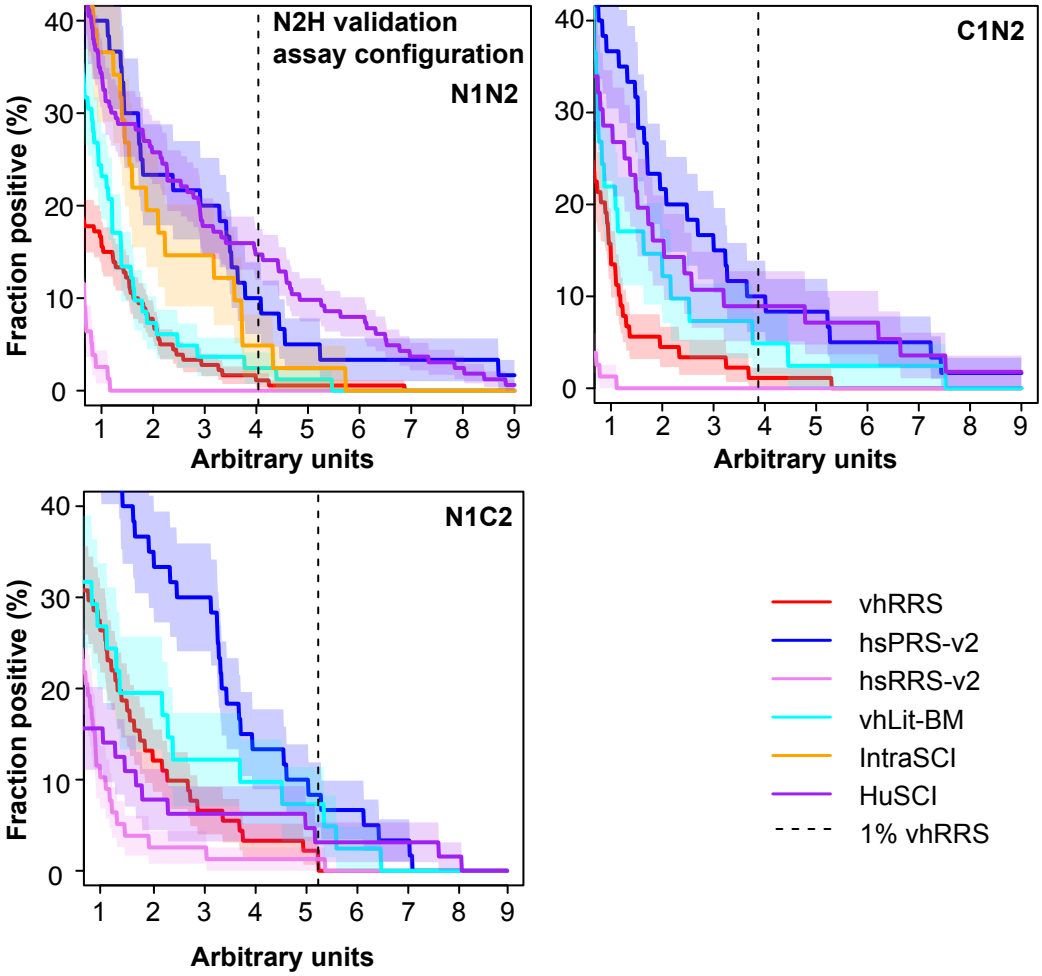
Data in Fig. 1C are presented as a fraction (%) of protein pairs in a given set identified as positive at 1% positive vhRRS threshold. Error bars represent standard error of the positive fractions of a given set from N replica experiments. *p* value was derived from Fisher's exact test. Data in Fig. 2B are presented as percentages of observed items over background items, *p* value was derived from hypergeometric distribution. Data in Fig. 2C are presented as fractions (%) of random interactors or protein interactions that were observed in either previous known human host targets or identified intraviral protein interactions. *p* value was derived from empirical testing.

