NAR Molecular Medicine, 2024, **1**, 1–18
https://doi.org/10.1093/narmme/ugad001
Advance access publication date: 23 January 2024
Standard Article

Predicting host-based, synthetic lethal antiviral targets from omics data

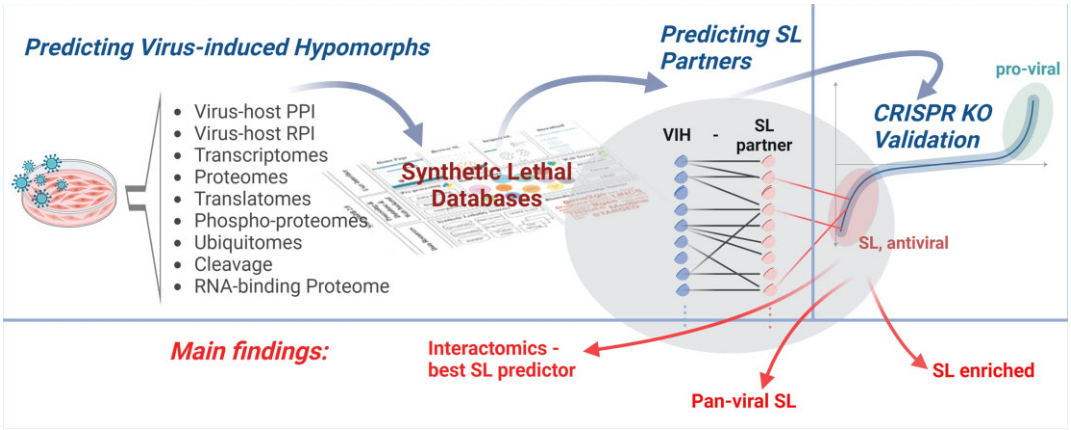
Jeannette P. Staheli, Maxwell L. Neal, Arti Navare, Fred D. Mast and John D. Aitchison*

Center for Global Infectious Disease Research, Seattle Children's Research Institute, Seattle, WA 98101, USA
*To whom correspondence should be addressed. Tel: +1 206 884 3125; Email: john.aitchison@seattlechildrens.org

Abstract

Traditional antiviral therapies often have limited effectiveness due to toxicity and the emergence of drug resistance. Host-based antivirals are an alternative, but can cause nonspecific effects. Recent evidence shows that virus-infected cells can be selectively eliminated by targeting synthetic lethal (SL) partners of proteins disrupted by viral infection. Thus, we hypothesized that genes depleted in CRISPR knockout (KO) screens of virus-infected cells may be enriched in SL partners of proteins altered by infection. To investigate this, we established a computational pipeline predicting antiviral SL drug targets. First, we identified SARS-CoV-2-induced changes in gene products via a large compendium of omics data. Second, we identified SL partners for each altered gene product. Last, we screened CRISPR KO data for SL partners required for cell viability in infected cells. Despite differences in virus-induced alterations detected by various omics data, they share many predicted SL targets, with significant enrichment in CRISPR KO-depleted datasets. Our comparison of SARS-CoV-2 and influenza infection data revealed potential broad-spectrum, host-based antiviral SL targets. This suggests that CRISPR KO data are replete with common antiviral targets due to their SL relationship with virus-altered states and that such targets can be revealed from analysis of omics datasets and SL predictions.

Graphical abstract



Introduction

We recently proposed that synthetic lethality (1), a well-established concept in cancer therapy (2–12), could mitigate multiple known shortcomings of current antiviral drugs such as the paucity of options, restricted target spectrum due to the small genome size of viruses, the often rapid development of drug resistance (13), marginal effectiveness due to dose-limiting toxicity (14) and the necessity to redesign drugs as new viral strains emerge (15,16). Of 92 approved antiviral drugs in 2018, two-thirds were aimed at human immunodeficiency virus (HIV) and hepatitis C virus and almost 90% were small molecule drugs (17,18). A mere 10% were host-based therapeutics, half of which were interferon-related biologics. The potential advantages of host-based drugs highlight the pressing need for new host-based methodologies to target infected cells and for strategies predicting these targets (19,20).

Viruses rely on host cell machinery and induce significant changes that create specific vulnerabilities (1). As obligate intracellular pathogens, viruses require the host cell machinery for every step of the viral life cycle, including attachment, penetration, uncoating, gene expression and replication, assembly and finally virion release (21). During these processes, viruses effectively remodel host cells, converting them into viral factories by leveraging numerous host cell functions, leading to wide-ranging cellular changes (22). These changes are initiated by the direct interaction of viral proteins with host proteins, usurping their functions that lead to cascading di-

Received: August 17, 2023. **Revised:** December 8, 2023. **Editorial Decision:** December 14, 2023. **Accepted:** January 3, 2024

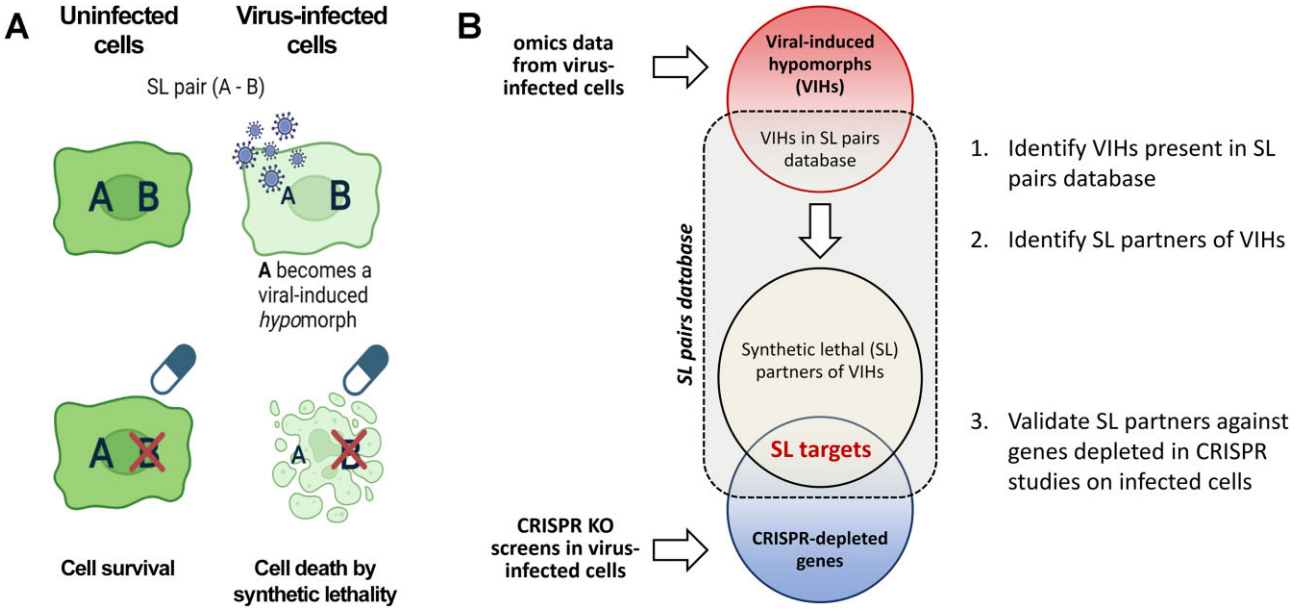


Figure 1. Strategy for identifying antiviral SL targets. (**A**) Synthetic lethality in virus-infected cells: In uninfected cells, disrupting one member of an SL pair (A–B) does not impact viability. In infected cells, where one of the SL partners is hypomorphic, disrupting its SL partner can compromise cells or cause cell death. That is, protein B suppresses the negative consequences of A becoming a VIH when cells are infected. Conversely, protein A suppresses the negative consequences of drugging B in uninfected cells. However, the combination of a VIH and targeting its SL partner with a drug has no recourse from these negative consequences and results in lethality. (**B**) Workflow used for validating predicted SL targets from various omics data classes. SL, synthetic lethal; VIH, viral-induced hypomorph.

rect and indirect effects, including altered protein complexes, changes in RNA and protein abundances, protein mislocalization, and changes in post-translational modifications (PTMs), protein cleavage, splice patterns, and metabolomic and regulatory networks, to name a few. Thus, virus-infected cells have distinct host cell states that introduce specific vulnerabilities to these infected cells.

The antiviral strategy of synthetic lethality capitalizes on these virus-induced vulnerabilities and the cell's reliance on its genetic network integrity. Notably, while many genes can be disrupted individually without undermining cell survival, the disruption of specific genes in combination can be lethal (10, 23–26). When a virus compromises a host protein's normal functionality, this host protein becomes a viral-induced hypomorph (VIH) (1) (Figure 1A). Synthetic lethal (SL) partners of these VIHs can be strategically targeted to impede the viral production machinery, cellular viability or both, providing a means of selectively disrupting the viral life cycle to contain viral dissemination and disease, while sparing uninfected cells (1). Since the interruption of viral replication can be achieved by either the impediment of the viral factory or the death of the virally infected cell, we do not differentiate between synthetic sick and SL cellular states.

The feasibility of targeting SL partners as a host-based antiviral strategy has been substantiated by two studies. Pal *et al.* demonstrated that by utilizing existing transcriptomics and CRISPR data, valid SL targets can be predicted in the context of SARS-CoV-2 (27). In that study, the authors predicted potential VIHs from a handful of transcriptomics datasets in SARS-CoV-2-infected cell lines and patient tissues. SL partners of these VIHs were then identified from cancer SL pairs based on the ISLE algorithm (28) and prioritized by CRISPR knockout (KO) screens from SARS-CoV-2-infected cells. The authors validated 24 predicted SL targets demonstrating reduced viral replication in infected cells combined with a decrease in infected cell counts. Similarly, Navare *et al.* illustrated that virus–host protein–protein interactions (PPIs) can render host proteins hypomorphic and expose cells to SL vulnerabilities (29). Specifically, this study demonstrated that the Golgi-specific brefeldin A resistance guanine nucleotide exchange factor 1 (GBF1) (30) was converted to a hypomorph by binding to poliovirus 3A protein and that depletion of ADP ribosylation factor 1, an SL partner of GBF1, led to cell death specifically in 3A-expressing cells, thus reducing viral replication (29). Importantly, GBF1 is a critical proviral factor and a common target of many viruses, suggesting that targeting SL partners of GBF1 could have broad antiviral activity. Taken together, this evidence suggests that PPIs and transcriptomic alterations can generate VIHs, the SL partners of which could be targeted to inhibit typical viral infection.

In this study, we aimed to determine whether various existing omics types can accurately predict VIHs and, by extension, SL targets. We also posited that a certain subclass of genes, identified through genome-wide CRISPR KO studies, might contain members that are synthetically lethal with VIHs and thus can be explored as antiviral therapeutics. We expected to find such SL targets among genes depleted in CRISPR KO screens alongside antiviral genes, since depletion of these subclasses of genes should lead to diminished cell viability or cell death.

To explore this, we have engineered a computational pipeline to discern the omics datasets best suited for predicting VIHs and the associated SL targets. Our findings indicate that multiple different omics data classes can be used to predict SL targets, as evidenced by enrichment in CRISPR screens. Among these potential data classes, virus–host protein interaction data were better predictors of VIHs. Furthermore, we highlight that many candidate SL targets of SARS-CoV-2 in-

fection are also targets of influenza infection, suggesting that shared SL targets may be suitable for host-based therapeutics against multiple viruses.

Materials and methods

Dataset selection

We utilized 54 published gene sets from 22 studies, encompassing 10 different omics data classes (see Table 1 and references under the 'Identification of VIHs' section), along with 11 SARS-CoV-2 and 3 influenza genome-wide CRISPR–Cas9 KO screens for this study (see Table 2 and references under the 'Compilation of publicly available CRISPR KO screen data' section). We specifically focused on genome-wide screens conducted in cell lines or tissues infected with SARS-CoV-2 or influenza viruses.

Identification of VIHs

We defined VIHs as gene products that exhibited significant reduction in gene function through direct interaction with viral gene products, reduced protein or RNA abundance, decreased translation rate, or changes in protein modifications and other measured behavior that may lead to reduced functional performance, but not complete loss. We determined such changes using the characteristics of 10 omics data classes when comparing virus-infected cells to noninfected cells. Omics data classes included bulk or single-cell RNA sequencing (transcriptomes) (31–35), specialized mass spectrometry (proteomes, RBP, phosphoproteomes, ubiquitomes, cleavage) (32,36–41), and combinations of affinity purification (AP-MS) (32,42–45) or proximity-dependent labeling (BioID-MS) (45–50) with mass spectrometry (Table 1). We also considered alterations in translation rates (translatome) (40) as well as changes in the host RBP and the host interactome with viral RNAs (51–53). All selected datasets were derived from genome-wide studies in cells infected with live SARS-CoV-2 virus, SARS-CoV-2 pseudovirus, or transduced with individual SARS-CoV-2 virus proteins or RNAs for interactome studies.

