Machine learning on large scale perturbation screens for SARS-CoV-2 host factors identifies β-catenin/CBP inhibitor PRI-724 as a potent antiviral

Maximilian A. Kelch[1], Antonella Vera-Guapi[2], Thomas Beder[3], Marcus Oswald[4], Alicia Hiemisch[4], Nina Beil[5], Piotr Wajda[5], Sandra Ciesek[1,6,7], Holger Erfle[5], Tuna Toptan[1*†] and Rainer Koenig[4*†]

[1]Institute for Medical Virology, University Hospital Frankfurt, Goethe University Frankfurt, Frankfurt, Germany, [2]Institute of Biochemistry II, University Hospital, Frankfurt, Germany, [3]Medical Department II, Hematology and Oncology, University Hospital Schleswig-Holstein, Kiel, Germany, [4]Institute for Infectious Diseases and Infection Control, Jena University Hospital, Jena, Germany, [5]Advanced Biological Screening Facility (ABSF), High-Content Analysis of the Cell (HiCell), BioQuant, Heidelberg University, Heidelberg, Germany, [6]German Centre for Infection Research (DZIF), External Partner Site Frankfurt, Frankfurt, Germany, [7]Fraunhofer Institute for Translational Medicine and Pharmacology ITMP, Frankfurt, Germany

Expanding antiviral treatment options against SARS-CoV-2 remains crucial as the virus evolves under selection pressure which already led to the emergence of several drug resistant strains. Broad spectrum host-directed antivirals (HDA) are promising therapeutic options, however the robust identification of relevant host factors by CRISPR/Cas9 or RNA interference screens remains challenging due to low consistency in the resulting hits. To address this issue, we employed machine learning, based on experimental data from several knockout screens and a drug screen. We trained classifiers using genes essential for virus life cycle obtained from the knockout screens. The machines based their predictions on features describing cellular localization, protein domains, annotated gene sets from Gene Ontology, gene and protein sequences, and experimental data from proteomics, phospho-proteomics, protein interaction and transcriptomic profiles of SARS-CoV-2 infected cells. The models reached a remarkable performance suggesting patterns of intrinsic data consistency. The predicted HDF were enriched in sets of genes particularly encoding development, morphogenesis, and neural processes. Focusing on development and morphogenesis-associated gene sets, we found β-catenin to be central and selected PRI-724, a canonical β-catenin/CBP disruptor, as a potential HDA. PRI-724 limited infection with SARS-CoV-2 variants, SARS-CoV-1, MERS-CoV and IAV in different cell line models. We detected a concentration-dependent reduction in cytopathic effects, viral RNA replication, and infectious virus production in SARS-CoV-2 and SARS-CoV-1-infected cells. Independent of virus infection, PRI-724 treatment caused cell cycle deregulation which substantiates its potential as a broad spectrum antiviral. Our proposed machine learning concept supports focusing and accelerating the discovery of host dependency factors and identification of potential host-directed antivirals.

KEYWORDS

coronavirus, SARS-CoV-2, influenza A virus, host dependency factors, CRISPR/Cas9 knockout screen, beta-catenin, machine learning, PRI-724

1. Introduction

Genetic plasticity of SARS-CoV-2 and herd immunity-derived evolutionary pressure have led to the emergence of variants of concern (VOC; Van Egeren et al., 2021; Flores-Vega et al., 2022). Some of the variants and recombinant lineages have evaded vaccine-elicited antibodies and monoclonal antibody-based treatments (Dejnirattisai et al., 2021; Planas et al., 2021; Wang P. et al., 2021; Iketani et al., 2022). Although antiviral drugs such as Remdesivir, Molnupiravir, or Nirmatrelvir/Ritonavir (Paxlovid) have significantly improved disease outcome in patients (Beigel et al., 2020; Hammond et al., 2022; Jayk Bernal et al., 2022; Wong et al., 2022), these treatments might select for more resistant variants over time (Gandhi et al., 2022; Hogan et al., 2023). Consequently, our current virus-directed countermeasures have limitations and there is an urgent need for developing host-directed antivirals (HDA) that can target host-dependency factors (HDF) required for the life cycle of the virus. HDA might provide variant-independent treatment opportunities to reduce disease severity and complement virus-directed antiviral therapies (Edinger TO et al., 2014; Mahajan et al., 2021; Wagoner et al., 2022). In addition, drug repurposing can offer an expedited timeline to bring HDA therapies into clinical settings in a cost-effective and timely manner. Still, a common challenge is the robust identification of relevant host factors among varying *in vitro* conditions, although recent advances in functional genomics have simplified the use of high-throughput technology platforms based on CRISPR/Cas9 and siRNAs (Daniloski et al., 2021; Wang R. et al., 2021; Wei et al., 2021; Zhu et al., 2021). However, when comparing the results of these experimental screens for SARS-CoV-2 host dependency factors, there is only a marginal overlap of hits. This may be due to different viral strains, host cells and/or different multiplicity of infection (MOI) or further different experimental settings such as the selection of knockout library constructs used. Interestingly, when grouping the identified HDF into gene sets of common function or cellular processes, the consistency increases suggesting that these lists of HDF in their specific contexts contain consistent patterns on a more complex level.

Machine learning methods can identify such common patterns in experimental data, as, e.g., the identification of common regulators explaining large scale transcription profiles (Alipanahi et al., 2015; Hörhold et al., 2020) and allow for integration of a broad variety of omics and genomics data. In this study, we followed a machine learning approach to identify HDF of SARS-CoV-2. From the knockout screens (Daniloski et al., 2021; Wang R. et al., 2021; Wei et al., 2021; Zhu et al., 2021), we assembled a gold standard of genes required for virus replication. With this gold standard, we trained and evaluated the classifier. The classifier based its predictions on features of the genes describing their encoded proteins´ cellular localization, protein domains, annotated cellular processes, and gene/protein sequences. Furthermore, we embedded experimental data comprising proteomic, PPI, and transcriptomic profiles of SARS-CoV-2 infected cells. In addition, another classifier was trained based on a gold standard derived from a drug screen (Ellinger et al., 2021), in which more than 6,000 drugs were screened to determine whether they had a protective effect on cells infected with SARS-CoV-2. Both classifiers delivered a list of predicted HDF. For a few, well selected HDF, we interrogated drug databases, and identified β-catenin and its small molecule inhibitor PRI-724, particularly antagonizing the interaction between β-catenin and CBP.

The β-catenin/CBP interaction displays a pivotal part of a switch-like signaling network within the Wnt/β-catenin pathway that controls the well-balanced interplay between cell proliferation and differentiation (Akiyama, 2000; Liu et al., 2022). Activation of Wnt signaling leads to a sequester of the β-catenin destruction complex enabling cytosolic accumulation of β-catenin and, in turn, its nuclear translocation where it interacts with CBP or p300 to drive either proliferation or differentiation, respectively (Bordonaro and Lazarova, 2016). In this study, we identified β-catenin as a SARS-CoV-2 host factor using machine-learning and sought to analyze the efficacy of one of its inhibitors, PRI-724, against different pandemic-related RNA viruses *in vitro*.

2. Materials and methods

2.1. Gene prioritizations

Two sets of machines were set up to predict HDF, one based on data from CRISPR/Cas9 based knockout screens, and one on data from a drug screen. For defining the gold standard for the first set of machines, the experimental results from four genome-wide CRISPR/Cas9 knockout screens were considered (Daniloski et al., 2021; Wang R. et al., 2021; Wei et al., 2021; Zhu et al., 2021; Table 1). The ranking of the screened genes was taken from the Supplementary material of the respective studies. The rank product was calculated for each gene for which information was available. The n = 500 top ranking genes of this new list were selected and used as the positive class for the

TABLE 1 HDF screens serving for a gold standard to train and validate the machines.

	Daniloski et al. (2021)	Wei et al. (2021)	Wang R. et al. (2021)	Zhu et al. (2021)
Number of screened genes	19,050	21,673	20,915	19,114
Cell line	A549-ACE2	Vero-E6	Huh7.5.1 with ACE2 and TMPRSS2 overexpression	A549-ACE2
SARS-CoV-2 strain	USA-WA1/2020 (NR-52281)	US-WA1/2020	USA/WA-1/2020	nCoV-SH01 (patient isolate)
MOI	0.01 (low) 0.3 (high)	0.1	0.01	0.3
CRISPR libraries	GeCKOv2 with 122,411 sgRNAs 6 sgRNAs/gene	*C. sabaeus* with 83,963 sgRNAs 4/sgRNAs gene	GeCKOv2	lentiCRISPRv2 76,441 sgRNAs

classifier, labeled as HDF. The bottom 1,000 genes were used as the negative class ("Non-HDF"). The rest of the genes in the list were used for prediction, i.e., assessing an HDF or non-HDF label. Daniloski et al. provided data of two screens with different MOI (0.01 and 0.3). To get a combined list of genes, we calculated the rank product for each gene listed in their score tables and used this combined ranking for our study. In addition, a second gold standard was established based on the screening of 5,632 compounds and their capacity to inhibit the cytopathic effect (CPE) of SARS-CoV-2 in a Caco-2 cell line (Ellinger et al., 2021). In line with the authors of the original study, compounds that blocked viral CPE by 75% were considered as hits. This led to 273 compounds inhibiting SARS-CoV-2 infection without inducing toxicity. Then, their respective targets were identified using the information of different drug databases [Drugbank (Wishart et al., 2018), ChEMBL (Mendez et al., 2019), TTD (Wang et al., 2020), PharmGKB (Thorn et al., 2013), and BindingDB (Gilson et al., 2016)]. By this, 178 gene targets were found and labeled as "HDF." For the remaining drugs with no inhibitory effect, their 1,881 targets were labeled as "non-HDF." Effective drugs with more than 11 targets were considered to be "promiscuous," and their targets were not included in the gold standard.