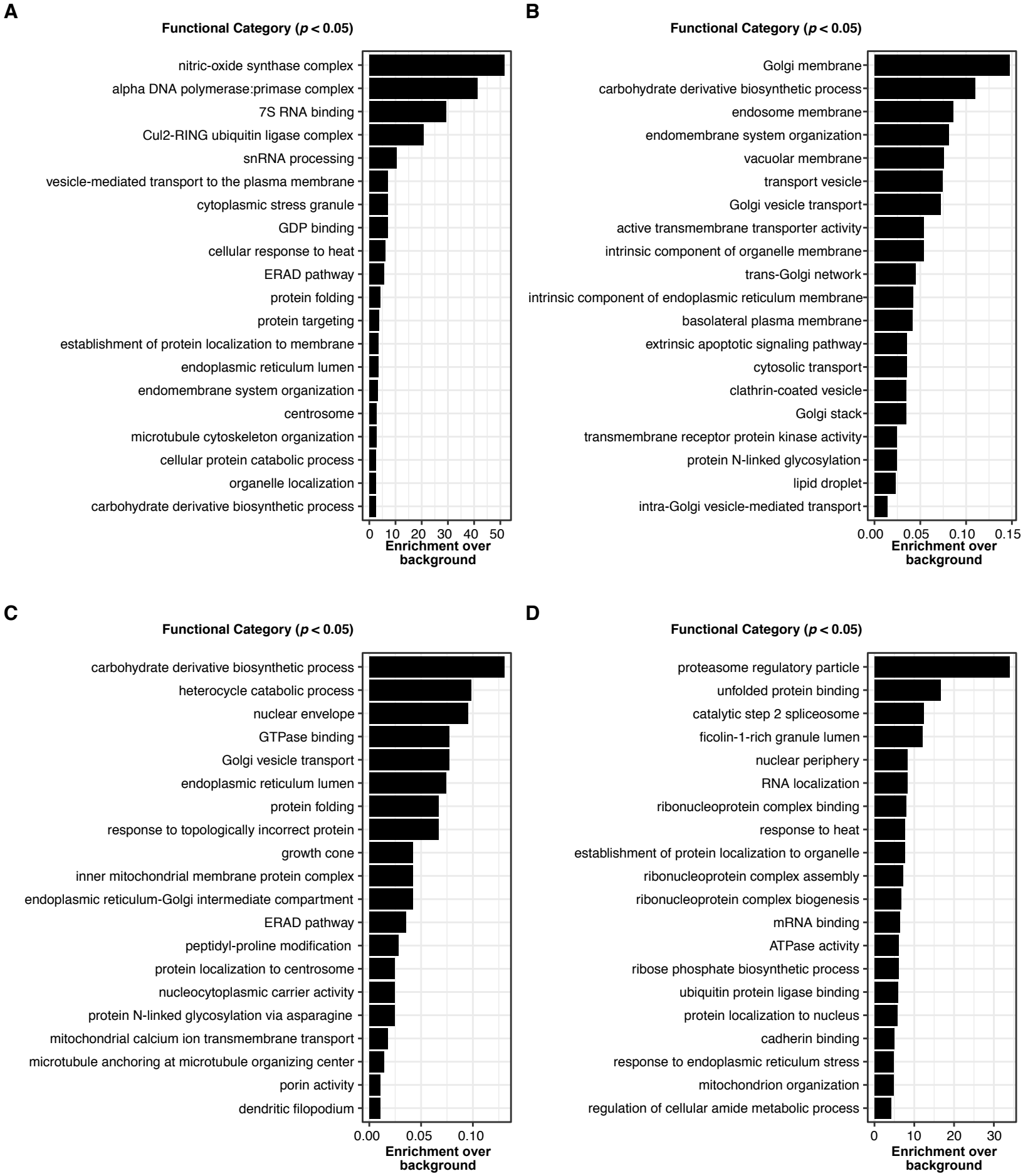

A
SARS-CoV-2 coding region synthesis
Human ORFeome 9.1
Subcloning into Y2H vectors
AD/DB
Yeast transformation
Bait-X
Prey-Y
Pooling
Y2H HIS3
Y2H GFP