To identify hypomorphs from available omics data, we adopted the statistical cutoff criteria used in the associated publication of the respective dataset or reported FDR-adjusted P-values <0.05 as significantly altered by infection. Since such P-values were not available for many datasets, meta-analysis of provided data would have been a complex undertaking fraught with additional pitfalls. Lists of significant interactors for virus–host PPIs and virus–host RNA–protein interactions (RPIs) were obtained from relevant publications. Using these criteria, we identified VIHs for each data class and then determined the union of VIHs across all datasets within each of the 10 omics data classes. The union for each data class was subsequently utilized in our computational pipeline for predicting SL targets.

Prediction of candidate SL targets

We defined candidate SL targets as the SL partners of VIHs listed in a combined dataset of SynLethDB 2.0 (last accessed 1 November 2023) (54), including human SL genes only, and SynLEGG (last accessed 20 November 2023) (55), which predicts SL pairs based on low gene expression of one SL partner combined with a gene depleted in CRISPR KO screens of 783 cell lines. Thus, we identified all VIHs that also appeared as an SL pair in this newly generated SL pair database and consolidated the union of all SL genes associated with those VIHs. This process was performed separately for each list of VIHs compiled from different omics data classes. The resulting list of candidate SL targets was then employed in downstream analyses to evaluate prediction accuracy against depleted products in CRISPR KO studies and to assess prediction agreement among different data classes.

Compilation of publicly available CRISPR KO screen data

We gathered publicly available data from genome-wide CRISPR KO screens conducted in cells infected with either SARS-CoV-2 (56–66) or influenza virus (67–69) and identified significantly depleted genes. Genes were considered significantly depleted if they exhibited FDR-adjusted P-values <0.05 along with negative $\log_2$ fold changes or z-scores <-1.96. The aggregated set of CRISPR-depleted genes from all the studies was then used as the surrogate 'gold standard' set of SL targets when evaluating the predictive performance of each omics data class in our study.

Enrichment analysis of predicted SL genes in CRISPR-depleted genes

We used hypergeometric tests to quantify the enrichment of predicted SL targets among depleted genes in CRISPR KO studies. The union of all genes tested across CRISPR KO studies was considered the background population. We considered FDR-adjusted P-values <0.05, computed using the Benjamini–Yekutieli method (70), as indicative of significance.

Sensitivity, specificity and precision metrics

To compute sensitivity, specificity and precision statistics, we designated genes depleted in CRISPR KO screens as positive cases. Negative cases consisted of products evaluated using CRISPR KO screens but were not depleted. Thus, the complete set of genes evaluated by CRISPR KO served as the ground truth against which we assessed SL target predictions. Predicted SL targets generated by our pipeline were considered positive tests, while negative tests consisted of genes not predicted to be SL targets. Therefore, a predicted SL target that was also depleted in CRISPR KO data was considered a true positive. A true negative referred to a gene not predicted to be an SL target and not depleted in CRISPR KO data.

Datasets used for characterizing broadly antiviral candidate SL targets

We employed several datasets to further characterize candidate SL target genes that maybe broadly antiviral for SARS-CoV-2, influenza A and possibly other viruses. Gene paralogs were obtained from Ensembl using the biomaRt package (70) (version 2.50.3). Extremely multifunctional genes (or hubs) were downloaded from MoonDB 2.0 (last accessed 3 November 2023) (71). Essential genes for 10 different cell lines were downloaded from supplemental tables in (72). Drugs targeting predicted broadly antiviral SL targets were downloaded from DrugBank (last accessed 22 September 2022) (73).

Analysis software

All analyses were performed using R (version 4.1.1). Heatmap analyses utilized the pheatmap package (version 1.0.12) with

Table 1. Omics data classes and datasets used to identify candidate SL targets based on VIHs

Data class	Study	Cell line	Tissue/cell type	MOI	HPI	Significance cutoffs
Transcriptome—RNA-seq						
(5 studies/14 datasets)	Blanco-Melo *et al.*	Calu3	Bronchial epithelial	2	24	$\|\log_2 FC\| > 1$ and FDR $P < 0.05$
	Blanco-Melo *et al.*	A549	Alveolar basal epithelial	2	24	$\|\log_2 FC\| > 1$ and FDR $P < 0.05$
	Blanco-Melo *et al.*	$A549^{ACE2}$	Alveolar basal epithelial	0.2	24	$\|\log_2 FC\| > 1$ and FDR $P < 0.05$
	Blanco-Melo *et al.*	$A549^{ACE2}$	Alveolar basal epithelial	2	24	$\|\log_2 FC\| > 1$ and FDR $P < 0.05$
	Blanco-Melo *et al.*	NHBE—primary cells	Bronchial epithelial	2	24	$\|\log_2 FC\| > 1$ and FDR $P < 0.05$
	Blanco-Melo *et al.*	Clinical—lung tissue	Alveolar and bronchial epithelial	NA	NA	$\|\log_2 FC\| > 1$ and FDR $P < 0.05$
	Stukalov *et al.*	$A549^{ACE2}$	Alveolar basal epithelial	2	6, 12, 24	$\|\log_2 FC\| > 1$ and FDR $P < 0.05$
	Huang *et al.*	NHBE—primary cells	Lung epithelial	5	24, 96	$\|\log_2 FC\| > 1$ and FDR $P < 0.05$
	Lieberman *et al.*	Clinical—nasal swabs	Nasal epithelial and immune cells	Low/high	NA	$\|\log_2 FC\| > 1$ and FDR $P < 0.05$
	Xiong *et al.*	Clinical—BALF	Alveolar macrophages (mostly)	NA	NA	$\|\log_2 FC\| > 1$ and FDR $P < 0.05$
Proteome—mass spectrometry						
(5 studies/12 datasets)	Appelberg *et al.*	Huh7	Liver epithelial-like	1	48, 72 versus 488	$\|\log_2 FC\| > 0$ and FDR $P < 0.05$
	Akgun *et al.*	Clinical—nasal swabs	Nasal epithelial and immune cells	NA	NA	$\|\log_2 FC\| > 0$ and FDR $P < 0.05$
	Bojkova *et al.*	Caco2	Colon epithelial	0.01	2, 6, 10, 24	$\|\log_2 FC\| > 0$ and FDR $P < 0.05$
	Stukalov *et al.* (global and individual)	$A549^{ACE2}$	Alveolar basal epithelial	2	6, 12, 24	$\|\log_2 FC\| > 0$ and FDR $P < 0.05$
	Weingarten-Gabbay *et al.*	$A549^{ACE2/TMPRSS2}$	Alveolar basal epithelial	3	24	$\|\log_2 FC\| > 0$ and FDR $P < 0.05$
		$HEK293T^{ACE2/TMPRSS2}$	Embryonic kidney epithelial	3	24	$\|\log_2 FC\| > 0$ and FDR $P < 0.05$
RBP (RNA-binding proteome) remodeling—cRIC						
(1 study/2 datasets)	Kamel *et al.*	Calu3	Bronchial epithelial	1	8, 24	$\|\log_2 FC\| > 0$ and FDR $P < 0.05$
Translatome—mass spectrometry						
(1 study/4 datasets)	Bojkova *et al.*	Caco2Spec	Bronchial epithelial	0.01	2, 6, 10, 24	$\|\log_2 FC\| > 0$ and FDR $P < 0.05$
Phosphoproteome—mass spectrometry (MDLC–MS/MS)						
(2 studies/4 datasets)	Klann *et al.*	Caco2	Colon epithelial	1	24	$\|\log_2 FC\| > 1$ and FDR $P < 0.05$
	Stukalov *et al.*	$A549^{ACE2}$	Alveolar basal epithelial	2	6, 12, 24	$\|\log_2 FC\| > 1$ and FDR $P < 0.05$
Ubiquitome—mass spectrometry						
(1 study/3 datasets)	Stukalov *et al.*	$A549^{ACE2}$	Alveolar basal epithelial	2	6, 12, 24	$\|\log_2 FC\| > 1$ and FDR $P < 0.05$
Cleavage—mass spectrometry (neo-N-termini)						
(1 study/1 dataset)	Meyer *et al.*	$A549^{ACE2}$	Alveolar basal epithelial	1	24	$\|\log_2 FC\| > 0$ and Storey's q-adjusted P-value ≤ 0.05
Virus–host PPI (BioID-MS)						
(6 studies/6 datasets)	Liu *et al.*	Tet-inducible HEK293T	Embryonic kidney epithelial	NA	NA	Significant interactors per authors
	Samavarchi-Tehrani *et al.*	A549	Alveolar basal epithelial			
	St-Germain *et al.*	HEK293 Flp-In T-REx	Embryonic kidney epithelial			
	Meyers *et al.*	HEK293T	Embryonic kidney epithelial			
	Laurent *et al.*	HEK293 Flp-In T-REx	Embryonic kidney epithelial			
	May *et al.*	A549	Alveolar basal epithelial			
Virus–host PPI (AP-MS)						
(5 studies/5 datasets)	Gordon *et al.*	HEK293T	Embryonic kidney epithelial	NA	NA	Significant interactors per authors
	Stukalov *et al.*	A549	Alveolar basal epithelial			
	Zhou *et al.*	Caco2	Colon epithelial			
	Li *et al.*	HEK293 Flp-In T-REx	Embryonic kidney epithelial			
	Liu *et al.*	Tet-inducible HEK293T	Embryonic kidney epithelial			
RPI—ChIRP-MS, vRIC, RAP-MS						
(3 studies/5 datasets)	Flynn *et al.*	Huh7.5	Liver epithelial-like	0.01	24, 48	$\|\log_2 FC\| > 0$ and FDR $P < 0.05$
	Kamel *et al.*	Calu3	Bronchial epithelial	1	8, 24	$\|\log_2 FC\| > 0$ and FDR $P < 0.05$
	Schmidt *et al.*	Huh7	Liver epithelial-like	10	24	$\|\log_2 FC\| > 0$ and FDR $P < 0.2$

This table includes experimental methods and tissue/cell type used as well as the multiplicity of infection (MOI) and hours post-infection (HPI). FDR, false discovery rate.

hierarchical clustering based on Euclidean distances. GO: biological process and Reactome enrichment tests were conducted using the clusterProfiler (version 4.2.2) (74) and ReactomePA (75) (version 1.38.0) packages, respectively. For functional enrichment tests on predicted SL targets derived from our pipeline, the background gene set was limited to the complement of genes in the SL pair database used for prediction. The Benjamini–Yekutieli method was used to FDR-correct enrichment test *P*-values (76). For all datasets analyzed in our study, gene and gene product identifiers were aligned to a reference list of gene symbols using the org.Hs.eg.db R package (77, https://bioconductor.org/packages/org.Hs.eg.db/).

Results

SL target prediction pipeline

Numerous mechanisms exist by which viruses may induce hypomorphic states within host cells that may introduce SL relationships. These include modifications to virus–host PPIs, transcription, protein abundance, PTMs, protein repositioning and proteolytic cleavage, among others (21,22). Consequently, a myriad of omics-level data acquisition techniques could potentially facilitate the identification of VIHs and subsequently their SL partners, which could serve as host-based therapeutic targets. To examine this, we developed a computational pipeline to integrate data from various omics data sources, identify potential VIHs, and predict their prospective SL partners as antivirals (Figure 1B). Here, we define VIHs as gene products with reduced levels as evidenced by transcriptomics, translatomics or proteomics datasets, as well as gene products with altered states as evidenced by phosphoproteomes, ubiquitomes, RNA or protein interaction datasets, and others that might result in altered protein function. For each omics data category, we selected individual datasets from as many applicable studies as were available in the literature, which led to the inclusion of 1–14 datasets per omics data class (Table 1). We aimed to encompass a large variety of infection conditions (varying cell lines, virus strains, infection multiplicity and timeline) cited in the literature and thus adopted an aggregate approach incorporating all available datasets within each data class. Following this, we applied a filtering methodology to identify the candidate SL genes with the highest likelihood of successful experimental validation (Figure 1).

Initial identification of putative VIHs was achieved by finding molecular species (transcripts, protein abundance or PTMs) that exhibited significant alterations following SARS-CoV-2 infection. Subsequently, the SL partners of these VIHs were identified within human SL pairs represented in the SynLethDB resource (54), a public repository of SL pairings developed for cancer research, as well as in SynLEGG (55), a database of SL pairs predicted on the basis of transcriptomics and CRISPR KO screen data from almost 800 cancer cell lines. Since genome-wide CRISPR KO screens in the setting of viral infection should contain SL genes due to reduced viability of cells with SL gene KOs (1), we used depleted CRISPR KO gene pools to quantify the power of each omics data class for predicting SL targets. These measures included enrichment of candidate SL targets as well as specificity, sensitivity and precision statistics. Finally, we used the same depleted CRISPR KO gene pool as a filter to further prioritize candidate SL targets. This step also filtered out genes in the CRISPR KO-depleted gene set that might not be in SL relationships such as antiviral genes, by virtue of not being an SL partner of a VIH, and genes that are essential for cell growth and survival based on cell death in uninfected cells. The remaining proteins represent candidate SL targets that have led to decreased cell viability when knocked out during viral infection compared to infection alone in CRISPR-based KO screens in at least one published viral infection scenario.