2.2. Feature generation

The procedure for feature generation was similar as published earlier for essential gene prediction (Acencio and Lemke, 2009; Aromolaran et al., 2020; Beder et al., 2021). Each gene served as a sample in the machine learning procedure, either labeled as HDF or non-HDF, or was not used for training the classifiers. For each gene, we generated a comprehensive set of 60,593 features comprising the seven categories of (1) gene sequence, (2) protein sequence, (3) protein domains, (4) gene sets from Gene Ontology, (5) conservation, (6) topology in the protein interaction network, and (7) subcellular location of the protein. We distinguished between intrinsic and extrinsic features. Intrinsic features denote features, which can be directly derived from gene and protein sequences, whereas extrinsic features describe network topology, homology, cellular localization of the expressed protein and functional (interaction) domains. Employing BioMart (Smedley et al., 2009), we downloaded the gene and protein sequences from Ensembl (v102). Protein and gene sequence features (categories 1 and 2) were calculated with several software tools [seqinR (Charif et al., 2005), protr (Xiao et al., 2015)], CodonW,[1] and rDNAse[2] spanning a large range of descriptors as explained in the following. Using seqinR, we calculated the number and percentage of each of the 20 amino acids in a protein. In addition, it was used to calculate other amino acid sequence features including the number of residues, the percentage of each of the physico-chemical classes and the theoretical isoelectric point. Using protr, we calculated the autocorrelation, Conjoint Triad Descriptors (CTD), quasi-sequence order and pseudo amino acid composition. CodonW was used calculating general gene features

like gene length and GC content, and frequencies of optimal codons (Hershberg and Petrov, 2009) and the number of codons in the coding sequence. rDNAse calculated additional gene features, like auto covariance, pseudo nucleotide composition, and kmer frequencies ($n = 2$–7). It also calculated the sequence attribute distribution. Amino acids were grouped into three classes: (1) polar, (2) neutral, or (3) hydrophobic. The second digit in the feature name accounted for the class. More fine grained, seven attributes comprising (1) hydrophobicity, (2) normalized van-der-Waals volume, (3) polarity, (4) polarizability, (5) charge, (6) secondary structure, and (7) solvent accessibility were represented by the first digit in the feature name. The last three digits described the location of the attribute in the sequence, i.e., either at the beginning of the sequence (000), around the 25% quantile of residues (025), 50% (050), 75% (075), or at the end of the sequence (100). For example, seq.attribute.distribution 42100 is the sequence attribute of amino acids being polarizable (4), neutral (2) and are located at the end of the sequence (100). The domain features (category 3) were calculated using BioMart yielding protein family domains (pfam domains), the number of coiled coils structures, prediction of membrane helices, post-translational modifications, β-turns, cofactor binding, acetylation and glycosylation sites, signal peptides and transmembrane helices, and the number and lengths of UTRs. To calculate the gene set features (category 4), gene sets of all terms from Gene Ontology (Biological Process, Molecular Function, Cellular Localization) were obtained from the gene annotation of Ensembl (v102). First, to each gene, its direct neighbors were added using the STRING database (Szklarczyk et al., 2019), and with this list a gene set enrichment test performed employing Fisher's exact tests. The negative log10 value of the p-value P was used as a feature. Highly redundant (overlapping) gene sets were removed by the following method. Overlap of genes among each pair of gene sets V and W was quantified by Jaccard similarity coefficients,

$$J(V,W) = \frac{|V \cap W|}{|V \cup W|} \qquad (1)$$

Pairs with J(V,W) above the threshold $\alpha = 0.3$ were included into a model and represented as an undirected graph G(X,E), in which X were the n gene sets and pairs above α the edges E. A Mixed Integer Linear Programming problem was set up,

Maximize $\qquad \Sigma w_i X_i \qquad (2)$

Subject to

$$X_i + X_j \leq 1 \text{ for every } \{i,j\} \, \epsilon \, E \qquad (3)$$

$$X_i \, \epsilon \, \{0,1\} \text{ for } 1 \leq i \leq n, \qquad (4)$$

in which w_i was the weight of a gene set and calculated from the significance value (p-value P) of the according Fisher's Exact test,

---

$$w_i = \frac{1 - \log_{10}(P)}{100} \qquad (5)$$

The problem was solved using Gurobi™ (version 7.5.1[3]). By this, up to one gene set of each pair with high overlap was selected and the gene set with the higher enrichment was privileged. Gene conservation (category 5) was assessed by sequence alignment of the protein sequence of each gene with all proteins listed in RefSeq (Pruitt et al., 2005) using PSI-Blast (Altschul et al., 1997). The number of obtained proteins with E-value cutoffs from 1e − 5 to 1e − 100 served as feature values of the respective features. For category 5, typical topology features like degree, degree distribution, several centrality descriptors and Page rank was obtained using protein interaction data ($n = 697,802$ interactions) from the BIOGRID database (Oughtred et al., 2019). In addition, we assembled a second network considering also the interactions between SARS-CoV-2 proteins and proteins of the human host cell. For this, in total 22,586 interactions between SARS-CoV-2 and human proteins were obtained from the BioGRID COVID-19 Coronavirus Curation Project (Oughtred et al., 2019). Viral proteins were added to the generic network and an interaction was added to the network between a host protein and a viral protein if the link was listed in the protein interactions between SARS-CoV-2 and human proteins. Using both protein interaction networks, in summary, 24 topological features were generated using the igraph R package for network analysis.[4] Subcellular localization (category 7) was derived using the prediction software DeepLoc (Almagro Armenteros et al., 2017), which assigns probability scores to 11 eukaryotic cellular compartments (cytoplasm, nucleus, mitochondria, ER, Golgi apparatus, lysosome, vacuole and peroxisome, plasma membrane, extracellular, chloroplast).

Furthermore, 69 features from experimental data derived from gene expression, proteomics and phospho-proteomics profiles of Caco-2 cells post infection with SARS-CoV-2 at different time points were used. Proteomics data was taken from a previous publication (Bojkova et al., 2020) and z-score transformed peptide counts for each time point post infection (0, 2, 6, 10, 24 h, MOI = 0.01) vs. time point zero taken. In line, own gene expression data from Caco-2 cells infected with SARS-CoV-2 (FFM1, Wuhan wildtype; 0, 3, 6, 12 and 24 hpi, MOI = 0.01). For both datasets, as features, $\log_{10}$ differential expression values between each time point post infection and time point zero (no infection) was taken. The phosphoproteomics data was taken from a previous study (SARS-CoV-2 infected Caco-2 cells, 24 hpi, MOI = 1) and features calculated similarly as for the proteomics data (Klann et al., 2020).

2.3. Machine learning pipeline

Each feature was normalized using z-score transformation. We trained and validated two different sets of machines, one based on the gold standard of the knockout screens (knockout screens based classifier) and one based on the gold standard from the drug screen (drugs screen based classifier). The machine learning procedure was

the same for both. For each, a 5-fold cross-validation (CV) was performed in which 4/5 of the data was used to train the model, and the remaining 1/5 of the data (validation set) was used to assess the performance leaving the test set unseen. In addition, we repeated these cross-validations five times and averaged the results over these five independent runs. The first step in the training procedure was feature selection. For this, Least Absolute Shrinkage and Selection Operator (LASSO) regression was performed using the "glmnet" R package (cv.glmnet function with parameters alpha = 1, type.measure = "auc"). In a second feature selection step, highly correlating features with Pearson's correlation coefficients r > 0.70 were removed. On average, 155 features remained. These features were fed into a Random Forest (RF) classifier using the caret package in R (tuneLength = 10, metric = "Kappa," nthread = 20, ntree = 500). All trained models were used to predict the complete set of 20,007 genes to find novel HDF candidates, and to confirm or reject the HDF annotation of the gold standard. We defined the top 10% of the predictions from a voting scheme as predicted HDF (predHDF) of the classifiers, i.e., of the knockout based classifier and of the drug based classifier, leading to 2,182 and 1,989 predHDF, respectively. To evaluate the performance of the learned models, the area under the receiver operating characteristic (ROC-AUC) and area under the precision-recall curve (PR-AUC) of the results from the validation data was calculated using the PRROC package in R. Average and standard variances of the performance estimates were calculated using the results of the individual cross validation runs. Most discriminative features of the model were obtained using the varImp function from the caret package which ranked the features based on their impact on the decision trees accuracy.

2.4. Gene set enrichment analyses

The lists of predHDF were investigated for enriched gene sets using the R package gprofiler2 (Raudvere et al., 2019). The statistical significance threshold for all GO terms was set to a false discovery rate $p < 0.05$. To focus on non-too-general and non-too-specific GO terms, only terms with more than 10 and less than 200 genes were considered. Redundancy of the terms was removed as described above (section 2.2) using $\alpha = 0.4$.

2.5. Drug selection

Aiming for repurposing known drugs including clinically approved drugs for the identified targets, a manually assembled data repository of drugs was set up based on seven public drug databases: Drugbank, ChEMBL, TTD, PharmGKB, BindingDB, IUPHAR/BPS, and DrugCentral (Thorn et al., 2013; Gilson et al., 2016; Ursu et al., 2017; Wishart et al., 2018; Mendez et al., 2019; Armstrong et al., 2020; Wang et al., 2020). The databases provide information about the mechanisms of actions of the drugs, drug-target interactions and therapeutic indications. We considered only drugs approved by either of the following regulatory agencies: Food and Drug Administration (FDA), Health Canada, European Medicines Agency (EMA), and Japan Pharmaceutical and Medical Devices Agency (PMDA), and drugs currently tested in clinical trials.

2.6. Tissue culture and viruses

Lung adenocarcinoma cell line A549 stably expressing ACE2 and TMPRSS2 (A549-AT; Widera et al., 2021a) was cultured in Minimum Essential Medium (MEM) supplemented with 10% (v/v) fetal calf serum (FCS), 100 U/mL penicillin +100 μg/mL streptomycin (Thermo Fisher; Waltham, Massachusetts, United States) and 2% (v/v) L-glutamine. Calu-3 was cultured in Dulbecco's Modified Eagles Medium (DMEM)/Ham's F12 (Thermo Fisher; Waltham, Massachusetts, United States) supplemented with 10% (v/v) FCS, 100 U/mL penicillin +100 μg/mL streptomycin. Vero E6 was cultivated in DMEM (Thermo Fisher; Waltham, Massachusetts, United States) supplemented with 10% (v/v) FCS and 100 U/mL penicillin +100 μg/mL streptomycin. Antibiotic concentrations were chosen based on the manufacturer's recommendations. Cells were incubated at 37°C, 5% CO_2. If not otherwise indicated, culture reagents were purchased from Sigma (St. Louis, MO, United States).