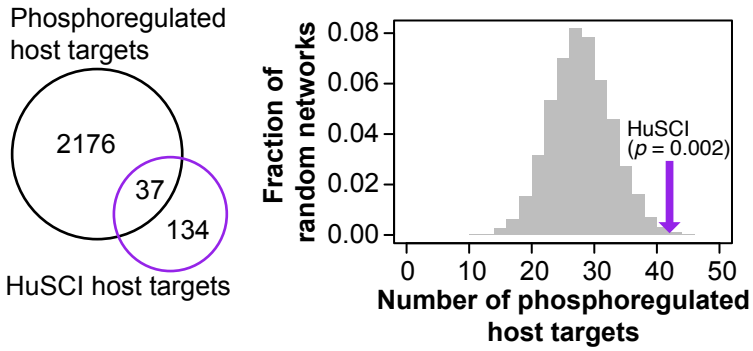

B
Y2H GFP
28 viral preys
14,627 human preys
28 viral baits
784 tested pairs
409,556 tested pairs
900 total pairs (70% overlap)
555,090 total pairs (61% overlap)
Y2H HIS3
27 viral preys
17,472 human preys
26 viral baits
702 tested pairs
454,272 tested pairs
15,236 human baits
411,372 tested pairs


C
Fraction positive
0.1
0.0
vhRRS
hsRRS-v2
vhLit-BM
hsPRS-v2
HuSCI
IntraSCI
*


A
AKAP9
ARPC3
CLTCL1
COG6
COG3
AMOT
CEP57L1
DAAM2
GOLGA2
KLC2
KRT13
KLC4
EXOC6
KLC1
KRT27
KRT32
KRT80
LAP3
KRTAP4-12
KRT31
KRT39
LBX1
PBXIP1
RAB3IL1
RGP1
RAB3IP
PAX7
PRPH
SESTD1
SNAP25
TEPSIN
TMEM258
TMED10
SLC51B
SPAG5
USO1
VEGFB
USP9X
VPS52
VIM
SNF8
MTUS2
CEP250
FXR2
FXR1
TRIM27
TRIM54
KLHL20
RINT1
UGGT2
NSP3
NSP1
NSP14
ORF9C
ORF7B
E
NSP10
NSP8
ORF3B
NSP15
NSP4
NSP6
NSP13
NSP12
NSP9
ORF6
ORF3D
N
NSP16
ASB12
MKRN2
RAD23B
CREB3L1
MKRN3
TRIM2
TTC23L
ZRANB1
TRIM50
USP25
BAG4
CARD9
CD27
EIF2AK2
IKBKG
SGTA
TRAF1
TRAF2
TRIM3
TRIM32
TRIML2
UBQLN1
UBQLN2
ABHD16A
ACOT11
AHNAK
ARAP3
ARRDC3
BYSL
C19orf18
C1QTNF6
ANKRD36B
ARL6IP1
CATSPER2
CHCHD3
C22orf42
CCNDBP1
CLUHP3
CTBP1-AS2
DHX8
ELOA
CPSF7
DDIT4L
EHHADH
F13A1
GBE1
HOMEZ
KLHL28
FKBP7
GOLGA6L9
KALRN
LCN15
ZNF572
LYPD5
MARS
LMX1A
LYAR
MAD1L1
MEI4
MORF4L2
NKRF
NPIPB5
MORF4L1
NDUFB7
NOTO
NUTM1
PARVG
PNMA1
PPFIBP1
OR1D5
PCDHB1
POLM
PPP4R3A
PYGM
RASSF10
SACM1L
PRB3
PYGO2
REEP6
WDYHV1
SF3B4
SLC25A2
SSX2IP
SERTAD1
ZNF19
SPART
SYCE1
TEKT4
TMEM14B
TSPAN8
TBXA2R
THAP9
TMEM45A
TTC23
ADARB1
ASZ1
CD40
CDC37
CDC57
FFAR2
G3BP1
G3BP2
HAVCR2
IKZF3
IL27RA
MBNL1
MFF
MIF4GD
PABPC3
PPARA
PROP1
PRKRA
REL
SP100
TARBP2
TCF4
TERF1
VBP1
LY6E
SLC10A1
SERPINB9
Immune regulation
Ubiquitination
Viral trafficking
Unassigned
Viral protein
Human protein
Interaction between viral and human proteins (HuSCI)
Interaction between viral proteins (IntraSCI)
B
Functional category (p < 0.05)
Positive regulation of ER-associated ubiquitin-dependent protein catabolic process
Polyubiquitin modification-dependent protein binding
Double-stranded RNA binding
Rab guanyl-nucleotide exchange factor activity
Cis-Golgi network
Tethering complex
Regulation of interleukin-2 production
Negative regulation of adaptive immune response
Cytoplasmic stress granule
Regulation of viral life cycle
Positive regulation of NF-kappaB transcription factor activity
Intermediate filament cytoskeleton
Growth cone
Trans-Golgi network
Organelle localization by membrane tethering
Ubiquitin protein ligase activity
Polymeric cytoskeletal fiber
Glycosylation
Spindle
Protein homodimerization activity
0
10
20
30
40
Enrichment over background
C
Previously known host targets
2,089
62
109
HuSCI host targets
Fraction of random networks
0.10
0.08
0.06
0.04
0.02
0.00
HuSCI
(p = 1x10−4)
0
10
20
30
40
50
60
70
Number of previously known host targets
D
Li, et al.
14
5
22
IntraSCI
Fraction of random networks
0.8
0.6
0.4
0.2
0.0
IntraSCI
(p = 2x10−4)
0
1
2
3
4
5
6
Number of PPIs in Li, et al