We are aware that shortcomings exist in all three main data sources: (i) altered gene products might still retain normal functionality; (ii) SL pairs in the combined SL pair database with human SL pairs from SynLethDB and SynLEGG might have relevance in certain cancer contexts but may not form SL relationships in the virus-infected cellular contexts we tested; and (iii) some antiviral genes might be present in the CRISPR-depleted genes and thus not represent SL interactors of VIHs. However, our overall rationale for this pipeline is that by combining information from the SL databases, the VIHs identified from viral infection omics studies, and the CRISPR KO screens on virus-infected cells, we strengthen the prediction of virus-specific SL targets: (i) a VIH with an SL partner in our SL pair database representing known or predicted SL interaction is more likely to form an SL interaction than one that is not; (ii) an SL gene that is partnered with a VIH in virus-infected cells is more likely to be virus-relevant than one that is not; (iii) an SL gene from a known SL interaction that is also depleted in a CRISPR KO screen in virus-infected cells is more likely to be virus-relevant than one that is not; and (iv) a CRISPR KO-depleted gene in a virus-infected cell is more likely to be in an SL relationship than one that is not depleted or is enriched.

Data selection

We curated datasets from published research focusing on SARS-CoV-2 and influenza viruses due to their ongoing global health and pandemic implications. Given the intensive research on these viruses and their relatives, many datasets describing various aspects of gene product functionality during orthomyxovirus and coronavirus infections are available, which is ideal for an exhaustive exploration of the most accurate predictors for virus-specific SL targets from a practical perspective. In total, we compiled 56 omics datasets from 22 distinct studies on SARS-CoV-2 infection, which analyzed a wide spectrum of cellular changes (Table 1). The datasets depict host interactors of viral proteins and RNA (32,42,44–53,78), infection-induced alterations in host gene expression (31–33,35,79) and protein abundance (32,36–38,40), changes in translation rates (40), and alterations in post-translational modifications (32,39), cleavage (41) and the RBP (51), comprising 10 VIH data classes in total. Host interactors of viral proteins were further split based on technology: AP-MS or BioID-MS (Table 1). Both techniques used individual viral open reading frames (ORFs) as baits to identify interacting host proteins. AP-MS was predominantly employed to capture virus–host multiprotein complexes (80) whereas BioID-MS (81) serves as a complementary method to examine transient interactions based on physical proximity, a frequent occurrence throughout the viral infection cycle (82).

SL target prediction

Identification of VIHs and their SL partners by analyzing omics data from virus-infected cells

Our pipeline's initial step (Figure 1B) involves the identification of potential VIHs by selecting gene products

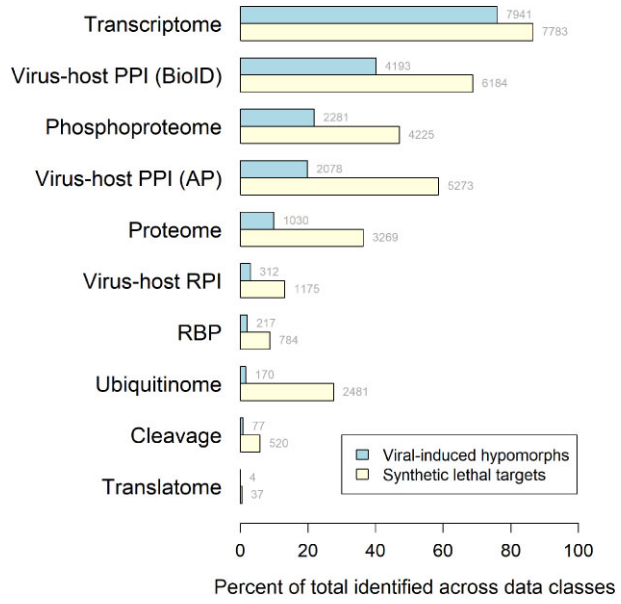


Figure 2. Percentage of predicted VIHs and SL targets for individual omics data classes. For each class, bar width indicates the percentage of VIHs (blue, upper bar) and SL targets (yellow, lower bar) predicted out of the combined set of all predictions across data classes. The numbers at the ends of the bars indicate the absolute number of VIHs and SL targets identified.

significantly impacted by viral infection, as evident from the 10 omics data classes outlined in Table 1. For VIH identification, we relied on the significance criteria stipulated in each dataset's corresponding publication cited here (31–33,35–42,44–51,53,78,79,83). We subcategorized VIHs based on the increase or decrease in molecular species abundance post-infection for transcriptomics, proteomics, RNA-binding proteomics and translatomics data. For phosphoproteome, ubiquitome and protein cleavage data, we ascertained the total absolute change for each protein based on the aggregate changes reported for all pertinent peptides.

Following VIH identification from each data class, we narrowed down the VIHs to human gene products that were also present in the commonly utilized SynLethDB version 2.0, and then identified the SL partners of putative VIHs (Figure 1B and Table 2). SynLethDB (54) comprises a collection of SL pairs procured from various sources: confirmed SL pairs from the literature based on genome-wide small interfering RNA (siRNA), CRISPR KO screens, bispecific SL short hairpin RNA screens, predictions from computational algorithms, text mining and other databases. The SynLethDB version we accessed for this study contains 35 943 human candidate SL pairs, encompassing 9855 unique candidate SL genes. We also utilized the SynLEGG database, containing 107 545 predicted SL gene pairs. SynLEGG had minimal overlap with SynLethDB SL pairs, adding nearly 7000 unique candidate SL genes to our combined SL pair database. In this database, each gene, on average, is in an SL relationship with three other genes.

Across all data classes, we identified a total of 10 454 unique VIHs and 6259 unique SL partners for those VIHs (Table 3). The number of unique VIHs and SL partner targets identified using each data class differed substantially (Table 3 and Figure 2). Transcriptomic, PPI and proteomic studies yielded the highest numbers of VIHs and SL partners, whereas translatome and cleavage studies produced the fewest. We found significant correlations between the number of datasets in a class and the total number of predicted VIHs (Pearson coefficient = 0.74, *P*-value = 0.01) as well as their predicted SL partners (Pearson coefficient = 0.71, *P*-value = 0.02). As expected, this confirmed the poor overlap observed between individual datasets within each data class, resulting from the wide range of infection scenarios included. The high numbers of predicted VIHs stem from the fact that the union of unique genes from two datasets is higher when the datasets have little overlap as compared to datasets that have large intersects. For example, no VIHs consistently appeared in all 14 of the transcriptomics datasets used. The maximum number of datasets in which a VIH consistently appeared was 9 and the median was 2. We chose the union of VIHs in each data class, as this approach offered a more comprehensive list of candidate SL targets from a broader range of relevant infection scenarios for further evaluation.

Performance of omics data classes in predicting valid SL targets

A central aim of our study was to evaluate the predictive accuracy of each data class in identifying legitimate, experimentally verified SL targets. Unfortunately, there is no universally accepted 'gold standard' for validating SL targets. To create a representative pool of SL targets that could serve as a surrogate gold standard bolstered by adequate experimental evidence, we aggregated depleted genes from multiple, diverse CRISPR KO screens (Supplementary Table S1). However, we note the unknown percentage of false negatives associated with antiviral activity and potentially some essential genes associated with these data. Moreover, there were no genes that exhibited universal depletion across all CRISPR KO datasets, potentially due to variable cell type selection among other experimental variables. As with our omics datasets, we opted to pool depleted genes from multiple disparate CRISPR screens to ensure representation of diverse SARS-CoV-2 infection scenarios. With these caveats, we used the list to assess the performance of each data class for predicting SL targets in SARS-CoV-2 infection. Performance measures included hypergeometric enrichment tests as well as sensitivity, specificity and precision metrics.

Sensitivity and specificity scores spanned a broad range across data classes, with the highest sensitivity observed in the transcriptome, virus–host PPI and proteome classes and the highest specificity within the translatome, RBP and cleavage data classes (Supplementary Figure S1). Our analysis showed significant enrichment of SL targets among genes depleted in CRISPR KO studies, regardless of the employed omics data class. FDR-adjusted *P*-values ranged from 0.001 (translatome) to 3.44E−60 [virus–host PPI (BioID)]. However, we also found that the complement of genes in SynLethDB itself is highly enriched for genes depleted in CRISPR KO studies (hypergeometric test *P*-value = 3.75E−63). Therefore, to compare data class predictive capabilities in a way that accounts for the existing enrichment of SynLethDB for CRISPR KO-depleted genes, we performed a randomization test to determine the level of enrichment one would expect by chance from each omics class (Figure 3). Thus, for each omics data class, we randomly selected *N* genes from SynLethDB, where *N* is the number of VIHs found in SynLethDB for the data class. We then collected the SL partners of the randomly selected

Table 2. CRISPR KO datasets used for validating predicted SL targets

Study	Cell line	Library	Virus strain	MOI	DPI	*N* depleted gene products
SARS-CoV-2						
Biering *et al.*	Calu3ACE2/TMPRSS2	Brunello	USA-WA1/2020	0.05	4	0
Baggen *et al.*	Huh7	Brunello	USA-WA1/2020	2.5, 0.1	41, 22	0
Daniloski *et al.*	A549ACE2	GeCKOv2 A&B	USA-WA1/2020	1	6	0
				3	6	0
Rebendenne *et al.*	Caco2ACE2	Brunello	BetaCoV/France/IDF0372/2020	0.005	10–13	206
	Calu3	Gattinara	BetaCoV/France/IDF0372/2020	0.005	10–13	0
Schneider *et al.*	Huh7.5 at 33°C	Brunello	USA-WA1/2020	0.01	5	63
	Huh7.5 at 37°C			0.1	5	73
Wang *et al.*	Huh7.5ACE2/TMPRSS2	GeCKO	USA-WA1/2020	0.01	12	3
Zhu *et al.*	A549ACE2	Brunello	SARS-CoV-2/human/CHN/SH01/2020 (nCoV-SH01)	3	7+	12
Hou *et al.*	A549ACE2	H1	Rec. 2019-nCoV/USA_WA1	5	2	84
Grodzki *et al.*	HEK239ACE2—initial infect	Brunello	UF-1	0.01	(40% CPE)	627
	HEK239ACE2—reinfect			0.01	(40% CPE)	
	HEK239ACE2—reinfect			0.1	(40% CPE)	
Israeli *et al.*	Calu3	Brunello	B.1.1.7 (Alpha), B.1.351 (Beta), B.1.617.2 (Delta)	0.02–0.04	7–9	7
Hoffmann *et al.*	Huh7.5 at 33°C	Custom[a]	USA-WA1/2020	0.25	7	19
	Huh7 .5 at 37°C			0.25	11	21
Influenza[b]						
Han *et al.*	A549	GeCKO	H5N1 (VNO4LOW)	5	2 (rd 2)	0
				5	2 (rd 5)	0
Yi *et al.*	A549	GeCKO	H7N9 H7N9 (A/chicken/Guangxi/97/2017)	0.01	10 (after 2 rds)	26
Tran *et al.*	A549	GeCKO	H1N1 (IAV-G, pseudotyped A/WSN/1933 strain)	0.3	20–21 (rd 1)	1
		GeCKO	H1N1 (A/WSN/1933)	0.3	40–42 (rd 2)	2
		Brunello	H7N7 (SC35M Flu-GFP; A/seal/Massachusetts/1-SC35M/1980)	0.1	63 (after 7 rds)	1669

Significance cutoffs: $\log_2$ fold change <0, FDR < 0.05 (Biering, Baggen Daniloski, Wang, Zhu, Hou, Grodzki, Israeli) or *z*-score <−1.96 (Rebendienne, Schneider, Hoffman). MOI, multiplicity of infection; DPI, days post-infection; rd(s), round(s) of infection.

[a]'Human SARS-COV-2 host protein interactome CRISPR library' including single-guide RNAs (sgRNAs) targeting 332 host factor proteins interacting with SARS-COV-2 proteins and control sgRNAs.

[b] Influenza virus-infected cells were generally allowed to expand after each round of infection.

Table 3. Number of VIHs identified among omics data classes, VIHs in SynLethDB and SynLEGG, and predicted SL partners

Data class	*N* VIHs	*N* VIHs in SynLethDB	*N* SL partners of VIHs (SynLethDB)	*N* VIHs in SynLethDB + SynLEGG	*N* SL partners of VIHs (SynLethDB + SynLEGG)
Cleavage	77	60	520	68	2643
Phosphoproteome	2281	1633	4225	1940	7619
Proteome	1030	662	3269	810	5790
RBP	217	175	784	191	2120
Transcriptome	7941	4663	7783	5798	10 802
Translatome	4	3	37	3	74
Ubiquitinome	170	131	2481	148	3672
Virus–host PPI (AP)	2078	1420	5273	1729	8403
Virus–host PPI (BioID)	4193	2877	6184	3494	9440
Virus–host RPI	312	251	1175	273	4063
Total unique	10 454	6259	8993	7784	11 673

Data classes are ordered alphabetically.

genes and computed hypergeometric *P*-values to test their enrichment for CRISPR KO-depleted genes. This procedure was repeated 10 000 times for each data class to build null distributions of enrichment *P*-values. We then computed percentile scores for each data class's actual enrichment *P*-value relative to the corresponding null distribution (Figure 3). We found that interactomic data classes (virus–host PPI, virus–host RPI) along with the cleavage data class were most significantly distinct from the null distributions (>95th percentile), indicating that SL targets derived from these classes were significantly enriched and the best predictors of SL targets.