For all viruses used in this study, the corresponding GenBank accession numbers, and references are listed in Table 2. Briefly, coronavirus isolates were obtained from nasopharyngeal swabs of infected individuals. Swab material was suspended in 1.5 mL phosphate-buffered saline (PBS) and propagated on Caco-2 cells. Cell-free aliquots were stored at −80°C and titers were determined by tissue-culture infectious dose ($TCID_{50}$) assay in A549-AT cells. The H5N1 influenza strain Hongkong/213/03 was received from the World Health Organization (WHO) Influenza Centre (National Institute for Medical Research, London, United Kingdom). Virus stocks were prepared by infecting Vero cells and aliquots were stored at −80°C. Titers were determined as described above.

According to the committee for Biological Safety (ZKBS), the entirety of infectious work presented in this study was conducted under BSL-3 conditions.

2.7. Inhibitors and chemicals

Compounds used in this study comprise Remdesivir (#HY-104077), ICG-001 (#HY-14428; MedChem Express; Monmouth

Junction, New Jersey, United States), C-82 (#S0990) and PRI-724 (#S8968; Selleck Chemicals; Houston, Texas, United States). Compounds were resuspended in DMSO (Carl Roth; Karlsruhe, Germany) and aliquots were stored at −80°C. Recombinant human Wnt-3a (rhWnt-3a; #5036-WN, R&D Systems; Minneapolis, Minnesota, United States) was resuspended according to manufacturer's instructions and stored at −80°C.

2.8. Compound treatment and infection assay

Inhibitors were diluted serially on confluent A549-AT or Calu-3 cells in 1% (v/v) serum MEM and were subsequently infected at an MOI of 0.1 with viruses for 48 h. Cells were fixed with 3% (v/v) paraformaldehyde (PFA) for 20 min and stored at 4°C in PBS until analysis. Cytopathic effects were quantified by confluency measurement using a Spark® Cyto 400 multimode plate reader (Tecan Group Ltd.; Zürich, Switzerland).

2.9. Scratch assay

Scratch assay was implemented to visualize growth hindrance in A549-AT cells upon treatment with PRI-724. $6 \cdot 10^5$ cells were seeded in a 12-well format 1 day prior to assay. After applying scratches with a pipette tip, cells were washed with 1x PBS and treated with 10 μM PRI-724 and the respective amount of DMSO in 1% FCS MEM. Cells were subjected to live cell imaging 0, 24, 48 and 72 h post treatment using a Spark® Cyto 400 multimode plate reader (Tecan Group Ltd.; Zürich, Switzerland).

2.10. Live cell imaging

In order to monitor the effectiveness of PRI-724 in terms of blocking syncytia formation and cytopathic effects in general, live cell imaging was carried out in A549-AT cells in a 96-well format. $3.5 \cdot 10^4$

TABLE 2 Clinical coronavirus isolates used for infection experiments in this study.

Virus	Specification	GenBank ID	References
SARS-CoV-2	B, FFM5/2020	MT358641	Toptan et al. (2020)
SARS-CoV-2	B.1, FFM7/2020	MT358643	Toptan et al. (2020)
SARS-CoV-2	B.1.1.7, FFM-UK7931/2021	MZ427280	Widera et al. (2021b)
SARS-CoV-2	B.1.351, FFM-ZAF1/2021	MW822592	Widera et al. (2021b)
SARS-CoV-2	B.1.617.2, FFM-IND8424/2021	MZ315141	Wilhelm et al. (2021)
SARS-CoV-2	P.2, FFM-BRA1/2021	MW822593	Widera et al. (2021b)
SARS-CoV-2	B.1.429, FFM-CALsprt/2021	MZ317896.2	Wilhelm et al. (2021)
SARS-CoV-2	B.1.617.1, FFM-IND5881/2021	MZ315140	Wilhelm et al. (2021)
SARS-CoV-2	B.1.1.529 BA.1, FFM-ZAF0396/2021	OL800703	Wilhelm et al. (2022)
SARS-CoV-2	B.1.1.529 BA.2, FFM-BA.2-3833/2022	OM617939	Wilhelm et al. (2022)
SARS-CoV-2	B.1.1.529 BA.5, FFM-BA.5-7501/2022	OQ568810	This study
SARS-CoV-1	Frankfurt 1	AY291315	Drosten et al. (2003)
MERS-CoV	EMC/2012	JX869059	Zaki et al. (2012)
Influenza A	H5N1/Hongkong/213/03	NCBI:txid432070	This study

cells were seeded and 1 day after were treated with 10, 3, 1, and 0.3 μM of PRI-724 just prior to infection. DMSO and 3 μM Remdesivir served as negative and positive inhibition controls, respectively. Infections with SARS-CoV-2 B.1.617.2 and SARS-CoV-1 Frankfurt-1 were carried out at an MOI of 0.1. Live cell imaging was performed by using a Spark® Cyto 400 multimode plate reader (Tecan Group Ltd.; Zürich, Switzerland) with hourly measurements of confluence and surface roughness for 48 h.

2.11. Immunofluorescence microscopy

A549-AT cells were infected with SARS-CoV-2 B.1.617.2, SARS-CoV-1 Frankfurt-1 and IAV H5N1/Hongkong/213/03 at an MOI of 0.1 for 18 h. Post fixation with 3% PFA, cells were washed and permeabilized using ice-cold MeOH. Epitopes were saturated by applying blocking solution (20% (v/v) goat serum, 2% (w/v) BSA, 0.3 M glycine, 0.1% Tween-20, 0.002% thimerosal) for at least 1 h. Primary antibodies (Supplementary Table S11) were diluted 1:1000 in blocking solution containing only 1% (v/v) goat serum and were then incubated on cells O/N at 4°C. Cells were washed three times with 1x PBS and were subsequently incubated with secondary anti-IgG-A488 (1:1000) together with 0.2 μg/mL DAPI and for 1 h. Fluorescence imaging was performed with an Operetta CLS™ High Content Analysis System (Perkin Elmer; Waltham, Massachusetts, United States) using integrated Harmony® software v4.9.

2.12. RNA kinetics and RT-qPCR

For assessing intracellular virus RNA replication upon treatment with PRI-724, medium of infected cells was removed 2 hpi and wells were washed three times with 1x PBS and were supplied with fresh 1% serum MEM. Lysates were collected 2, 4, 6, 8, 10, 12, and 24 hpi in A549-AT and 2, 4, 6, 8, 10, 12, 24, 36, and 48 hpi in Calu-3. Samples were prepared and RNA was extracted using RNeasy QIAcube HT (Qiagen; Hilden, Germany) kit according to manufacturer's instructions. RNA was stored at −20°C until analysis. RT-qPCR was carried out using Reliance One-Step Supermix (Bio-Rad; Hercules, CA, United States) to measure total viral N RNA and sg-N RNA (Supplementary Table S10, P1–P12) in a multiplex experiment[5] (Veleanu et al., 2022). Quantifications of cellular genes CDC25A, BIRC5, RNase P (Supplementary Table S10, P13–P19; Spandidos et al., 2008) were carried out using Luna Universal One-Step RT-qPCR Kit (New England Biolabs; Frankfurt (Main), Germany) according to manufacturer's protocol. Fold changes were calculated using the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001).

2.13. TCID$_{50}$ and plaque assay

For the quantification of infectious virus, supernatants were snap-frozen at −80°C. After thawing, supernatants were diluted

on confluent Vero E6 in 96-well plates. Four days post infection, dilutions were analyzed for CPE development. TCID$_{50}$ was calculated according to the Spearman and Kärber algorithm. For plaque assays, supernatants were diluted on confluent Vero E6 in 24-well plates. After 1 h, supernatants were coated with overlay medium (1.5% methyl cellulose (w/v; Merck KgaA; Darmstadt, Germany), 1× MEM, 1% (v/v) FCS, 100 U/mL penicillin +100 μg/mL streptomycin). When plaques reached sufficient size, cells were washed and fixed with 3% PFA for 20 min. Cells were stained with CV staining solution (1% (w/v) crystal violet, 20% (v/v) MeOH).

2.14. Luciferase assay

A549-AT cells were seeded in a 96-well format and were transfected with 0.1 ng of pRL-SV40 (Addgene; #27163) together with either 0.1 μg M50 Super 8x TOPFlash (Addgene; #12456) or M51 Super 8x FOPFlash (TOPFlash mutant; Addgene; #12457) using TransIT®-LT1 Transfection Reagent (Mirus Bio LLC; Madison, Wisconsin, United States). After 24 h, cells were treated with either of the following substances: 10 mM LiCl (Merck KgaA; Darmstadt, Germany), 100 μg/mL rhWnt-3a, 10/3/1/0.3 μM PRI-724. Additionally, cells were treated simultaneously with LiCl and PRI-724. After 24 h, cells were analyzed for luciferase activity using a Dual Luciferase Reporter Assay Kit (Promega; Madison, Wisconsin, United States) according to manufacturer's instructions.

2.15. Western blot

For preparation of protein extracts, cells were lysed using RIPA buffer (150 mM NaCl, 50 mM Tris/HCl pH 8.0, 1% (v/v) Triton X-100, 0.5% (v/v) sodium deoxycholate, 0.3% (v/v) SDS, 2 mM MgCl$_2$). Buffer was supplemented with 5 mM NaF, 1 mM Na$_3$VO$_4$, 20 mM β-glycerophosphate, 1x cOmplete Mini, EDTA-free protease inhibitor cocktail (Merck KgaA; Darmstadt, Germany) and 1 U/mL Pierce™ Universal Nuclease (Thermo Fisher; Waltham, Massachusetts, United States). Lysates were stored at −20°C. Protein concentrations were determined using DC Protein Assay Kit (Bio-Rad; Hercules, CA, United States). Proteins were separated by SDS-PAGE and were blotted onto an Amersham™ Protran® nitrocellulose membrane (Merck KgaA; Darmstadt, Germany) for 1.5 h using const. 120 V. Blots were washed and blocked for at least 1 h using 5% (w/v) BSA/TBS-t (150 mM NaCl, 20 mM Tris, 0.1% (v/v) Tween20). Primary antibody (Supplementary Table S11) incubation was performed overnight at 4°C in 5% BSA-TBS-t. After washing, secondary antibody was incubated for 1 h in 5% BSA-TBS-t. Blots were imaged using CLx imaging device (LI-COR; Lincoln, Nebraska, United States). All full-length western blots are presented in Supplementary Figure S5.