N2H validation
assay configuration
N1N2
C1N2
N1C2
Fraction positive (%)
Arbitrary units
0
10
20
30
40
1
2
3
4
5
6
7
8
9
vhRRS
hsPRS-v2
hsRRS-v2
vhLit-BM
IntraSCI
HuSCI
1% vhRRS


A
Functional Category (p < 0.05)
nitric-oxide synthase complex
alpha DNA polymerase:primase complex
7S RNA binding
Cul2-RING ubiquitin ligase complex
snRNA processing
vesicle-mediated transport to the plasma membrane
cytoplasmic stress granule
GDP binding
cellular response to heat
ERAD pathway
protein folding
protein targeting
establishment of protein localization to membrane
endoplasmic reticulum lumen
endomembrane system organization
centrosome
microtubule cytoskeleton organization
cellular protein catabolic process
organelle localization
carbohydrate derivative biosynthetic process
0 10 20 30 40 50
Enrichment over background
B
Functional Category (p < 0.05)
Golgi membrane
carbohydrate derivative biosynthetic process
endosome membrane
endomembrane system organization
vacuolar membrane
transport vesicle
Golgi vesicle transport
active transmembrane transporter activity
intrinsic component of organelle membrane
trans-Golgi network
intrinsic component of endoplasmic reticulum membrane
basolateral plasma membrane
extrinsic apoptotic signaling pathway
cytosolic transport
clathrin-coated vesicle
Golgi stack
transmembrane receptor protein kinase activity
protein N-linked glycosylation
lipid droplet
intra-Golgi vesicle-mediated transport
0.00 0.05 0.10 0.15
Enrichment over background
C
Functional Category (p < 0.05)
carbohydrate derivative biosynthetic process
heterocycle catabolic process
nuclear envelope
GTPase binding
Golgi vesicle transport
endoplasmic reticulum lumen
protein folding
response to topologically incorrect protein
growth cone
inner mitochondrial membrane protein complex
endoplasmic reticulum-Golgi intermediate compartment
ERAD pathway
peptidyl-proline modification
protein localization to centrosome
nucleocytoplasmic carrier activity
protein N-linked glycosylation via asparagine
mitochondrial calcium ion transmembrane transport
microtubule anchoring at microtubule organizing center
porin activity
dendritic filopodium
0.00 0.05 0.10
Enrichment over background
D
Functional Category (p < 0.05)
proteasome regulatory particle
unfolded protein binding
catalytic step 2 spliceosome
ficolin-1-rich granule lumen
nuclear periphery
RNA localization
ribonucleoprotein complex binding
response to heat
establishment of protein localization to organelle
ribonucleoprotein complex assembly
ribonucleoprotein complex biogenesis
mRNA binding
ATPase activity
ribose phosphate biosynthetic process
ubiquitin protein ligase binding
protein localization to nucleus
cadherin binding
response to endoplasmic reticulum stress
mitochondrion organization
regulation of cellular amide metabolic process
0 10 20 30
Enrichment over background


Phosphoregulated host targets
2176
37
134
HuSCI host targets
Fraction of random networks
0.08
0.06
0.04
0.02
0.00
HuSCI
(p = 0.002)
0
10
20
30
40
50
Number of phosphoregulated host targets