We also examined precision results from the same randomization test and found that across all randomization iterations, predictive precision using SynLethDB was low. Median values for each data class's null distribution of precision scores were 0.13 or below, with the third quartile not exceeding 0.2. This indicates that, across the different omics data classes, statistical precision for predicting CRISPR-depleted genes in a SARS-CoV-2 context is limited by a high proportion of false positives that arise when using SynLethDB to select SL partners from VIHs. This is most likely because SynLethDB contains SL pairs derived from cancer research, and many of those relationships are not applicable in the context of viral infection. Thus, the low statistical precision we observed for each omics data class may not reflect issues with the data derived from the data classes themselves but rather from false positive predictions inherent to using SynLethDB in this research context. We also tested whether limiting SynLethDB pairs to those with confidence scores ≥0.5 or 0.8 would enhance precision but found that, even at the 0.8 cutoff, the highest precision value for a data class was 0.28, and increases in precision came at the cost of reduced sensitivity, which averaged 0.11 across classes. Therefore, to accurately predict SL targets from any omics data class investigated here, it is essential to further filter candidates using data from studies that validate the impact of target disruption on cell viability, such as CRISPR KO and siRNA data, to reduce false positives.

The inclusion of antiviral genes that are not in an SL relationship with our 'gold standard' CRISPR KO-depleted genes likely contributes to the observed low precision. To further understand this, we cross-referenced a list of 77 known antiviral proteins from DRAVP (84), a comprehensive database of antiviral peptides and proteins based on the literature, with our SARS-CoV-2 and influenza candidate SL targets. Only a small percentage, 0.8% and 0.5%, of the SARS-CoV-2 and influenza candidate SL targets, respectively, matched known antiviral genes, confirming that our pipeline generally omits antiviral genes when predicting SL targets.

We also investigated whether the predictive accuracy of each omics data class correlated with the number of datasets compiled for that class. Using results from the randomization test described above, we saw no significant correlations between the number of datasets in a data class and percentile scores for enrichment, sensitivity, specificity or precision (data not shown).

Discovering that the highest predictive sensitivity across all omics data classes was 0.61 for transcriptome, we aimed to boost the sensitivity of our pipeline by including additional SL pairs from the SynLEGG database. Unlike SynLethDB, the SL pairs in SynLEGG are derived from genes paired based on low expression and CRISPR KO depletion in 783 diverse cell lines without prior selection for cancer-related genes (55). SynLEGG depends upon an algorithm (MultiSEp) for unsupervised assignment of cell lines into gene expression clusters, which provides the basis for analysis of CRISPR scores to discover mutually exclusive loss signatures and propose candidate SL pairs. Using the combined set of SL pairs from SynLethDB and SynLEGG, we increased the overall sensitivity of our pipeline: The count of unique genes predicted as SL targets by at least one omics data class and also depleted in CRISPR studies rose by 16%, from 664 to 772. This increase in sensitivity was accompanied by a small drop in average predictive precision across data classes from 0.11 to 0.09 and a decrease in average specificity from 0.86 to 0.75 (Table 4 and Supplemental Figure S2). The remaining analyses and results reported here are derived from the version of our pipeline that uses the merged SynLethDB and SynLEGG SL pairs.

Upon examining the pattern of positive predictions for candidate SL genes that were also CRISPR KO-depleted SL targets, we identified two distinct data class clusters: one cluster for the groups of candidate SL targets predicted by the transcriptome, proteome, phosphoproteome, ubiquitome and virus–host PPI omics data classes, and another cluster that includes the virus–host RPI, RBP, cleavage and translatome data classes (Figure 4). The two primary clusters share overlapping SL targets but differ in the number of SL targets predicted. Overall, the cluster consisting of transcriptome, virus–host PPI, phosphoproteome, proteome and ubiquitinome classes

Table 4. Metrics of performance among omics data classes for predicting valid SL targets against SARS-CoV-2 infection

Data class	*N* candidate SL targets also CRISPR-depleted	Sensitivity	Specificity	Precision
Cleavage	184	0.19	0.88	0.07
Phosphoproteome	542	0.56	0.65	0.07
Proteome	480	0.50	0.74	0.08
RBP	222	0.23	0.91	0.10
Transcriptome	716	0.75	0.50	0.07
Translatome	11	0.01	1.00	0.15
Ubiquitinome	307	0.32	0.83	0.08
Virus–host PPI (AP)	620	0.65	0.62	0.07
Virus–host PPI (BioID)	676	0.70	0.57	0.07
Virus–host RPI	317	0.33	0.81	0.08
Total unique	772			

Table 5. Metrics of performance for predicting valid SL targets against influenza infection using omics data from SARS-CoV-2 infections

Data class	*N* candidate SL targets also CRISPR-depleted	Sensitivity	Specificity	Precision
Cleavage	235	0.14	0.87	0.09
Phosphoproteome	694	0.41	0.62	0.09
Proteome	514	0.31	0.71	0.09
RBP	185	0.11	0.89	0.09
Transcriptome	965	0.57	0.46	0.09
Translatome	3	0.00	1.00	0.04
Ubiquitinome	325	0.19	0.82	0.09
Virus–host PPI (AP)	744	0.44	0.58	0.09
Virus–host PPI (BioID)	857	0.51	0.53	0.09
Virus–host RPI	345	0.20	0.80	0.09
Total unique	1063			

predicted a greater number of valid SL targets than the second cluster consisting of virus–host RPI, RBP, cleavage and translatome data classes. While this second cluster identified a smaller number of candidate SL targets (lower sensitivity), its members predicted a higher proportion of true positives (higher precision) (Table 4).

Our pipeline identified 772 genes as high-confidence SL targets (Table 4, Figure 4 and Supplementary Table S2), supported by three crucial lines of evidence: (i) CRISPR KO screen data from virus-infected cells, highlighting that disrupting these genes impairs cell survival under such conditions; (ii) data indicating a negative genetic relationship with gene products dramatically altered by viral infection; and (iii) indications of their involvement in SL relationships in the context of cellular changes due to cancer. The remaining 188 CRISPR-depleted genes (Table 4 and Figure 4), not flagged as SL partners by our pipeline, may reflect host dependency factors not identified as VIHs by the omics data classes we investigated, or they may reflect SL relationships that remain undiscovered.

Functional profiling of high-confidence SL targets

We performed GO:biological process and Reactome enrichment analyses to identify the cellular pathways and functions associated with the 772 high-confidence SL targets mentioned above. There were 137 GO:biological process classes enriching for the SL targets, and they were primarily associated with RNA processing (including RNA splicing/processing as well as ribosomal RNA- and noncoding RNA-related processes), ribosomal biogenesis, mitochondrial gene expression and translation, aerobic cellular respiration, metabolism, DNA replication, chromosome organization, various DNA repair processes and the cell cycle (Supplementary Table S3). Significantly enriched biological processes ($P < 0.05$) also included virus-associated terms such as virion assembly and viral budding via the endosomal sorting complex required for transport (ESCRT) complex. Reactome pathways showing enrichment ($N = 123$) were primarily associated with translation, mitochondrial processes, RNA splicing/processing, transcription including transcriptional termination, the cell cycle, DNA repair, stress responses, metabolism, ribosomal pathways, DNA replication, chromosome organization, p53 regulation, nuclear envelope reassembly and cytosolic iron–sulfur cluster assembly (Supplementary Table S4). Virus-associated pathways were also significantly enriched ($P < 0.05$) and included viral life cycle, viral transcription, translation and replication, as well as maturation and budding of virus.

Overall, these findings align well with those of Pal *et al.* in associating cellular stress responses, RNA splicing/processing, DNA repair and translation with SL targets (27). However, because our results are based on larger and more diverse datasets, they revealed a broader array of biological functions associated with SL targets in the context of SARS-CoV-2 infection and, importantly, included multiple processes corresponding to viral infection themes that included the entire viral life cycle.

Identification of common SL targets for SARS-CoV-2 and influenza viruses

To investigate the existence of pan-viral SL targets, we explored whether some SL targets predicted from SARS-CoV-2 omics data might also be present in CRISPR-depleted gene sets from cells infected with other viruses. We selected the influenza virus as a contrasting global viral threat, noting its distinct features from coronaviruses: influenza viruses use sia-

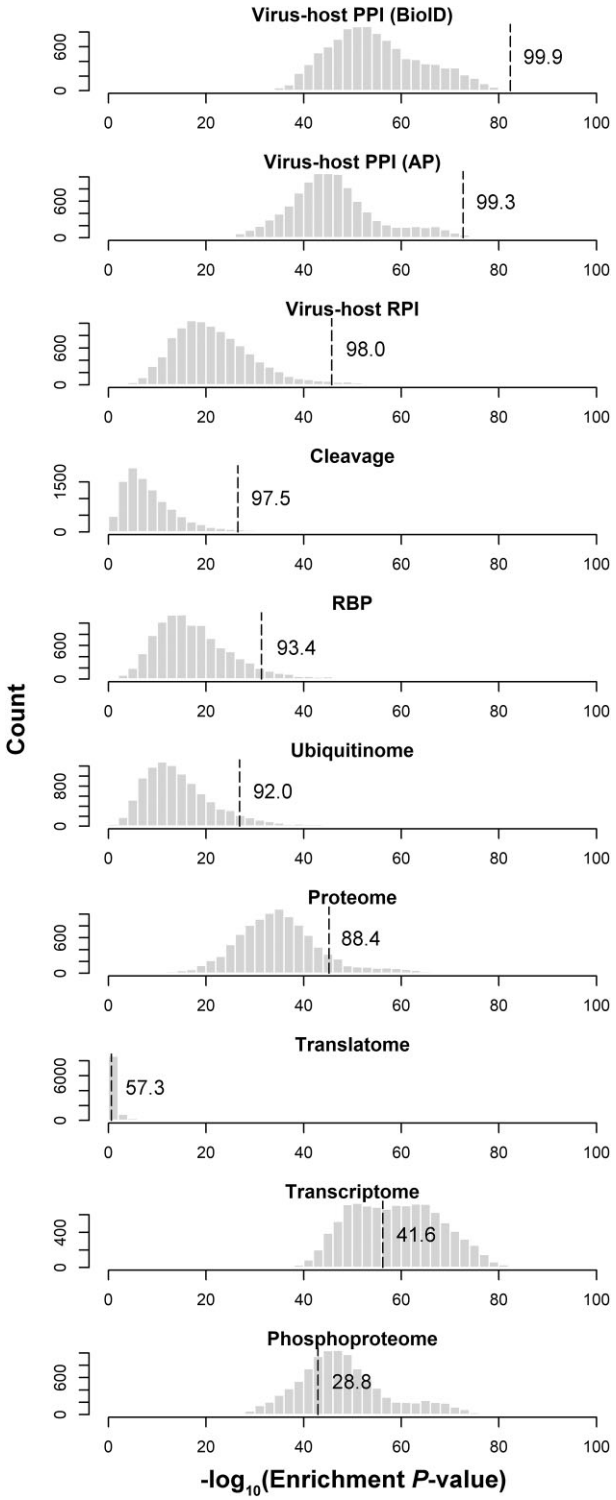


Figure 3. Null distributions of scores from tests quantifying enrichment of SL targets in CRISPR KO-depleted genes. Histograms indicate the distribution of hypergeometric test *P*-values expected by chance for each omics data class. The dashed line indicates where the omics class's actual enrichment test value fell. The number to the right of the line is the percentile of that test with respect to the null distribution.