2.16. Cytotoxicity measurement by fluorescence microscopy

Confluent A549-AT cells were treated with 10, 3, 1, 0.3 μM PRI-724 and with the highest corresponding amount of DMSO for

---

48 h in 1% MEM. Cells were fixed with 3% PFA and stained with 0.5 μg/mL Hoechst 33342. Imaging was performed with Olympus IX 81 scanning a total of 9 fields/well with a 3 × 3 scatter at a total magnification of 10x (pixel size = 0.645 μm). Screening was performed using ScanR Acquisition v2.4.0.13 software.

The analysis was performed using KNIME and KNIP (Dietz and Berthold, 2016). The images were filtered using a gaussian convolution (σ = 2 px). A consecutive image thresholding assigned pixels with values greater than the image-mean to the foreground. Cells were detected and labeled by connected component analysis. Cell clumps were afterwards split into single cells using the Wählby algorithm (Wählby et al., 2004). The resulting components were filtered based on their size, components too large or too small were rejected as they most likely do not depict cells. Cell count was measured for the different PRI-724 concentrations (Supplementary Figure S4).

2.17. PI staining and flow cytometry

In a 6-well plate, confluent A549-AT cells were treated with 10 μM PRI-724 and the corresponding amount of DMSO for 24 h and were infected with SARS-CoV-2 (B.1.617.2 isolate, MOI 0.01) in 1% MEM. After that, cells were gently detached and fixed in ice-cold 70% ethanol for 1 h. Cells were washed three times with FACS-PBS (2% (v/v) FCS, 0.1% NaN$_3$) and were then stained overnight at 4°C by applying PI staining buffer (50 μg/mL PI, 2% (v/v) FCS, 200 μg/mL Monarch® RNase A (New England Biolabs; Frankfurt (Main), Germany), 0.1% Igepal CA-630). Cells were analyzed for PI staining using FACSVerse™ Flow Cytometer (BD Biosciences; Mississauga, ON, Canada).

2.18. Statistics

The curve fittings presented for dose-responses to Remdesivir and PRI-724 presented in Figure 1A and Supplementary Figure S1 were performed by applying robust non-linear regression comprising a sigmoidal 4-parameter model for a total of three biological replicates per concentration:

$$Y = Bottom + \frac{Top - Bottom}{1 + \left(\frac{IC50}{X}\right)^{HillSlope}} \quad (6)$$

The exact IC$_{50}$ values calculated from these analyses are listed in Figure 1A and Supplementary Tables S7, S8. For simple group comparisons such as data presented in Figures 1B, 2, 3B, C, 4A, B, 5, one-way ANOVA was generally applied to identify significant differences between treatments. In all tests, equal variances according to the mean were assumed and further analyses were not corrected for multiple testing (Fisher's LSD test). Results were rated significant when $p < 0.05$. For RNA kinetics depicted in Figure 3A, two-way ANOVA was applied to identify significant differences in vRNA levels according to the timepoint of collection and the sort of treatment. Again, equal variances according to the mean were assumed and further analyses were not corrected for multiple testing (Fisher's LSD test). Results were rated significant when $p < 0.05$.

3. Results

3.1. Host dependency factors identified by genome scale screens showed marginal overlap on the gene level, but improved consistency for common cellular processes and functions

First, the overlap of the results from published screening studies were investigated, in the following denoted as Daniloski et al. (2021), Wang R. et al. (2021), Wei et al. (2021), and Zhu et al. (2021). For a balanced comparison, we regarded the top 500 genes with the highest scores of every screen (Figure 6A). Only ACE2 was common among all candidates. The highest overlap (n = 37 genes) was observed between the screens of Daniloski et al., and Zhu et al. Besides ACE2, only 11 genes were common among Wang et al., and Wei et al. All these overlaps were not significantly enriched (using Fisher's Exact Test). Similarly, we found very low correlations while analyzing the ranking of all genes of the screens (independent of the cutoff; Figure 6B). Next, we investigated how grouping of genes into gene sets of common cellular processes and gene functions affect the results. For this, gene set enrichment tests were performed with the same lists of the 500 highest scoring genes (Figure 6C). Interestingly, the overlaps were considerably higher but still not significant. Daniloski et al. and Zhu et al. shared 44/51 gene sets, Daniloski et al., Wang et al., and Zhu et al. shared 5/51 gene sets and 2/51 were identical among Daniloski et al., Wei et al., and Zhu et al. The overlapping gene sets included vesicle-mediated transport to the plasma membrane, endosomal and lysosomal transport, phagosome maturation, transforming growth factor-beta production, post-Golgi vesicle-mediated transport, transition metal ion transport, negative regulation of cell growth, and phosphatidylinositol biosynthetic process (Supplementary Table S1). In summary, we observed limited overlap among the investigated screens. The overlap was higher when regarding gene sets of common cellular processes or molecular function. This supported our concept to set up a machine learning procedure to identify more common patterns among these different datasets.

3.2. Classifiers based on data from the knockout screens and a drug based screen performed well in predicting HDF from the gold standards

We assembled a list of 500 genes commonly high ranking among the investigated knockout screens (top 500 genes of the rank products, see Methods). This list served as the input for the machine learning classifier. Similarly, we selected negative controls from the lowest ranking genes. Furthermore, we set up a descriptor for gene predictions by compiling a comprehensive list of more than 60,000 features for each gene describing its nucleic acid and protein sequences, potential associated cellular processes, its associated compartments, functional domains, molecular functions, its conservation, and its network topology within a protein interaction network. Here, also other omics data of SARS-CoV-2 infected cells, like gene transcription profiles, proteomics and interactions between viral and host proteins were integrated. Performing a cross-validation,
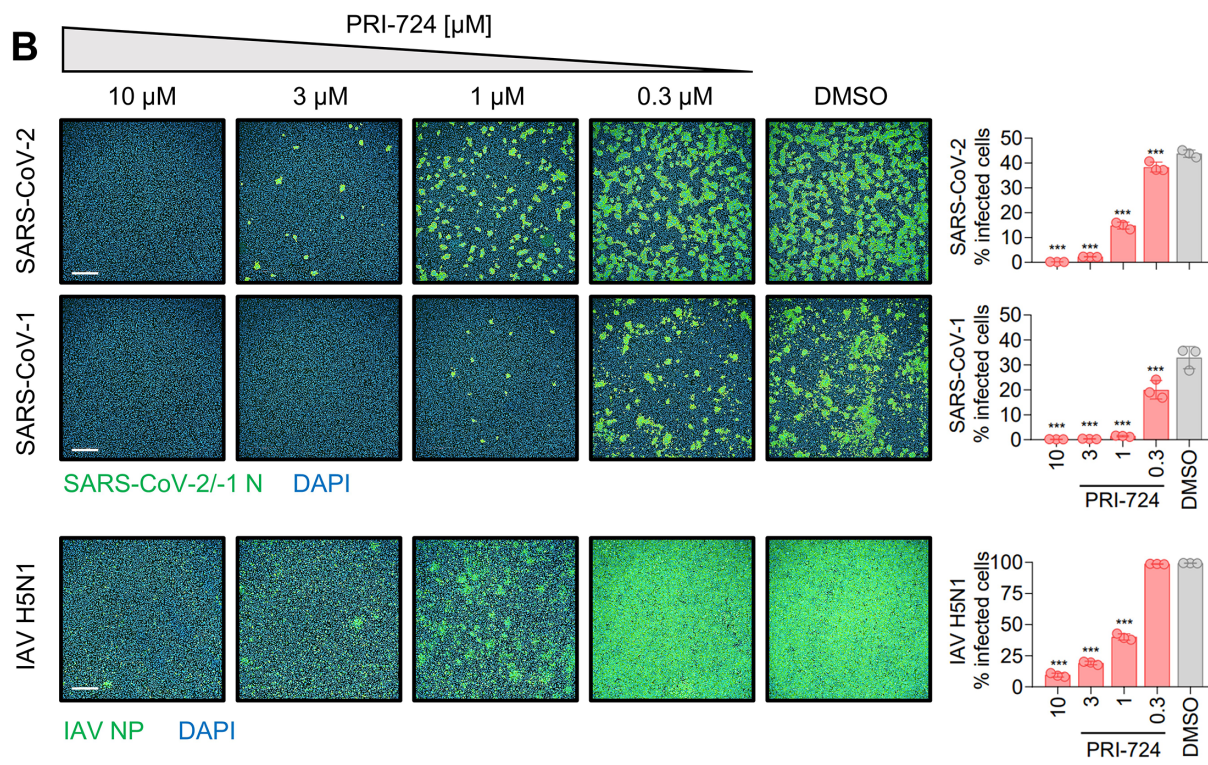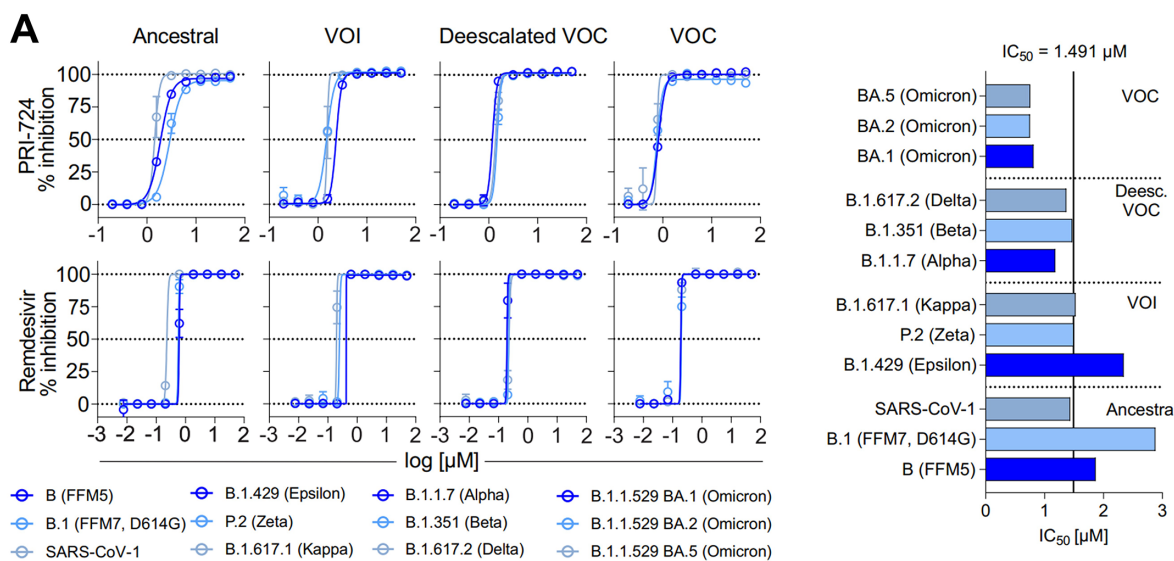
FIGURE 1

Inhibition of SARS-CoV-2, SARS-CoV-1 and IAV H5N1 by PRI-724. **(A)** Dose–response curves of diverse SARS-CoV-2 mutational variants and SARS-CoV-1 to PRI-724 and Remdesivir. Cell confluence was analyzed 48hpi using a Spark® Cyto Imaging System. Dose–response curves were fit to normalized % inhibition. Data represent mean±SD of three biological replicates. The experiment was repeated twice with similar results. IC$_{50}$ values were calculated by applying robust non-linear regression. Mann–Whitney U-test confirmed no significant changes between any dose-responses analyzed. **(B)** Staining of SARS-CoV-2 (B.1.617.2 strain) and SARS-CoV-1 N protein, as well as IAV H5N1 NP 18hpi. Data points represent mean±SD of three biological replicates. The experiment was repeated twice with similar results. Infected cells were determined by Alexa 488/DAPI co-localization using an Operetta® High-Throughput Imaging system. Scale bars=500µm.

in which training and validation data is strictly separated, we observed good performance results with an area under the curve (AUC) of a Receiver Operator Characteristics of 0.82 (1σ = 0.03; Figure 7A). In addition, the same procedure was performed and a classifier learned based on data of a drug screen (in the following drug screen based classifier; Ellinger et al., 2021). Here, we obtained slightly worse performance compared to the knockout-based classifier receiving an AUC of the Receiver Operator Characteristics of 0.71 (1σ = 0.11). Additional performance parameters are supplemented (Supplementary Table S2).
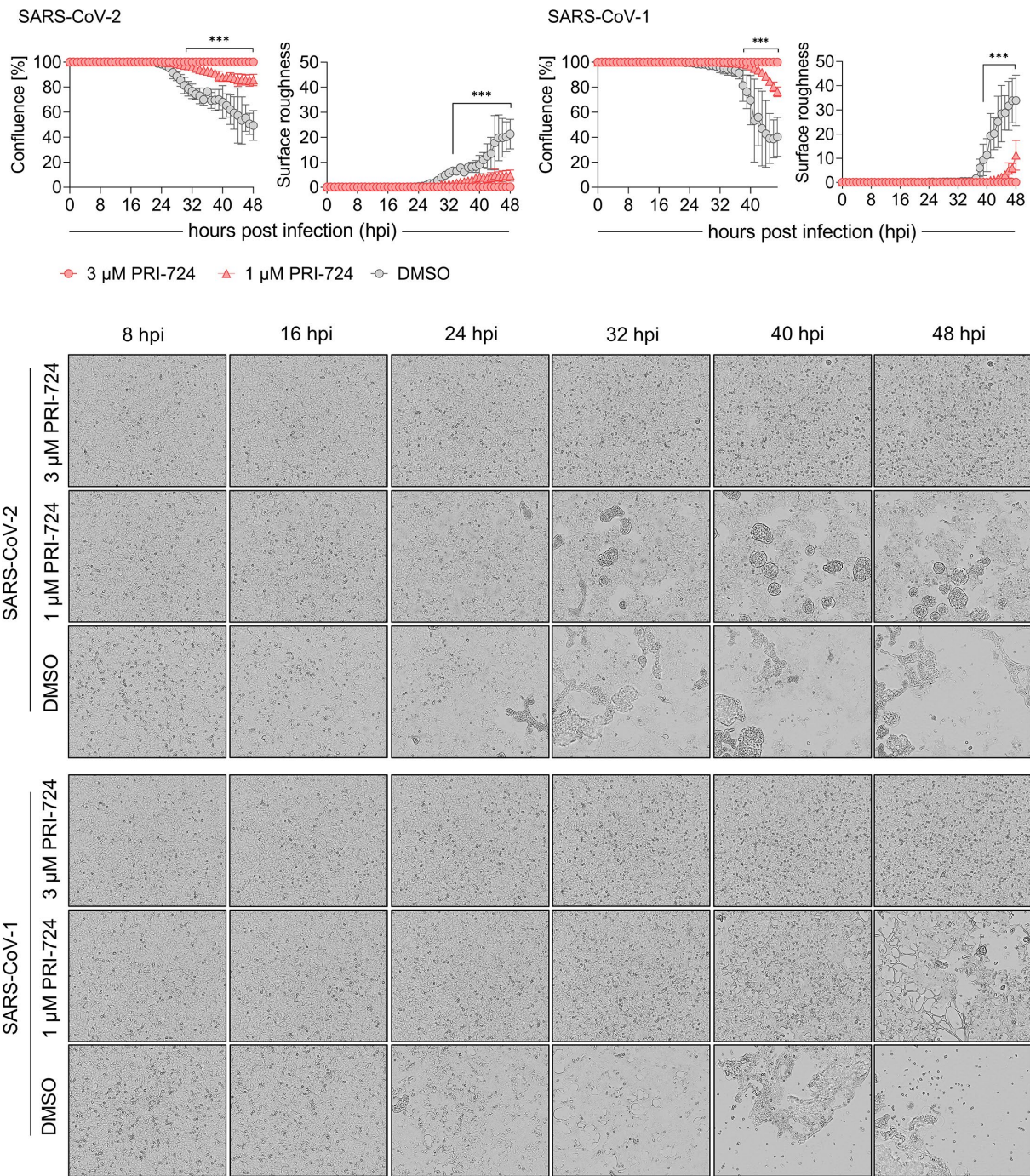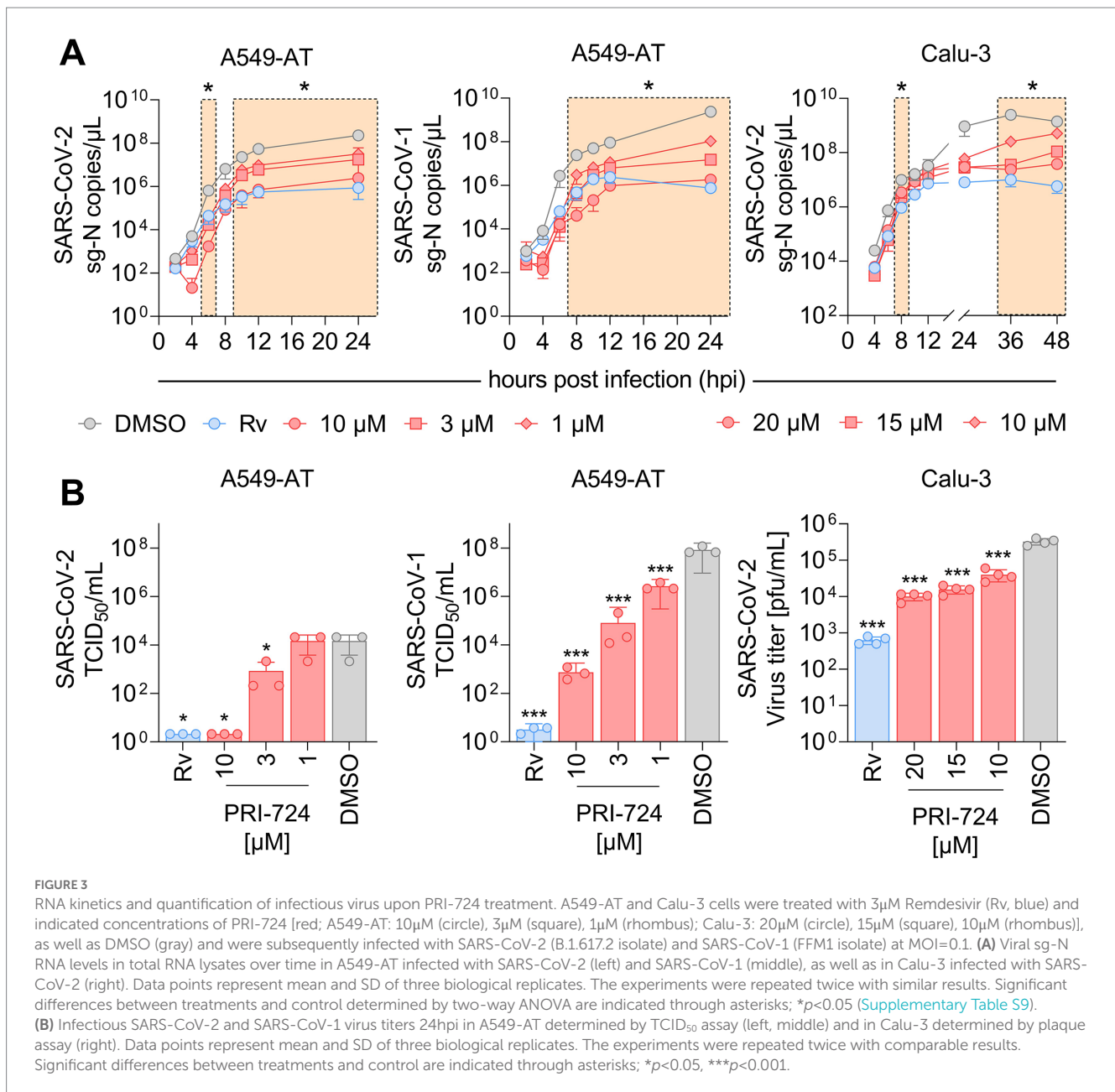
FIGURE 2
Growth kinetics of PRI-724 treated A549-AT cells. A549-AT cells treated with high (3μM) and low (1μM) dose of PRI-724 were subsequently infected with SARS-CoV-2 (B.1.617.2 isolate) and SARS-CoV-1 (FFM1 isolate; MOI=0.1). Cells were incubated at 37°C and 5% CO_2 for 48h in a Spark® Cyto Imaging system. **Top**: Graphical presentation of % confluence and surface roughness for both SARS-CoV-2 and SARS-CoV-1 are shown for measurements hourly performed. **Bottom**: Respective brightfield microscopy images represent critical time points of CPE formation over the course of infection during treatment with PRI-724 and without. Scale bar: 200μm. See also Supplementary Videos S1–S6. Every condition comprises three biological replicates and graphs represent means ± SD; significant differences (***$p<0.001$) are indicated by asterisks obtained by performing multiple t-tests. The experiment was performed twice yielding similar results.