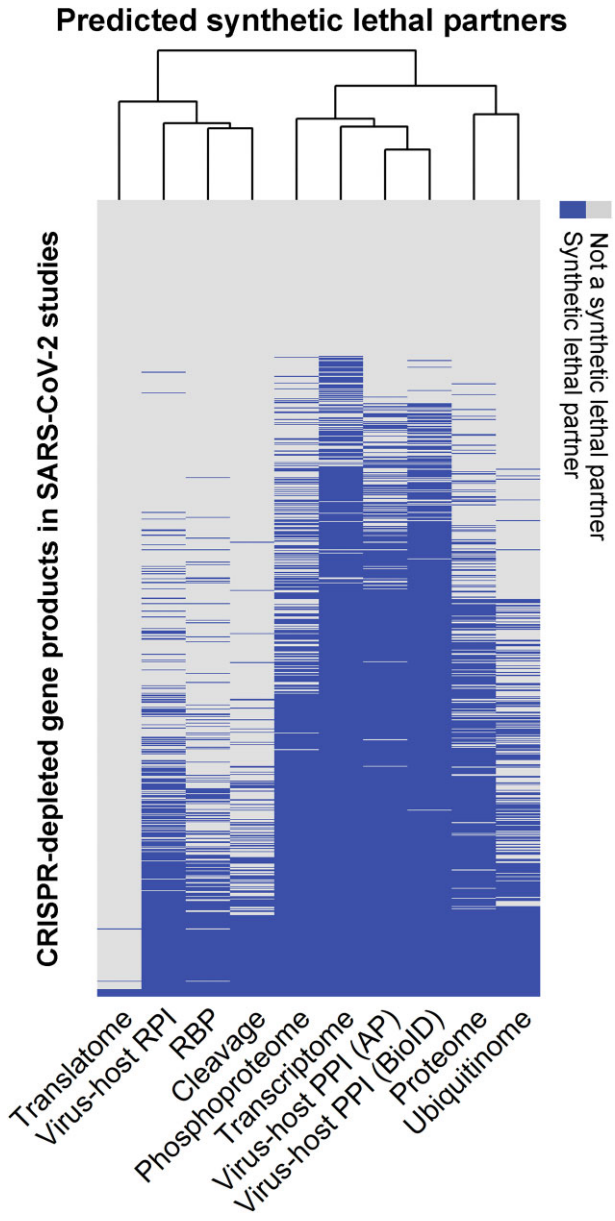


Figure 4. Candidate SL targets across omics data classes depleted in SARS-CoV-2 CRISPR KO studies. Binary heatmap showing which genes in the SARS-CoV-2 CRISPR-depleted gene pool (rows) were SL partners of VIHs and thus predicted to be SL targets for each omics data class. Predicted targets are indicated in dark blue (refer to Supplementary Tables S1 and S2 for gene names). Light gray cells indicate CRISPR-depleted gene products that are not SL partners of VIHs and are thus not predicted to be SL targets. Rows are sorted by number of omics data classes predicting the product as an SL target.

lyloligosaccharides for cell entry, have a segmented genome and replicate in the host cell nucleus (85) in contrast to the continuous double-stranded RNA genome and cytoplasmic replication of coronaviruses (82). We compiled 1684 unique genes significantly depleted in three influenza CRISPR KO studies (Supplementary Table S5) and assessed their prediction as SL targets using our pipeline. This comparison showed that precision remained uniformly low, akin to the SARS-CoV-2 context, with sensitivity and specificity values suggesting no greater accuracy than random chance (with the sum of these values across data classes approximately 1.0) (Table 5).

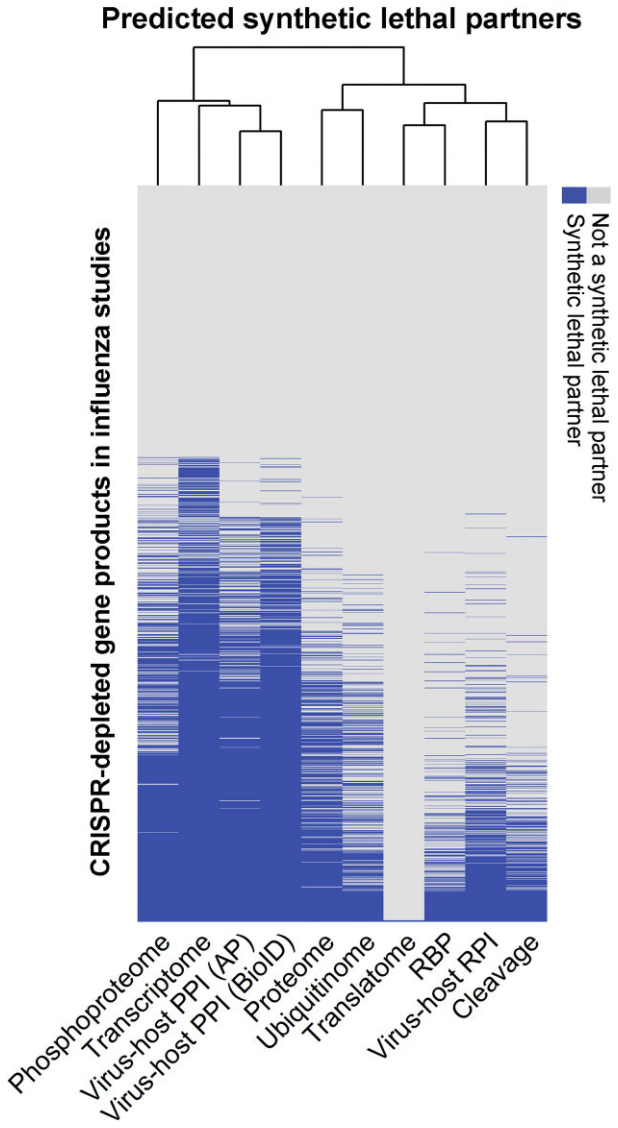


Figure 5. Candidate SL targets across omics data classes depleted in influenza A CRISPR KO studies. Binary heatmap showing which genes in the influenza CRISPR-depleted gene pool (rows) were predicted to be SL targets for each omics data class. As in Figure 4, predicted targets are indicated in dark blue (refer to Supplementary Tables S5 and S6 for gene names). Light gray cells indicate CRISPR-depleted products not predicted to be an SL target. Rows are sorted by number of data classes predicting the product as an SL target.

However, this analysis revealed that more than half of the influenza CRISPR KO-depleted genes were predicted as SL targets using the SARS-CoV-2 omics data-based predictions (Figure 5, Supplementary Table S6). This substantial overlap led us to hypothesize that a set of SL targets common to both viral infections existed. Examining the overlap between significantly depleted genes in SARS-CoV-2 and influenza CRISPR KO data (Supplementary Tables S1 and S5), we identified 78 shared genes, of which 61 (78%) were predicted as SL targets (Supplementary Table S7). These 61 genes potentially represent broad viral SL targets. Intriguingly, a GO:biological process enrichment test found no significantly enriched processes among these genes using an FDR-correction cutoff of 0.1. A Reactome enrichment test identified one significantly enriched pathway: translation (FDR-adjusted P-value = 0.09).

These 61 broadly antiviral candidate SL targets need to be further scrutinized and prioritized before undergoing experimental testing. Notably, only one of our broadly antiviral candidates was also predicted by Pal *et al.* as an SL target for SARS-CoV-2: poly(RC) binding protein 1 (PCBp1), which is crucial for pre-mRNA processing, translation, ferroptosis and the initiation of viral replication and viral translation (86).

A key consideration for deprioritization includes high functional interconnectivity, which risks off-target effects, as identified in the MoonDB database (71). Additionally, genes essential in a specific or multiple cell lines should be avoided, as their disruption could kill both infected and uninfected cells (72). The presence of numerous paralogs in candidate SL genes also demands careful consideration due to potential functional redundancy hindering effective synthetic lethality. Interestingly, gene pairs that are paralogs have been found to engage in SL relationships more often than by chance (87). The prioritization of VIH–SL pairs can consider these aspects, with final candidates selected for their druggability and the availability of existing drugs or ligands (Table 6).

Discussion

This study seeks to extend previous research establishing that viruses expose SL relationships during infection, and to identify VIH-SL partnerships that may be exploited in the search for host-based antiviral therapeutics. An important step in this effort is the establishment of a computational pipeline for predicting SL targets. Since synthetic lethality is a type of genetic interaction unexplored in infectious diseases, such a pipeline, once fully established, can expand the number of antiviral drug targets and add to the knowledgebase of pathogen biology. Our goal is for this pipeline to generate SL targets to be examined, prioritized and experimentally validated.

Our pipeline generated a list of potential SL targets that are enriched with genes depleted in CRISPR KO studies specific to SARS-CoV-2 infection. These targets were predicted by using an array of available omics datasets, which were specific for virus-infected cells, in combination with two established SL databases. One assumption we made was that the SL partners of gene products undergoing significant changes in virus-infected cells would be depleted in CRISPR KO screens, thus potentially serving as SL targets. Our results supported this assumption, notwithstanding the fact that the total number of SL partners of VIHs encompassed almost all SL genes (11 673/12 348) represented in SynLethDB and SynLEGG due to multiple SL partners per VIH. This confirms what has become evident from studies in yeast, namely that almost every gene can be in an SL partnership (88). An approach that would further aid in the prediction of SL targets from CRISPR KO screens is the addition of viral replication measurements in addition to cell viability, a practice not typically adopted for genome-wide functional screens.

We also investigated whether different classes of omics data would differ in their ability to accurately predict SL targets by measuring whether SL prediction by some data classes was better supported by CRISPR KO screens than others. Exploring various omics classes, each interrogating different facets of host–virus biology, we observed that despite varying VIH predictions, there was substantial overlap in the SL vulnerabilities predicted from different omics classes. While we have to take into consideration the unifying effect of filtering our

Table 6. Characteristics of candidate pan-viral SL targets potentially useful for prioritizing their confirmation

Gene	EMF gene	Essential in *N* cell lines	*N* paralogs	Targeting drug (DrugBank ID)
ACAD9	No	5	14	
ATXN7L3	No	3	4	
CHERP	Yes	9	0	
CHMP2A	No	10	4	
COX4I1	No	9	1	DB02659; DB04464
CPSF2	No	10	3	
DDX11	No	2	3	
DDX49	No	10	38	
DNAJC19	No	1	1	
DPF1	No	0	0	
DRAM1	No	0	4	
DUX4	No	0	0	
EEF2	No	10	18	DB02059; DB03223; DB04315; DB08348; DB11823; DB12688
EIF2S3	No	8	18	DB04315
F2R	No	0	16	DB00086; DB05361; DB09030; DB11300
FDX2	No	0	1	
GOLGA8O	No	0	18	
GPN2	No	9	2	
ISCA2	No	5	1	
KDM2A	No	4	4	
LIG3	No	2	2	DB00290
LRP1	No	0	0	DB00025; DB00031; DB00100; DB06245; DB13152; DB13998; DB13999
MAOA	No	0	0	DB00191; DB00752; DB00780; DB00805; DB00909; DB01037; DB01171; DB01247; DB01363; DB01381; DB01442; DB01472; DB01577; DB01626; DB03147; DB04017; DB04820; DB04821; DB04832; DB04850; DB05205; DB07641; DB07919; DB09244; DB09245; DB13876; DB14914;DB00182; DB00721; DB01168; DB01171
MRPL43	No	8	0	
MRPS14	No	9	0	
MRPS21	No	5	0	
MTIF2	No	7	18	
MYBBP1A	No	10	0	
NOL10	No	10	0	
P4HB	No	0	13	DB01593; DB03615; DB09130; DB11638; DB14487; DB14533; DB14548
PCBP1	No	9	12	
PDE4DIP	No	9	1	
PDPK1	Yes	3	5	DB00482; DB01863; DB01933; DB01946; DB02010; DB03777; DB04522; DB06932; DB07033; DB07132; DB07300; DB07456; DB07457; DB12010
PLEKHF2	No	0	13	
POLR2A	No	10	2	
POLR2J	No	0	3	
PPP2R1A	No	9	2	DB02506; DB06905
PTRH1	No	1	0	
PWP2	No	10	0	
RAD51	No	10	6	DB04395; DB12742
ROS1	No	0	53	DB11986; DB12010
RPL28	No	0	0	
RPL3	No	8	1	DB02494; DB04865; DB07374; DB08437; DB09092
RPL36	No	7	0	
RUVBL1	No	10	1	
SBNO1	No	9	1	
SDHA	No	8	0	DB04141; DB04657; DB04795; DB08689; DB09270
SERINC3	No	0	4	
SLX4	No	1	0	
SMC2	No	10	7	
SMC5	No	8	1	
TAF6	No	10	1	
TCIRG1	No	0	0	DB01133
TP53RK	No	9	0	
TPT1	No	7	0	DB11093; DB11348; DB14481
TRRAP	No	10	5	
TUBB4B	No	0	23	DB00518; DB00643; DB01873; DB03010; DB04910; DB05147; DB12695
TXN	No	5	4	DB12695
UBE2I	Yes	9	24	
VPS25	No	10	0	
WAPL	No	0	0	

EMF, extremely multifunctional (MoonDB 2.0). The 17 bolded genes are not EMF, are essential in ≤3 cell lines and have ≤5 paralogs. However, we recommend that each gene be examined on a wider range of merits and/or by possible ranking algorithms, before testing on the bench. While we consider these to be important characteristics, no characteristic alone indicates a high-level SL target.

data through SynLethDB, this may also indicate that common biological pathways and functions may be altered upon SARS-CoV-2 infection rendering them vulnerable. Our findings demonstrate that valid SL targets can indeed be derived from different omics data types. Moreover, our data show that predictions from host proteins that interact directly with viral proteins or viral RNA outperform other omics data classes in terms of enrichment for valid SL targets when controlling for the unifying effect of SynLethDB. This suggests that the hypomorphic effects of viral infection may be strongest for proteins that physically interact with viral gene products, which may lead to disruption of host protein complexes, interruption of cofactor binding or mislocalization. Many direct interactors have also been shown to be host dependency factors and thus are essential for viral replication. Furthermore, we speculate that omics datasets describing other ways in which viral infection alters normal gene and protein functions, such as changes in RNA splicing or additional post-translational modifications, might also be beneficial for SL prediction, but they are currently fewer in number.