3.3. Predicted HDF are prominently enriched in gene sets related to morphogenesis and development

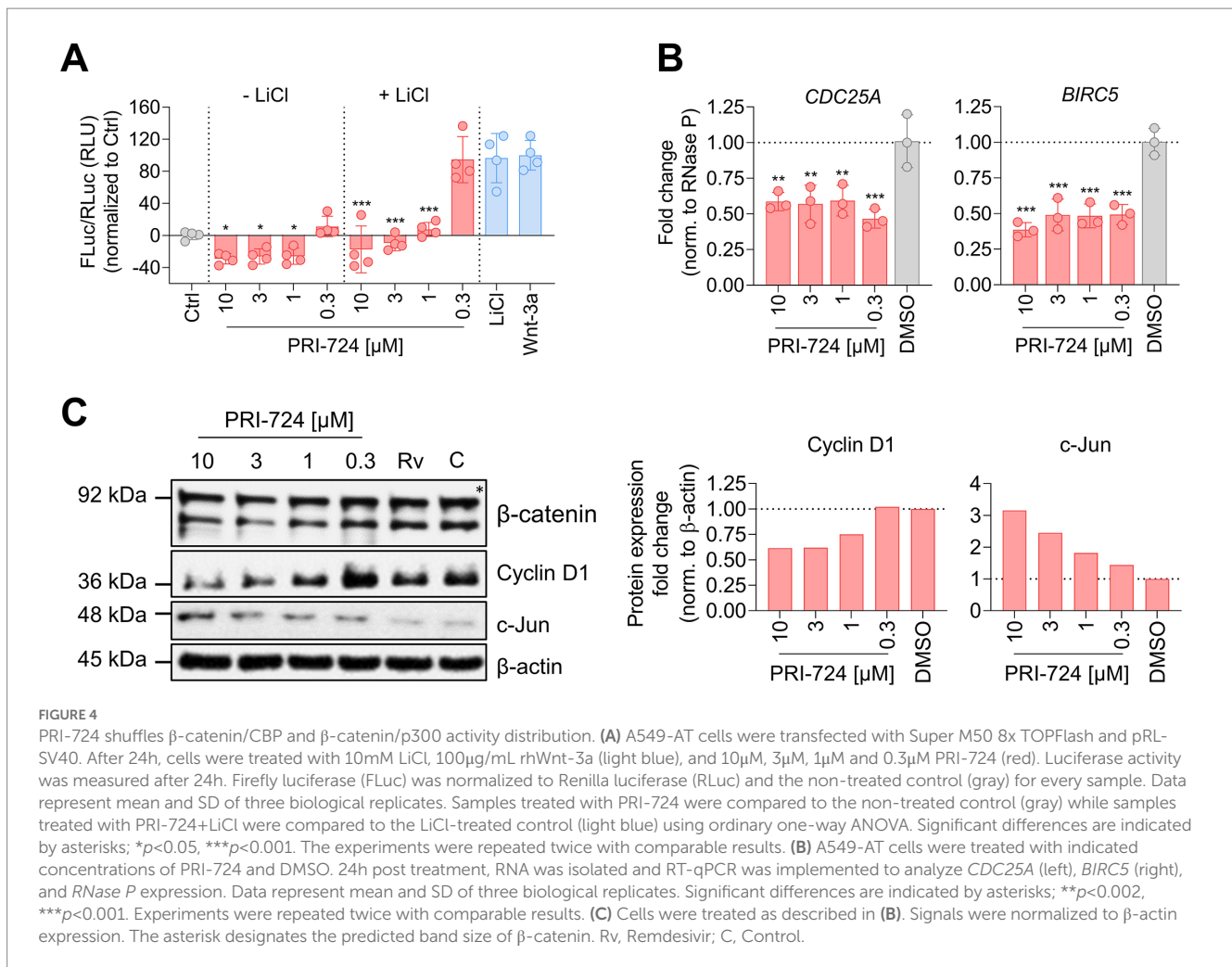We selected 2,182 and 1,989 top-scoring genes identified by knockout- and drug screen-based classifiers, respectively. To identify the most involved cellular mechanisms in the life cycle of the virus, we performed gene set enrichment analysis for each list illuminating 1,098 and 688 gene sets of the respective classifiers (Supplementary Tables S3, S4). Interestingly, we observed a high number of common gene sets ($n=313$) showing consistency among both classifiers (Figure 7B). Regarding these common gene sets,

FIGURE 3

RNA kinetics and quantification of infectious virus upon PRI-724 treatment. A549-AT and Calu-3 cells were treated with 3μM Remdesivir (Rv, blue) and indicated concentrations of PRI-724 [red; A549-AT: 10μM (circle), 3μM (square), 1μM (rhombus); Calu-3: 20μM (circle), 15μM (square), 10μM (rhombus)], as well as DMSO (gray) and were subsequently infected with SARS-CoV-2 (B.1.617.2 isolate) and SARS-CoV-1 (FFM1 isolate) at MOI=0.1. **(A)** Viral sg-N RNA levels in total RNA lysates over time in A549-AT infected with SARS-CoV-2 (left) and SARS-CoV-1 (middle), as well as in Calu-3 infected with SARS-CoV-2 (right). Data points represent mean and SD of three biological replicates. The experiments were repeated twice with similar results. Significant differences between treatments and control determined by two-way ANOVA are indicated through asterisks; *$p<0.05$ (Supplementary Table S9).
(B) Infectious SARS-CoV-2 and SARS-CoV-1 virus titers 24hpi in A549-AT determined by TCID$_{50}$ assay (left, middle) and in Calu-3 determined by plaque assay (right). Data points represent mean and SD of three biological replicates. The experiments were repeated twice with comparable results. Significant differences between treatments and control are indicated through asterisks; *$p<0.05$, ***$p<0.001$.

we further found a high fraction of these gene sets (69 out of 313 gene sets) to be related to morphogenesis and development, followed by 60 gene sets being related to neural processes (Supplementary Table S5). Besides these gene sets, further gene sets were related to signaling, gene regulation, and immune and stress response (Supplementary Table S5).

During development, a large variety of distinct different cellular identities need to be established and maintained in the embryo. Particularly, during developmental lineage reprogramming, a somatic cell can be reprogrammed into a distinct cell type by forced expression of lineage-determining factors (Vierbuchen and Wernig, 2012). Similarly, coronaviruses reprogram their host cell for their specific needs, including their replication (Spagnolo and Hogue, 2000; Shi and Lai, 2005; Reggiori et al., 2010; Nagy and Pogany, 2011; Pizzato et al., 2022). We followed this interesting parallelism and selected a gene list

from the gene sets of development and morphogenesis for further prioritization. Aiming to interfere "reprogramming" with a high impact, we prioritized genes coding for proteins with high connectivity (number of nearest neighbors) and centrality measures in a protein–protein interaction (PPI) network to enhance the impact of a treatment by small molecule inhibitors. This led to a short list of genes with highest ranks (employing rank products) based on these measures in this PPI network. Table 3 lists the top 10, Supplementary Table S6 lists the top ranking genes and their connectivity and centrality values when protein interactions to SARS-CoV-2 proteins are also taken into account. Out of these, β-catenin (CTNNB1) was selected for further analysis due to its high gene expression in SARS-CoV-2 infected cells. Interrogating publicly available compound and drug databases led to the selection of PRI-724 for experimental follow up described in the next sections.

FIGURE 4

PRI-724 shuffles β-catenin/CBP and β-catenin/p300 activity distribution. **(A)** A549-AT cells were transfected with Super M50 8x TOPFlash and pRL-SV40. After 24h, cells were treated with 10mM LiCl, 100μg/mL rhWnt-3a (light blue), and 10μM, 3μM, 1μM and 0.3μM PRI-724 (red). Luciferase activity was measured after 24h. Firefly luciferase (FLuc) was normalized to Renilla luciferase (RLuc) and the non-treated control (gray) for every sample. Data represent mean and SD of three biological replicates. Samples treated with PRI-724 were compared to the non-treated control (gray) while samples treated with PRI-724+LiCl were compared to the LiCl-treated control (light blue) using ordinary one-way ANOVA. Significant differences are indicated by asterisks; $*p<0.05$, $***p<0.001$. The experiments were repeated twice with comparable results. **(B)** A549-AT cells were treated with indicated concentrations of PRI-724 and DMSO. 24h post treatment, RNA was isolated and RT-qPCR was implemented to analyze *CDC25A* (left), *BIRC5* (right), and *RNase P* expression. Data represent mean and SD of three biological replicates. Significant differences are indicated by asterisks; $**p<0.002$, $***p<0.001$. Experiments were repeated twice with comparable results. **(C)** Cells were treated as described in **(B)**. Signals were normalized to β-actin expression. The asterisk designates the predicted band size of β-catenin. Rv, Remdesivir; C, Control.

3.4. PRI-724 inhibits SARS-CoV-2 variants, SARS-CoV-1, and influenza a virus in A549-AT cells

In order to validate *in silico* findings regarding PRI-724 in the context of an infection, various SARS-CoV-2 isolates [Ancestral variants: B (FFM5), B.1 (FFM7, D614G); variants of interest (VOI): P.2 (Zeta), B.1.429 (Epsilon), B.1.617.1 (Kappa); deescalated variants of concern (VOC): B.1.1.7 (Alpha), B.1.351 (Beta), B.1.617.2 (Delta); VOC: B.1.1.529 BA.1, BA.2, and BA.5 (Omicron)], and SARS-CoV-1 Frankfurt-1 were tested. A549-AT cells were infected at an MOI of 0.1 and treated with PRI-724 or Remdesivir for 48h. CPE-related confluency changes were measured using automated label-free brightfield microscopy (Figure 1A). For all viruses tested, a mean IC_{50} of 1.491μM (95%CI 1.087–1.896) and IC_{50} of 0.2916μM (95%CI 0.1948–0.3885) was observed for PRI-724 and Remdesivir, respectively. No significant differences were detected among SARS-CoV-2 variants, even though the B.1 and B.1.429 variants showed ~1.5-fold and ~2-fold higher IC_{50}, respectively, and BA.1, BA.2, and BA.5 variants had ~2-fold lower IC_{50} (Figure 1A; Supplementary Table S7). The overall response pattern among different variants remained similar for Remdesivir. Treatment with the PRI-724 active metabolite C-82 and the analog ICG-001 likewise

inhibited SARS-CoV-2 (B.1.617.2 isolate) in A549-AT with IC_{50} values of 1.005 and 14.85μM, respectively (Supplementary Figure S1A). In addition to that, dose-responses to PRI-724 for SARS-CoV-2 variants B.1, B.1.1.7, B.1.351, P.2, and B.1.617.2 were obtained in Calu-3 cells, with overall increased IC_{50} of 8.448μM (95%CI 7.431–9.465) for PRI-724 and mean IC_{50} of 0.4962 (95%CI 0.2686–0.7237) for Remdesivir (Supplementary Figure S1B; Supplementary Table S8). In Calu-3, PRI-724 further demonstrated inhibition of MERS-CoV ($IC_{50}=22.4$μM) and showed significantly decreased SARS-CoV-1 nucleocapsid (N) expression upon treatment with 10μM, 30μM, and 100μM PRI-724 (Supplementary Figure S1C). In an attempt to validate our results and to test whether PRI-724 can be used against a broader spectrum of RNA viruses, we infected A549-AT cells with Influenza A virus (IAV), SARS-CoV-2 (B.1.617.2 strain), or SARS-CoV-1 (Frankfurt-1 strain), and performed immunofluorescence staining for IAV nucleoprotein (NP) and the coronavirus N (Figure 1B). Significant concentration-dependent reduction of NP expression was evident for treatment with 10, 3, and 1μM PRI-724. Treatment with 10 and 3μM PRI-724 led to ~10-fold and ~4-fold reduction of H5N1 NP+ cells, whereas for these treatments no cells stained positive for SARS-CoV-1 N. Treatment with 3μM reduced SARS-CoV-2 N+ cells ~15-fold and staining was negative with 10μM PRI-724.
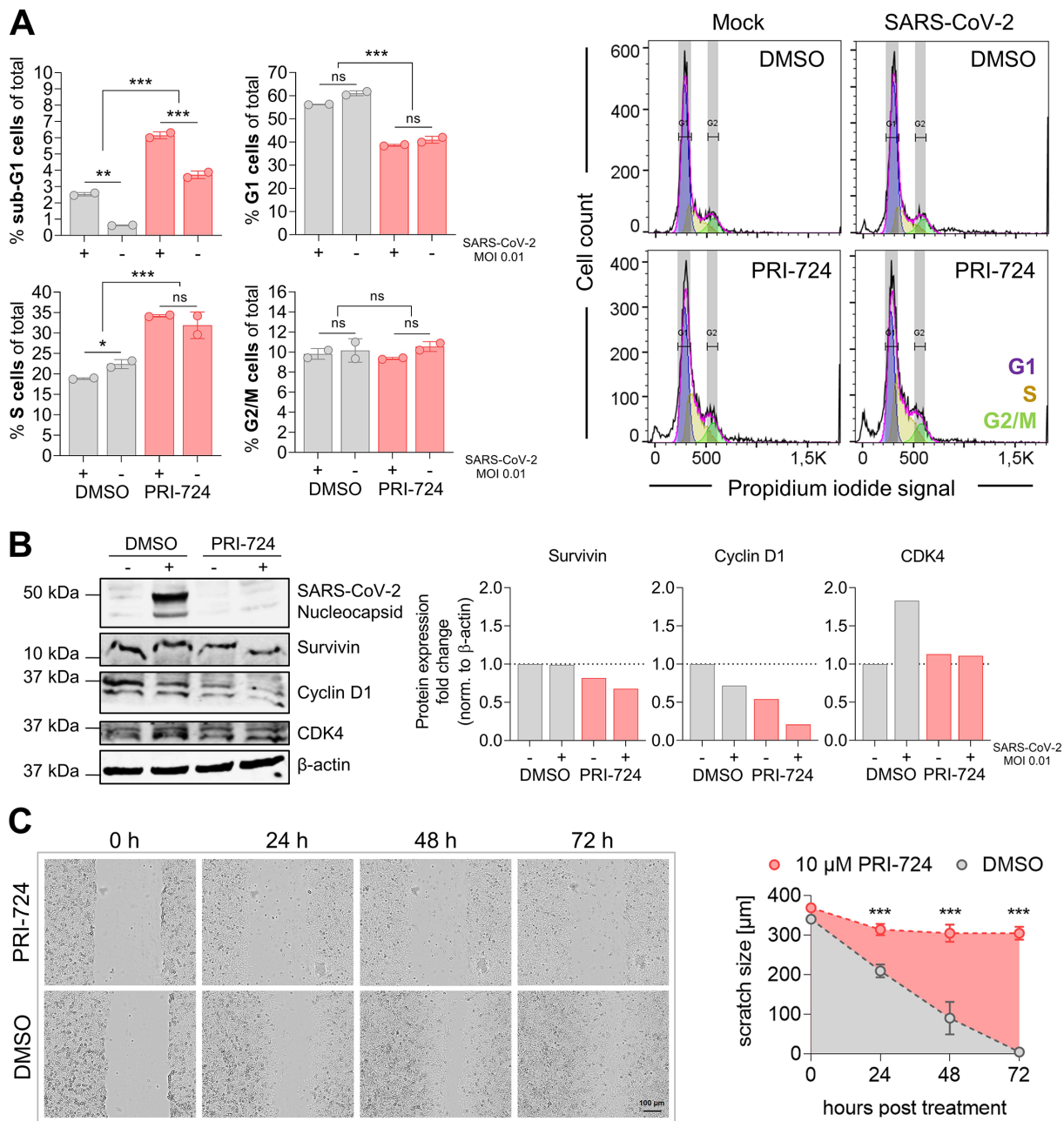
FIGURE 5

Cell cycle analysis in A549-AT upon treatment with PRI-724. **(A)** A549-AT cells were treated with 10 μM PRI-724 or corresponding amounts of DMSO, and were subsequently infected with SARS-CoV-2 (B.1.617.2 isolate, MOI 0.01). After 24h, cells were stained with propidium iodide (PI) following flow cytometric analysis. G1, S and G2/M populations were calculated using the Watson model implemented in FlowJo (v10.8.1). Sub-G1 populations were measured manually. Additionally, cell lysates were analyzed by western blot for N prourvivingrvivin, cyclin D1 and CDK4 expression. Expression levels were normalized to β-actin expression. **(B)** Confluent A549-AT cells were scratched prior to treatment with 10 μM PRI-724 or corresponding DMSO amounts. Live cell imaging was performed 0, 24, 48, and 72h post treatment to assess scratch healing properties using a Spark® Cyto imaging system. All data represent mean and SD of three biological replicates. Experiments were repeated twice. Significant results obtained by one-way ANOVA are indicated by asterisks; *p<0.05, **p<0.002, ***p<0.001.

3.5. PRI-724 inhibits syncytium formation and CPE onset in A549-AT cells

After confirming PRI-724-mediated inhibition of CPE in A549-AT cells at 48 hpi, we examined the timeline of CPE onset and development. Label-free live-cell brightfield imaging was implemented to address viral growth kinetics upon treatment with PRI-724. Cells were treated with 3 μM (high-dose), 1 μM (low-dose) PRI-724, and the highest corresponding amount of DMSO and were subsequently infected with SARS-CoV-2 (B.1.617.2 isolate) or SARS-CoV-1 (Frankfurt-1 strain; MOI = 0.1). Imaging, confluence measurements and surface roughness measurements were performed over the course of 48 h. A549-AT confluence as well as surface roughness were mostly comparable among all treatment conditions until 24 hpi for both
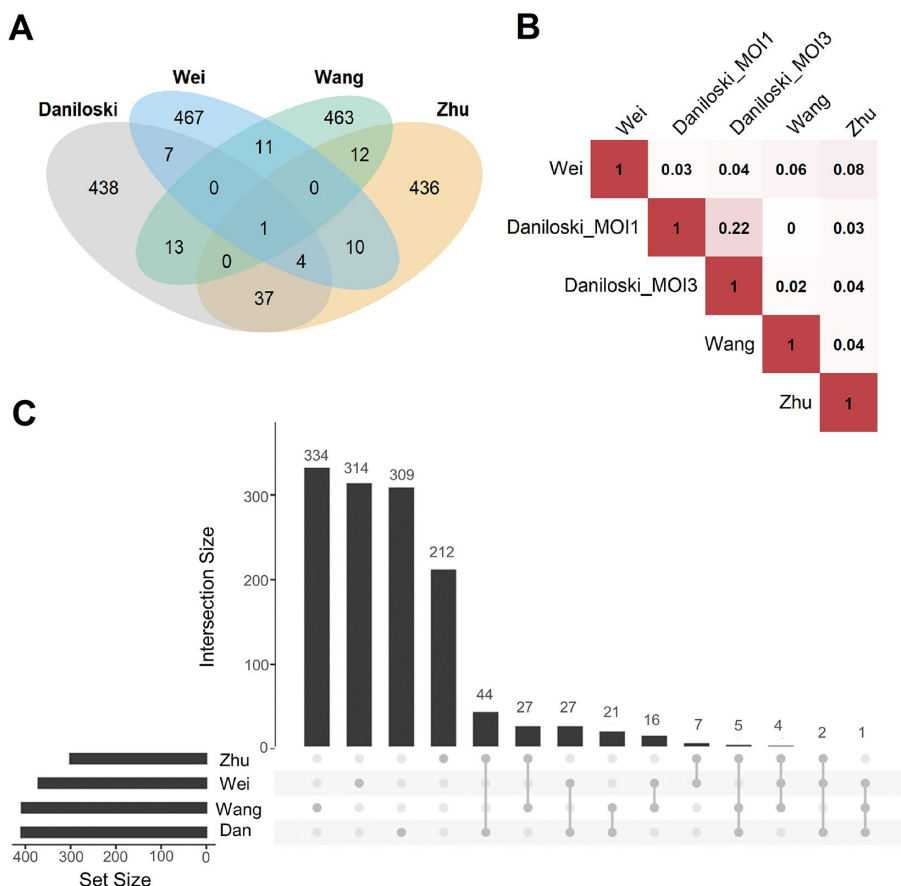
FIGURE 6
Overlap analysis of publicly available host factor identification screens. **(A)** Intersections of the experimentally obtained HDF (500 genes with the highest scores in the screens) among the different knockout screening studies. Only ACE2 (*n*=1, center of the figure) was identified as an HDF in all screens. **(B)** Pairwise correlations between the ranking of all commonly screened genes of the four studies (Daniloski et al., Zhu et al., Wang et al., Wei et al.). Red represents a positive correlation, white no correlation, Daniloski_MOI1: data of the Daniloski screen with MOI=0.01, Daniloski_MOI3: data of the Daniloski screen with MOI=0.3. **(C)** This diagram (upset plot) shows the overlaps of gene sets being enriched of HDF (again considering the 500 genes with the highest scores in the screens), Dan: Daniloski.
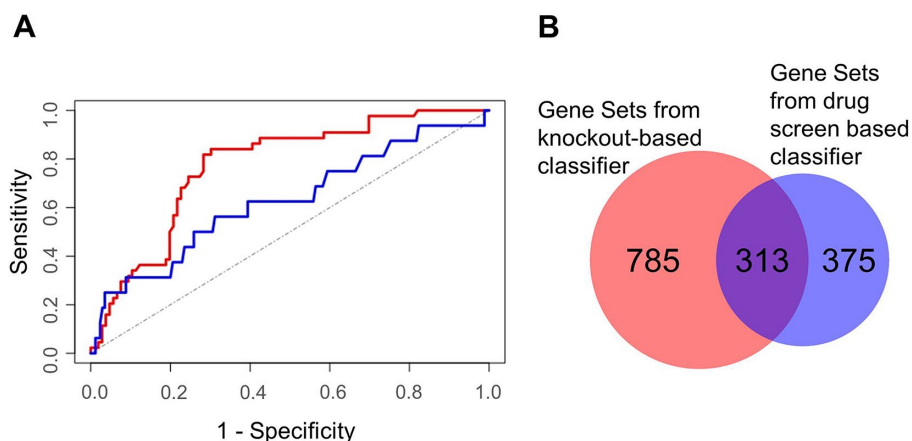