Analyzing the high-confidence SL targets predicted by different omics classes, we found that each data class varied in sensitivity and specificity in predicting SL targets based on functional assays such as CRISPR KO screens (Supplementary Figure S1), indicating that some omics data classes might be more suitable than others for accurate SL prediction. Our randomization analysis confirms this and highlights host–virus interactions as the best data classes for SL prediction. We also discovered that genes in existing SL databases are themselves enriched for depleted genes in CRISPR KO studies. However, we also found that using the databases to predict SL targets led to low statistical precision across all omics data classes (Table 4). This finding emphasizes the need to reduce false positives when predicting SL targets. An important way to do so is to limit predicted targets to those found depleted in CRISPR KO studies.

Recent findings by Pal *et al.* (27) confirmed that transcriptomics data alone could be used to predict SL target genes leading to reduced viral replication and cell death in SARS-CoV-2-infected cells when depleted. The authors reported differences in cell viability between infected and uninfected cells for several of their top 26 candidate SL targets. We note substantial overlap in enriched cellular pathways and functions between our study and theirs, but also identified additional pathways that may become vulnerable in the virus-infected cellular state, such as aerobic respiration, RNA splicing, transcription, p53 regulation, ribosomal pathways, DNA replication, chromosome organization and iron–sulfur cluster assembly. This last process involves one of the oldest classes of bioinorganic cofactors that support and modulate innate immune responses restricting viral replication, and some of the proteins interacting with Fe–S clusters also support replication of viruses (89). Six of the proteins involved in such clusters were predicted as SL partners in SARS-CoV-2 infection (ERCC2, RTEL, CIAO1, CIAO3, NUBP1, NUBP2) by at least one and up to six omics classes in our screen and might thus be worthwhile host targets for follow-up targeted KO studies in infected cells. Additionally, our 772 high-confidence SL targets were enriched for multiple Reactome and GO terms associated with viral infection, supporting the relevance of the targets within the infection context.

Inclusion of many diverse omics datasets as well as multiple CRISPR KO studies generated a large pool of SL candidate genes for further prioritization and validation. Broadening our investigation, we identified SL targets that may be pan-viral. Even with their distinct life cycles and biology, both coronaviruses and influenza viruses shared common host SL targets. One example is 60S ribosomal protein L28 (RPL28) that plays a role in negatively regulating an influenza A virus (IAV) encoded peptide for antigen presentation, thus potentially modulating immunosurveillance (90). Two other candidate SL targets for both SARS-CoV-2 and IAV that emerged from our screen, VPS25 and charged multivesicular body protein 2A (CHMP2A), are of particular interest. They are part of the ESCRT complex (91–93), which is an important component of a major pathway for the lysosomal degradation of monoubiquitinated membrane proteins (94) and is involved in the abscission step of cytokinesis and in the budding process of several enveloped as well as nonenveloped viruses. A recent report shows that depletion of VPS25 by RNA interference leads to decreased rotavirus replication in CaCo2 cells via reduced viral entry (95), confirming the possible efficacy of targeting VPS25 as an antiviral strategy. Both VPS25 and CHMP2A represent candidate pan-viral SL targets involved in multivesicular body assembly, which is directed by the ESCRT complex and has been shown to be involved in the early viral life cycle of multiple viruses (96–98). Furthermore, CHMP2A interacts with SARS-CoV-2 Orf9b and has been shown to contribute to the budding of a variety of viruses including HIV (87), equine infectious anemia virus (99), and murine leukemia virus (88), suggesting a critical role for virus release and thus a potential role as a pan-viral SL target. While these three candidate SL targets could be promising, none are listed as drug targets in DrugBank. Additionally, CHMP2A and VPS25 are essential in 10/10 cell lines tested in a recent study (72).

When considering genes in our list of 61 potentially pan-viral SL targets with more ideal characteristics, we found four genes that are known drug targets, are neither multifunctional nor widely essential, and have zero or very few paralogs. These genes include ligase 3 (LIG3), low-density lipoprotein receptor-related protein 1 (LRP1), monoamine oxidase A (MAOA) and T-cell immune regulator 1, ATPase H^+ transporting V0 subunit A3 (TCIRG1). Of these, LIG3 fulfills several other criteria that make this gene potentially a strong SL target. For example, the cellular protein PARP1 interacts directly with the ORF14 encoded protein of SARS-CoV-2 as well as with the polymerase complex of IAV (42,44,100) rendering it a VIH. LIG3 is a predicted SL partner of PARP1 and the two genes are paralogs, a strong predictor for SL interactions. Additionally, the LIG3 and PARP1 proteins physically interact and participate in DNA repair, mitochondrial organization, and regulation of DNA metabolism, among other processes (101,102). Previous research has shown that PARP1 is involved in the replication process of SARS-CoV-2 and other viruses (103); therefore, it is possible that depletion of its SL partner LIG3 could critically impair viral replication and/or lead to the killing of infected cells. Thus, our results widen the future exploration scope for therapeutic strategies against multiple viruses, for example by using combined drug regimens that could target two or more host proteins at otherwise subtherapeutic levels, potentially minimizing off-target effects. However, extensive and systematic investigation is required to determine the suitability of these and other candidate SL genes as drug targets for antiviral therapy.

It is important to note that our methodology, while promising, is not without limitations. We recognize that no single dataset or database can predict or validate virus-specific SL targets with absolute accuracy. For example, not all gene products with altered expression or changed modifications are necessarily true VIHs with reduced functionality. Furthermore, CRISPR KO-depleted genes are not a true 'gold standard' for SL prediction, but they currently provide the best available type of data suitable for assessing performance in support of a gene's SL candidacy. Additionally, the SynLethDB database used here contains SL gene pairs derived primarily from cancer cell lines and tumors (54), and additional SL pairs from SynLEGG were evaluated in cancer cell lines as well (55). Thus, while these databases include many SL interactions that are relevant in both cancer and viral infection states, it will contain some SL targets that are irrelevant in the virus-infected cellular state. It may also lack some targets that are important in viral infections. Yet, when available data are combined in the manner demonstrated in our study, they can predict a significant number of candidate SL genes that can be further constrained and validated in laboratory settings using virus-infected cell lines and animal models.

To prioritize SL targets for experimental validation, fold-change values for VIHs in omics (27) and CRISPR datasets could be used to rank their associated SL partners as potential targets. The 'SL statistics' score provided by SynLethDB (a weighted confidence score based on the source of the SL interaction) and the number of VIH partners paired with a predicted SL target (27) could also provide additional metrics for ranking. Approaches that have been shown to enrich for SL interactions in a more generic and thus less context-specific manner include pairing paralogs (87) as well as coregulated or mutually exclusive genes (104,105), in addition to pairing genes that exhibit certain network topology features (106). Some of these data-driven algorithms are part of SL-Cloud (105), a platform that was developed for the cancer field but can be utilized for infectious disease research. Any of the above metrics alone or in combination could be integrated into our pipeline to potentially improve predictive performance and help prioritize SL targets for validation.

To summarize, our research has demonstrated that virus-specific SL targets can be predicted from various omics data classes with varying levels of predictive power. We found that SynLethDB is highly enriched for CRISPR-depleted genes and has a unifying effect on the predictive capacities of the omics data classes. We also found that interactomic data classes performed better in predicting valid SL targets than others. This may be due to the fact that direct interactions of host proteins with viral factors are among the earliest cellular changes following infection and all other alterations occur as a consequence of these interactions. Moreover, physical protein interactions may function more like binary switches controlling protein functions, while other alterations tend to occur more gradually, likely leading to less abrupt disturbances in protein functions. Regardless of data class, predicted virus-exposed SL genes are highly enriched in depleted CRISPR KO data from virus-infected cells. Thus, our pipeline allows for a new application of many already existing high-throughput infectious disease datasets that are often underutilized. In human cells, each VIH was seen to form SL interactions with an average of three genes; in yeast, every gene is in at least one SL relationship across different growth conditions (11,88). Hence, it was not surprising that our study found many candidate SL targets for the two viruses investigated, all having the potential to specifically disrupt virus-infected cells. These candidate SL targets add to the number of potential host-based antiviral targets, mostly host dependency factors, which have been previously described in the literature. With further validation, our presented strategy for identifying SL targets has the potential to identify additional SL targets for other viruses and cell-dependent pathogens, leading to new broadly targeting host-based therapies against infectious diseases.

Data availability

No new data were generated or analyzed in support of this research. All available data used in this article are referenced in the text or tables and have been made public by the respective authors.

Supplementary data

Supplementary Data are available at NARMME Online.

Acknowledgements

We thank Fergal J. Duffy for guidance on questions related to specific R packages. We thank Jason P. Wendler for guidance and helpful discussions related to the generation of the pipeline. We also thank the rest of the Aitchison lab for helpful discussions and development of the virus-induced SL concept. Our code base does not include any novel algorithms but relies on R base functions, widely used publicly available R packages and standard predictive power measurements such as receiver operating characteristic curve analysis. It is available upon request.

Author contributions: J.P.S.: conceptualization, formal analysis, methodology, visualization, writing—original draft. M.L.N.: methodology, formal analysis, visualization, writing—review & editing. A.N.: conceptualization, writing—review & editing. F.D.M.: conceptualization, writing—review & editing. J.D.A.: conceptualization, writing—review & editing.

Funding

National Institutes of Health [P41 GM109824 to J.D.A.]; W.M. Keck Foundation [12, 2022 to J.D.A.]; Seattle Children's Research Institute. Funding for open access charge: National Institutes of Health.

Conflict of interest statement

None declared.

References

1. Mast,F.D., Navare,A.T., van der Sloot,A.M., Coulombe-Huntington,J., Rout,M.P., Baliga,N.S., Kaushansky,A., Chait,B.T., Aderem,A., Rice,C.M., *et al.* (2020) Crippling life support for SARS-CoV-2 and other viruses through synthetic lethality. *J. Cell Biol.*, **219**, e202006159.
2. Lee,J.S., Nair,N.U., Dinstag,G., Chapman,L., Chung,Y., Wang,K., Sinha,S., Cha,H., Kim,D., Schperberg,A.V., *et al.* (2021) Synthetic lethality-mediated precision oncology via the tumor transcriptome. *Cell*, **184**, 2487–2502.

3. Mengwasser,K.E., Adeyemi,R.O., Leng,Y., Choi,M.Y., Clairmont,C., D'Andrea,A.D. and Elledge,S.J. (2019) Genetic screens reveal FEN1 and APEX2 as BRCA2 synthetic lethal targets. *Mol. Cell*, **73**, 885–899.
4. Lord,C.J. and Ashworth,A. (2017) PARP inhibitors: synthetic lethality in the clinic. *Science*, **355**, 1152–1158.
5. Castells-Roca,L., Tejero,E., Rodríguez-Santiago,B. and Surrallés,J. (2021) CRISPR screens in synthetic lethality and combinatorial therapies for cancer. *Cancers (Basel)*, **13**, 1–22.
6. Guo,L., Li,S., Qian,B., Wang,Y., Duan,R., Jiang,W., Kang,Y., Dou,Y., Yang,G., Shen,L., *et al.* (2020) Integrative omics analysis reveals relationships of genes with synthetic lethal interactions through a pan-cancer analysis. *Comput. Struct. Biotechnol. J.*, **18**, 3243–3254.
7. Shieh,G.S. (2022) Harnessing synthetic lethal interactions for personalized medicine. *J. Pers. Med.*, **12**, 1–10.
8. Akimov,Y. and Aittokallio,T. (2021) Re-defining synthetic lethality by phenotypic profiling for precision oncology. *Cell Chem. Biol.*, **28**, 246–256.
9. Tang,L., Chen,R. and Xu,X. (2020) Synthetic lethality: a promising therapeutic strategy for hepatocellular carcinoma. *Cancer Lett.*, **476**, 120–128.
10. Hartwell,L.H., Szankasi,P., Roberts,C.J., Murray,A.W. and Friend,S.H. (1997) Integrating genetic approaches into the discovery of anticancer drugs. *Science*, **278**, 1064–1068.
11. Srivas,R., Shen,J.P., Yang,C.C., Sun,S.M., Li,J., Gross,A.M., Jensen,J., Licon,K., Bojorquez-Gomez,A., Klepper,K., *et al.* (2016) A network of conserved synthetic lethal interactions for exploration of precision cancer therapy. *Mol. Cell*, **63**, 514–525.
12. Huang,A., Garraway,L.A., Ashworth,A. and Weber,B. (2020) Synthetic lethality as an engine for cancer drug target discovery. *Nat. Rev. Drug Discov.*, **19**, 23–38.
13. Van Poelvoorde,L.A.E., Saelens,X., Thomas,I. and Roosens,N.H. (2020) Next-generation sequencing: an eye-opener for the surveillance of antiviral resistance in influenza. *Trends Biotechnol.*, **38**, 360–367.
14. Jefferson,T. (2014) Oseltamivir for influenza in adults and children: systematic review of clinical study reports and summary of regulatory comments. *BMJ*, **348**, 1–18.
15. Chitalia,V.C. and Munawar,A.H. (2020) A painful lesson from the COVID-19 pandemic: the need for broad-spectrum, host-directed antivirals. *J. Transl. Med.*, **18**, 390–395.
16. Castro,L.A., Bedford,T. and Meyers,L.A. (2020) Early prediction of antigenic transitions for influenza A/H3N2. *PLoS Comput. Biol.*, **16**, e1007683.
17. Chaudhuri,S., Symons,J.A. and Deval,J. (2018) Innovation and trends in the development and approval of antiviral medicines: 1987–2017 and beyond. *Antiviral Res.*, **155**, 76–88.
18. De Clercq,E. and Li,G. (2016) Approved antiviral drugs over the past 50 years. *Clin. Microbiol. Rev.*, **29**, 695–747.
19. Adamson,C.S., Chibale,K., Goss,R.J.M., Jaspars,M., Newman,D.J. and Dorrington,R.A. (2021) Antiviral drug discovery: preparing for the next pandemic. *Chem. Soc. Rev.*, **50**, 3647–3655.
20. Newman,D.J. and Cragg,G.M. (2020) Natural products as sources of new drugs over the nearly four decades from 01/1981 to 09/2019. *J. Nat. Prod.*, **83**, 770–803.
21. Ravindran,V., Nacher,J.C., Akutsu,T., Ishitsuka,M., Osadcenco,A., Sunitha,V., Bagler,G., Schwartz,J.M. and Robertson,D.L. (2019) Network controllability analysis of intracellular signalling reveals viruses are actively controlling molecular systems. *Sci. Rep.*, **9**, 1–11.
22. 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.
23. Ashworth,A., Lord,C.J. and Reis-Filho,J.S. (2011) Genetic interactions in cancer progression and treatment. *Cell*, **145**, 30–38.
24. Kaelin,W.G. (2005) The concept of synthetic lethality in the context of anticancer therapy. *Nat. Rev. Cancer*, **5**, 689–698.
25. Turner,N.C., Lord,C.J., Iorns,E., Brough,R., Swift,S., Elliott,R., Rayter,S., Tutt,A.N. and Ashworth,A. (2008) A synthetic lethal siRNA screen identifying genes mediating sensitivity to a PARP inhibitor. *EMBO J.*, **27**, 1368–1377.
26. Wiltshire,T.D., Lovejoy,C.A., Wang,T., Xia,F., O'Connor,M.J. and Cortez,D. (2010) Sensitivity to poly(ADP-ribose) polymerase (PARP) inhibition identifies ubiquitin-specific peptidase 11 (USP11) as a regulator of DNA double-strand break repair. *J. Biol. Chem.*, **285**, 14565–14571.
27. Pal,L.R., Cheng,K., Nair,N.U., Martin-Sancho,L., Sinha,S., Pu,Y., Riva,L., Yin,X., Schischlik,F., Lee,J.S., *et al.* (2022) Synthetic lethality-based prediction of anti-SARS-CoV-2 targets. *iScience*, **25**, 104311–104328.
28. Lee,J.S., Das,A., Jerby-Arnon,L., Arafeh,R., Auslander,N., Davidson,M., McGarry,L., James,D., Amzallag,A., Park,S.G., *et al.* (2018) Harnessing synthetic lethality to predict the response to cancer treatment. *Nat. Commun.*, **9**, 2546–2557.
29. Navare,A.T., Mast,F.D., Olivier,J.P., Bertomeu,T., Neal,M.L., Carpp,L.N., Kaushansky,A., Coulombe-Huntington,J., Tyers,M. and Aitchison,J.D. (2022) Viral protein engagement of GBF1 induces host cell vulnerability through synthetic lethality. *J. Cell Biol.*, **221**, e202011050.
30. Claude,A., Zhao,B.P., Kuziemsky,C.E., Dahan,S., Berger,S.J., Yan,J.P., Armold,A.D., Sullivan,E.M. and Melançon,P. (1999) GBF1: a novel Golgi-associated BFA-resistant guanine nucleotide exchange factor that displays specificity for ADP-ribosylation factor 5. *J. Cell Biol.*, **146**, 71–84.
31. 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.
32. Stukalov,A., Girault,V., Grass,V., Karayel,O., Bergant,V., Urban,C., Haas,D.A., Huang,Y., Oubraham,L., Wang,A., *et al.* (2021) Multilevel proteomics reveals host perturbations by SARS-CoV-2 and SARS-CoV. *Nature*, **594**, 246–252.
33. Xiong,Y., Liu,Y., Cao,L., Wang,D., Guo,M., Jiang,A., Guo,D., Hu,W., Yang,J., Tang,Z., *et al.* (2020) Transcriptomic characteristics of bronchoalveolar lavage fluid and peripheral blood mononuclear cells in COVID-19 patients. *Emerg. Microbes Infect.*, **9**, 761–770.
34. Lieberman,N.A.P., Peddu,V., Xie,H., Shrestha,L., Huang,M.L., Mears,M.C., Cajimat,M.N., Bente,D.A., Shi,P.Y., Bovier,F., *et al.* (2020) *In vivo* antiviral host transcriptional response to SARS-CoV-2 by viral load, sex, and age. *PLoS Biol.*, **18**, e3000849.
35. Huang,J., Hume,A.J., Abo,K.M., Werder,R.B., Villacorta-Martin,C., Alysandratos,K.D., Beermann,M.L., Simone-Roach,C., Lindstrom-Vautrin,J., Olejnik,J., *et al.* (2020) SARS-CoV-2 infection of pluripotent stem cell-derived human lung alveolar type 2 cells elicits a rapid epithelial-intrinsic inflammatory response. *Cell Stem Cell*, **27**, 962–973.
36. Appelberg,S., Gupta,S., Akusjärvi,S.S., Ambikan,A.T., Mikaeloff,F., Saccon,E., Végvári,Á., Benfeitas,R., Sperk,M., Ståhlberg,M., *et al.* (2020) Dysregulation in Akt/mTOR/HIF-1 signaling identified by proteo-transcriptomics of SARS-CoV-2 infected cells. *Emerg. Microbes Infect.*, **9**, 1748–1760.
37. Akgun,E., Tuzuner,M.B., Sahin,B., Kilercik,M., Kulah,C., Cakiroglu,H.N., Serteser,M., Unsal,I. and Baykal,A.T. (2020) Proteins associated with neutrophil degranulation are upregulated in nasopharyngeal swabs from SARS-CoV-2 patients. *PLoS One*, **15**, e240012.
38. Weingarten-Gabbay,S., Klaeger,S., Sarkizova,S., Pearlman,L.R., Chen,D.Y., Gallagher,K.M.E., Bauer,M.R., Taylor,H.B., Dunn,W.A., Tarr,C., *et al.* (2021) Profiling SARS-CoV-2 HLA-I

peptidome reveals T cell epitopes from out-of-frame ORFs. *Cell*, **184**, 3962–3980.
39. Klann,K., Bojkova,D., Tascher,G., Ciesek,S., Münch,C. and Cinatl,J. (2020) Growth factor receptor signaling inhibition prevents SARS-CoV-2 replication. *Mol. Cell*, **80**, 164–174.
40. Bojkova,D., Klann,K., Koch,B., Widera,M., Krause,D., Ciesek,S., Cinatl,J. and Münch,C. (2020) Proteomics of SARS-CoV-2-infected host cells reveals therapy targets. *Nature*, **583**, 469–472.
41. Meyer,B., Chiaravalli,J., Gellenoncourt,S., Brownridge,P., Bryne,D.P., Daly,L.A., Grauslys,A., Walter,M., Agou,F., Chakrabarti,L.A., *et al.* (2021) Characterising proteolysis during SARS-CoV-2 infection identifies viral cleavage sites and cellular targets with therapeutic potential. *Nat. Commun.*, **12**, 5553–5568.
42. 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.* (2020) A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature*, **583**, 459–468.
43. Li,G., Tang,Z., Fan,W., Wang,X., Huang,L., Jia,Y., Wang,M., Hu,Z. and Zhou,Y. (2023) Atlas of interactions between SARS-CoV-2 macromolecules and host proteins. *Cell Insight*, **2**, 100068.
44. Zhou,Y., Liu,Y., Gupta,S., Paramo,M.I., Hou,Y., Mao,C., Luo,Y., Judd,J., Wierbowski,S., Bertolotti,M., *et al.* (2022) A comprehensive SARS-CoV-2–human protein–protein interactome reveals COVID-19 pathobiology and potential host therapeutic targets. *Nat. Biotechnol.*, **41**, 128–139.
45. Liu,X., Huuskonen,S., Laitinen,T., Redchuk,T., Bogacheva,M., Salokas,K., Pöhner,I., Öhman,T., Tonduru,A.K., Hassinen,A., *et al.* (2021) SARS-CoV-2–host proteome interactions for antiviral drug discovery. *Mol. Syst. Biol.*, **17**, e10396.
46. 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 doi: https://doi.org/10.1101/2020.08.28.272955, 29 August 2020, preprint: not peer reviewed.
47. May,D.G., Martin-Sancho,L., Anschau,V., Liu,S., Chrisopulos,R.J., Scott,K.L., Halfmann,C.T., Peña,R.D., Pratt,D., Campos,A.R., *et al.* (2022) A BioID-derived proximity interactome for SARS-CoV-2 proteins. *Viruses*, **14**, 611.
48. Meyers,J.M., Ramanathan,M., Shanderson,R.L., Beck,A., Donohue,L., Ferguson,I., Guo,M.G., Rao,D.S., Miao,W., Reynolds,D., *et al.* (2021) The proximal proteome of 17 SARS-CoV-2 proteins links to disrupted antiviral signaling and host translation. *PLoS Pathog.*, **17**, e1009800.
49. 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 doi: https://doi.org/10.1101/2020.09.03.282103, 4 September 2020,preprint: not peer reviewed.
50. 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 doi: https://doi.org/10.1101/2020.08.28.269175, 28 August 2020, preprint: not peer reviewed.
51. Kamel,W., Noerenberg,M., Cerikan,B., Chen,H., Järvelin,A.I., Kammoun,M., Lee,J.Y., Shuai,N., Garcia-Moreno,M., Andrejeva,A., *et al.* (2021) Global analysis of protein–RNA interactions in SARS-CoV-2-infected cells reveals key regulators of infection. *Mol. Cell*, **81**, 2851–2867.
52. Flynn,R.A., Belk,J.A., Qi,Y., Yasumoto,Y., Schmitz,C.O., Mumbach,M.R., Limaye,A., Wei,J., Alfajaro,M.M., Parker,K.R., *et al.* (2020) Discovery and functional interrogation of SARS-CoV-2 RNA–host protein interactions. *Cell*, **184**, 2394–2411.
53. Schmidt,N., Lareau,C.A., Keshishian,H., Ganskih,S., Schneider,C., Hennig,T., Melanson,R., Werner,S., Wei,Y., Zimmer,M., *et al.* (2021) The SARS-CoV-2 RNA–protein interactome in infected human cells. *Nat. Microbiol.*, **6**, 339–353.
54. Wang,J., Wu,M., Huang,X., Wang,L., Zhang,S., Liu,H. and Zheng,J. (2022) SynLethDB 2.0: a web-based knowledge graph database on synthetic lethality for novel anticancer drug discovery. *Database*, **2022**, 1–10.
55. Wappett,M., Harris,A., Lubbock,A.L.R., Lobb,I., McDade,S. and Overton,I.M. (2021) SynLeGG: analysis and visualization of multiomics data for discovery of cancer 'Achilles heels' and gene function relationships. *Nucleic Acids Res.*, **49**, W613–W618.
56. Baggen,J., Persoons,L., Vanstreels,E., Jansen,S., Van Looveren,D., Boeckx,B., Geudens,V., De Man,J., Jochmans,D., Wauters,J., *et al.* (2021) Genome-wide CRISPR screening identifies TMEM106B as a proviral host factor for SARS-CoV-2. *Nat. Genet.*, **53**, 435–444.
57. Daniloski,Z., Jordan,T.X., Wessels,H.H., Hoagland,D.A., Kasela,S., Legut,M., Maniatis,S., Mimitou,E.P., Lu,L., Geller,E., *et al.* (2021) Identification of required host factors for SARS-CoV-2 infection in human cells. *Cell*, **184**, 92–105.
58. Rebendenne,A., Roy,P., Bonaventure,B., Chaves Valadão,A.L., Desmarets,L., Arnaud-Arnould,M., Rouillé,Y., Tauziet,M., Giovannini,D., Touhami,J., *et al.* (2022) Bidirectional genome-wide CRISPR screens reveal host factors regulating SARS-CoV-2, MERS-CoV and seasonal HCoVs. *Nat. Genet.*, **54**, 1090–1102.
59. Hoffmann,H.H., Schneider,W.M., Rozen-Gagnon,K., Miles,L.A., Schuster,F., Razooky,B., Jacobson,E., Wu,X., Yi,S., Rudin,C.M., *et al.* (2021) TMEM41B is a pan-flavivirus host factor. *Cell*, **184**, 133–148.
60. Schneider,W.M., Luna,J.M., Hoffmann,H.H., Sánchez-Rivera,F.J., Leal,A.A., Ashbrook,A.W., Le Pen,J., Ricardo-Lax,I., Michailidis,E., Peace,A., *et al.* (2021) Genome-scale identification of SARS-CoV-2 and pan-coronavirus host factor networks. *Cell*, **184**, 120–132.
61. Zhu,Y., Feng,F., Hu,G., Wang,Y., Yu,Y., Zhu,Y., Xu,W., Cai,X., Sun,Z., Han,W., *et al.* (2021) A genome-wide CRISPR screen identifies host factors that regulate SARS-CoV-2 entry. *Nat. Commun.*, **12**, 1–11.
62. Wang,R., Simoneau,C.R., Kulsuptrakul,J., Bouhaddou,M., Travisano,K.A., Hayashi,J.M., Carlson-Stevermer,J., Zengel,J.R., Richards,C.M., Fozouni,P., *et al.* (2021) Genetic screens identify host factors for SARS-CoV-2 and common cold coronaviruses. *Cell*, **184**, 106–119.
63. Grodzki,M., Bluhm,A.P., Schaefer,M., Tagmount,A., Russo,M., Sobh,A., Rafiee,R., Vulpe,C.D., Karst,S.M. and Norris,M.H. (2022) Genome-scale CRISPR screens identify host factors that promote human coronavirus infection. *Genome Med.*, **14**, 1–18.
64. Israeli,M., Finkel,Y., Yahalom-Ronen,Y., Paran,N., Chitlaru,T., Israeli,O., Cohen-Gihon,I., Aftalion,M., Falach,R., Rotem,S., *et al.* (2022) Genome-wide CRISPR screens identify GATA6 as a proviral host factor for SARS-CoV-2 via modulation of ACE2. *Nat. Commun.*, **13**, 1–16.
65. Biering,S.B., Sarnik,S.A., Wang,E., Zengel,J.R., Leist,S.R., Schäfer,A., Sathyan,V., Hawkins,P., Okuda,K., Tau,C., *et al.* (2022) Genome-wide bidirectional CRISPR screens identify mucins as host factors modulating SARS-CoV-2 infection. *Nat. Genet.*, **54**, 1078–1089.
66. Hou,J., Wei,Y., Zou,J., Jaffery,R., Sun,L., Liang,S., Zheng,N., Guerrero,A.M., Egan,N.A., Bohat,R., *et al.* (2024) Integrated multi-omics analyses identify anti-viral host factors and pathways controlling SARS-CoV-2 infection. *Nat. Commun.*, **15**, 109–122.

67. Yi,C., Cai,C., Cheng,Z., Zhao,Y., Yang,X., Wu,Y., Yin,Z., Xiang,Y., Yin,M., Han,L., *et al.* (2022) Genome-wide CRISPR-Cas9 screening identifies the CYTH2 host gene as a potential therapeutic target of influenza viral infection. *Cell Rep.*, **38**, 110559.
68. Tran,V., Ledwith,M.P., Thamamongood,T., Higgins,C.A., Tripathi,S., Chang,M.W., Benner,C., García-Sastre,A., Schwemmle,M., Boon,A.C.M., *et al.* (2020) Influenza virus repurposes the antiviral protein IFIT2 to promote translation of viral mRNAs. *Nat. Microbiol.*, **5**, 1490–1503.
69. Han,J., Perez,J.T., Chen,C., Li,Y., Benitez,A., Kandasamy,M., Lee,Y., Andrade,J., tenOever,B. and Manicassamy,B. (2018) Genome-wide CRISPR/Cas9 screen identifies host factors essential for influenza virus replication. *Cell Rep.*, **23**, 596–607.
70. Durinck,S., Spellman,P.T., Birney,E. and Huber,W. (2009) Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat. Protoc.*, **4**, 1184–1191.
71. Ribeiro,D.M., Briere,G., Bely,B., Spinelli,L. and Brun,C. (2019) MoonDB 2.0: an updated database of extreme multifunctional and moonlighting proteins. *Nucleic Acids Res.*, **47**, D398–D402.
72. Bertomeu,T., Coulombe-Huntington,J., Chatr-Aryamontri,A., Bourdages,K.G., Coyaud,E., Raught,B., Xia,Y. and Tyers,M. (2018) A high-resolution genome-wide CRISPR/Cas9 viability screen reveals structural features and contextual diversity of the human cell-essential proteome. *Mol. Cell. Biol.*, **38**, e00302-17.
73. Wishart,D.S., Feunang,Y.D., Guo,A.C., Lo,E.J., Marcu,A., Grant,J.R., Sajed,T., Johnson,D., Li,C., Sayeeda,Z., *et al.* (2018) DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res.*, **46**, D1074–D1082.
74. Wu,T., Hu,E., Xu,S., Chen,M., Guo,P., Dai,Z., Feng,T., Zhou,L., Tang,W., Zhan,L., *et al.* (2021) clusterProfiler 4.0: a universal enrichment tool for interpreting omics data. *Innovation (Camb.)*, **2**, 100141.
75. Yu,G. and He,Q.-Y. (2016) ReactomePA: an R/Bioconductor package for Reactome pathway analysis and visualization. *Mol. Biosyst.*, **12**, 477–479.
76. Benjamini,Y. and Yekutieli,D. (2001) The control of the false discovery rate in multiple testing under dependency. *Ann. Stat.*, **29**, 1165–1188.
77. Carlson, M. (2019) org.Hs.eg.db: Genome wide annotation for Human. R package version 3.8.2.. https://bioconductor.org/packages/org.Hs.eg.db/, (1 November 2023, date last accessed).
78. Li,J., Guo,M., Tian,X., Wang,X., Yang,X., Wu,P., Liu,C., Xiao,Z., Qu,Y., Yin,Y., *et al.* (2021) Virus–host interactome and proteomic survey reveal potential virulence factors influencing SARS-CoV-2 pathogenesis. *Med*, **2**, 99–112.
79. Lieberman,N.A.P., Peddu,V., Xie,H., Shrestha,L., Huang,M.-L., Mears,M.C., Cajimat,M.N., Bente,D.A., Shi,P.-Y., Bovier,F., *et al.* (2020) *In vivo* antiviral host response to SARS-CoV-2 by viral load, sex, and age. *PLoS Biol.*, **18**, e3000849.
80. Varjosalo,M., Keskitalo,S., VanDrogen,A., Nurkkala,H., Vichalkovski,A., Aebersold,R. and Gstaiger,M. (2013) The protein interaction landscape of the human CMGC kinase group. *Cell Rep.*, **3**, 1306–1320.
81. Roux,K.J., Kim,D.I., Raida,M. and Burke,B. (2012) A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *J. Cell Biol.*, **196**, 801–810.
82. Gillen,J. and Nita-Lazar,A. (2019) Experimental analysis of viral–host interactions. *Front. Physiol.*, **10**, 1–12.
83. Flynn,R.A., Belk,J.A., Qi,Y., Yasumoto,Y., Wei,J., Alfajaro,M.M., Shi,Q., Mumbach,M.R., Limaye,A., DeWeirdt,P.C., *et al.* (2021) Discovery and functional interrogation of SARS-CoV-2 RNA–host protein interactions. *Cell*, **184**, 2394–2411.
84. Liu,Y., Zhu,Y., Sun,X., Ma,T., Lao,X. and Zheng,H. (2023) DRAVP: a comprehensive database of antiviral peptides and proteins. *Viruses*, **15**, 820–832.
85. Dou,D., Revol,R., Östbye,H., Wang,H. and Daniels,R. (2018) Influenza A virus cell entry, replication, virion assembly and movement. *Front. Immunol.*, **9**, 1581.
86. Makeyev,A.V. and Liebhaber,S.A. (2002) The poly(C)-binding proteins: a multiplicity of functions and a search for mechanisms. *RNA*, **8**, 265–278.
87. Xin,Y. and Zhang,Y. (2023) Paralog-based synthetic lethality: rationales and applications. *Front. Oncol.*, **13**, 1168143.
88. Costanzo,M., Kuzmin,E., van Leeuwen,J., Mair,B., Moffat,J., Boone,C. and Andrews,B. (2019) Global genetic networks and the genotype-to-phenotype relationship. *Cell*, **177**, 85–100.
89. Honarmand Ebrahimi,K., Ciofi-Baffoni,S., Hagedoorn,P.-L., Nicolet,Y., Le Brun,N.E., Hagen,W.R. and Armstrong,F.A. (2022) Iron–sulfur clusters as inhibitors and catalysts of viral replication. *Nat. Chem.*, **14**, 253–266.
90. Wei,J., Kishton,R.J., Angel,M., Conn,C.S., Dalla-Venezia,N., Marcel,V., Vincent,A., Catez,F., Ferré,S., Ayadi,L., *et al.* (2019) Ribosomal proteins regulate MHC class I peptide generation for immunosurveillance. *Mol. Cell*, **73**, 1162–1173.
91. Vietri,M., Schink,K.O., Campsteijn,C., Sem Wegner,C., Schultz,S.W., Christ,L., Thoresen,S.B., Brech,A., Raiborg,C. and Stenmark,H. (2015) Spastin and ESCRT-III coordinate mitotic spindle disassembly and nuclear envelope sealing. *Nature*, **522**, 231–235.
92. Morita,E., Sandrin,V., Mccullough,J., Katsuyama,A., Hamilton,I.B. and Sundquist,W.I. (2011) ESCRT-III protein requirements for HIV-1 budding. *Cell Host Microbe*, **9**, 235–242.
93. Bartusch,C. and Prange,R. (2016) ESCRT requirements for murine leukemia virus release. *Viruses*, **8**, 1–18.
94. Wollert,T. and Hurley,J.H. (2010) Molecular mechanism of multivesicular body biogenesis by ESCRT complexes. *Nature*, **464**, 864–869.
95. Silva-Ayala,D., López,T., Gutiérrez,M., Perrimon,N., López,S. and Arias,C.F. (2013) Genome-wide RNAi screen reveals a role for the ESCRT complex in rotavirus cell entry. *Proc. Natl Acad. Sci. U.S.A.*, **110**, 10270–10275.
96. Gruenberg,J. (2020) Life in the lumen: the multivesicular endosome. *Traffic*, **21**, 76–93.
97. Votteler,J. and Sundquist,W.I. (2013) Virus budding and the ESCRT pathway. *Cell Host Microbe*, **14**, 232–241.
98. Meng,B. and Lever,A.M.L. (2021) The interplay between ESCRT and viral factors in the enveloped virus life cycle. *Viruses*, **13**, 324.
99. Sandrin,V. and Sundquist,W.I. (2013) ESCRT requirements for EIAV budding. *Retrovirology*, **10**, 104.
100. Mayer,D., Molawi,K., Martínez-Sobrido,L., Ghanem,A., Thomas,S., Baginsky,S., Grossmann,J., García-Sastre,A. and Schwemmle,M. (2007) Identification of cellular interaction partners of the influenza virus ribonucleoprotein complex and polymerase complex using proteomic-based approaches. *J. Proteome Res.*, **6**, 672–682.
101. Leppard,J.B., Dong,Z., Mackey,Z.B. and Tomkinson,A.E. (2003) Physical and functional interaction between DNA ligase IIIα and poly(ADP-ribose) polymerase 1 in DNA single-strand break repair. *Mol. Cell. Biol.*, **23**, 5919–5927.
102. Hu,Y., Lin,J., Fang,H., Fang,J., Li,C., Chen,W., Liu,S., Ondrejka,S., Gong,Z., Reu,F., *et al.* (2018) Targeting the MALAT1/PARP1/LIG3 complex induces DNA damage and apoptosis in multiple myeloma. *Leukemia*, **32**, 2250–2262.
103. Fehr,A.R., Singh,S.A., Kerr,C.M., Mukai,S., Higashi,H. and Aikawa,M. (2020) The impact of PARPs and ADP-ribosylation on inflammation and host–pathogen interactions. *Genes Dev.*, **34**, 341–359.
104. Jerby-Arnon,L., Pfetzer,N., Waldman,Y.Y., McGarry,L., James,D., Shanks,E., Seashore-Ludlow,B., Weinstock,A., Geiger,T., Clemons,P.A., *et al.* (2014) Predicting cancer-specific vulnerability via data-driven detection of synthetic lethality. *Cell*, **158**, 1199–1209.

105. Tercan,B., Qin,G., Kim,T.-K., Aguilar,B., Phan,J., Longabaugh,W., Pot,D., Kemp,C.J., Chambwe,N. and Shmulevich,I. (2022) SL-Cloud: a Cloud-based resource to support synthetic lethal interaction discovery. *F1000Research*, **11**, 493.

106. Benstead-Hume,G., Chen,X., Hopkins,S.R., Lane,K.A., Downsid,J.A. and Pearlid,F.M.G. (2019) Predicting synthetic lethal interactions using conserved patterns in protein interaction networks. *PLoS Comput. Biol.*, **15**, e1006888.

Received: August 17, 2023. **Revised:** December 8, 2023. **Editorial Decision:** December 14, 2023. **Accepted:** January 3, 2024