FIGURE 7
Characterization of machine learning performance and overlap of the gene sets derived from the two classifier approaches. **(A)** Machine learning performance results depicted by the Receiver Operator Characteristics of the validation sets. The area under the curve was 0.82 for the knockout screens based classifier (red) and 0.71 for the drug screen based classifier (blue). The (theoretical) result from random guessing is indicated by the dotted line. **(B)** Gene sets being enriched in HDF predicted by the knockout screening based classifier (left) and the drug screening based classifier (right).

TABLE 3 Selected genes of morphogenesis and development with high connectivity, closeness and betweenness centrality.

Gene symbol	Degree[1,2]	Closeness centrality[1–3]	Betweenness centrality[1–3]	Quantile gene expression[4]
EP300	87	0.421	0.104	74
CUL3	82	0.405	0.148	89
TP53	65	0.390	0.068	82
AR	60	0.383	0.052	93
FASN	53	0.373	0.078	94
CTNNB1	40	0.394	0.046	98
HDAC1	57	0.381	0.047	58
MYCN	53	0.368	0.050	Not expressed
STAT3	39	0.382	0.046	94
AKT1	44	0.369	0.041	93

[1]Based on a generic protein interaction network of developmental genes. [2]Determined using Python package Networkx. [3]Decimal places are rounded. [4]Based on the expression levels of the listed gene across several time points of SARS-CoV-2 infected A549 cells (details, see Materials and Methods).

viruses. After 24 hpi, virus-induced cell lysis gradually reduced the confluence and increased surface roughness in DMSO-treated cells. SARS-CoV-2 and SARS-CoV-1 infected cells treated with low-dose PRI-724 started to lyse at 29 and 39 hpi, respectively, while high-dose PRI-724 treated cells showed no substantial loss in confluence over the course of the experiment (Figure 2, top; Supplementary Videos S1–S6). These findings were confirmed by brightfield microscopy at indicated time points (Figure 2, bottom). Here, cell–cell-fusion, i.e., syncytium formation, caused by both viruses, was initially detectable at 16 and 24 hpi, respectively, in DMSO-treated cells and was not present in high-dose PRI-724 treated cells. In low-dose PRI-724 treated cells, SARS-CoV-2 syncytia and CPE were detectable, but much smaller compared to the control (Figure 2, bottom, row 2–3). Although, we found likewise results for SARS-CoV-1, cell lysis initiated at 32 hpi compared to the control (Figure 2, bottom, row 5–6).

3.6. PRI-724 reduces virus RNA replication and infectious virus progeny

Since PRI-724 inhibited CPE development, we questioned whether it impaired virus replication and/or progeny virus production. For the replication analysis, the expression of intracellular subgenomic nucleocapsid (sg-N) RNA was monitored over the course of 24 and 48 h in A549-AT and Calu-3 cells, respectively (Figure 3A). In A549-AT cells, intracellular SARS-CoV-2 (B.1.617.2 isolate) sg-N RNA was significantly reduced at 6, 10, 12, and 24 hpi when treated with 10, 3, and 1 μM PRI-724 just prior to infection. Most remarkable changes were detected 6 and 24 hpi where vRNA was reduced ~370-fold/~96-fold, ~38-fold/~13-fold, and ~20-fold/~7-fold (Supplementary Table S9) compared to the control when treated with 10, 3, and 1 μM PRI-724, respectively. Similarly, in Calu-3, SARS-CoV-2 vRNA was significantly reduced at 8, 36, and 48 hpi, with most noticeable changes at 36 hpi, namely ~105-, ~69-, and ~10-fold (Supplementary Table S9) after treatment with 20, 15, and 10 μM PRI-724, respectively. For SARS-CoV-1 in A549-AT, PRI-724 treatment reduced vRNA at 8, 10, 12, and 24 hpi. In comparison to SARS-CoV-2, the impact on SARS-CoV-1 vRNA levels was more pronounced. 24 hpi SARS-CoV-1 vRNA was reduced ~1,288-, ~157-,

and ~22-fold (Supplementary Table S9) upon treatments with 10, 3, and 1 μM PRI-724, respectively. In A549-AT and Calu-3, Remdesivir treatment appeared more effective in terms of vRNA reduction compared to PRI-724, but timepoints of significant reductions were comparable between both treatments. After 24 h in A549-AT, Remdesivir reduced SARS-CoV-2 and SARS-CoV-1 vRNA 270-fold and ~3,142-fold, respectively (Supplementary Table S9). In Calu-3, SARS-CoV-2 RNA was reduced ~249-fold (Supplementary Table S9) at 24 hpi upon treatment compared to the DMSO control (Figure 3A). In general, reduced sg-N copies correlated with lower total N RNA in A549-AT and Calu-3 cells (Supplementary Figure S2).

To quantify infectious virus in supernatant, $TCID_{50}$ and plaque assays were implemented (Figure 3B). In the supernatant of SARS-CoV-2 infected A549-AT cells, infectious virus was not detectable following treatment with Remdesivir and 10 μM PRI-724. Virus titer was ~17.5-fold reduced upon treatment with 3 μM PRI-724 and no significant reduction was detected with 1 μM PRI-724, which is in line with the obtained RNA levels (Figure 3A). SARS-CoV-1 titer was reduced ~10^8-, ~10^5-, 10^3-, and ~31-fold upon treatment with Remdesivir, 10, 3, and 1 μM PRI-724, respectively. Quantification of SARS-CoV-2 in Calu-3 supernatants at 24 hpi by plaque assay revealed 32.5-, 20.8-, and 8.1-fold reduction in infectious virus titer, when treated with 20, 15, and 10 μM PRI-724, respectively. Overall, a concentration-dependent reduction of SARS-CoV-2 and SARS-CoV-1 vRNA and infectious virus titers upon PRI-724 treatment in A549-AT and Calu-3 cells was observed.

3.7. PRI-724 downregulates β-catenin/CBP target genes *CDC25A*, *BIRC5* and cyclin D1 protein levels and upregulates β-catenin/ p300 target c-Jun

After illuminating the protective *in vitro* properties of PRI-724 during coronavirus infection in A549-AT and Calu-3 cells, we sought to investigate the underlying mechanism. In order to assess β-catenin/CBP activity in A549-AT cells, we implemented a TOPFlash reporter assay (Figure 4A), The TOPFlash plasmid contains TCF/LEF binding sites coupled to a firefly luciferase reporter. Nuclear translocated

β-catenin binds to TCF/LEF transcription factors to recruit the co-factor CBP to the regulatory element (Lee et al., 2010) and thereby activates transcription. Treatment with 10, 3, and 1 μM PRI-724 inhibited TOPFlash reporter activity significantly in comparison to the uninduced control. Moreover, induction of TOPFlash activity by 10 mM LiCl was effectively blocked when using the same concentrations. The control construct with mutant TCF/LEF binding sites, showed almost no luciferase activity and no differences among treatments (Supplementary Figure S3).

We further verified the effect of PRI-724 on β-catenin/CBP activity by RT-qPCR and western blot analysis. The expression of the direct target genes *CDC25A* and *BIRC5* were reduced ~2-fold in PRI-724-treated A549-AT cells (Figure 4B). Also, t protein levels of cyclin D1 were reduced when treated with 10 and 3 μM PRI-724. Reciprocally, β-catenin/p300 target c-Jun was increased in a concentration dependent fashion (Figure 4C). β-catenin expression remained unchanged upon treatments.

3.8. PRI-724 causes cell cycle deregulation in A549-AT cells

It has been reported that PRI-724 causes cell cycle deregulation and block proliferation in several cancer cell lines, such as hepatocellular carcinoma (Gabata et al., 2020), pancreatic (Martinez-Font et al., 2020), and colon cancer (Emami et al., 2004; Kleszcz et al., 2020). We therefore determined whether PRI-724 and/or SARS-CoV-2 infection (MOI 0.01) effects host cell cycle progression. Flow cytometry analysis revealed a significant reduction in G1 (55–60% to 40%) and an increase in S cells (20–22% to 32–35%) upon treatment with PRI-724 compared to DMSO. No significant differences were obtained in G2/M populations. Sub-G1 populations were significantly elevated upon PRI-724 treatment and infected samples showed significantly higher sub-G1 populations than non-infected samples. SARS-CoV-2 infection did not seem to have any significant influence on the cell cycle distribution, although we noticed a slight but significant reduction in the S population of DMSO-treated infected cells compared to non-infected cells. Furthermore, western blot confirmed decreased amounts of Survivin and Cyclin D1 upon treatment with PRI-724. Also, infected cells treated with DMSO but not with PRI-724 showed slightly increased CDK4 expression (Figure 5A). PRI-724-treated A549-AT cells also showed reduced proliferation (Figure 5B). To visualize whether PRI-724 induces lateral growth hindrance, a scratch assay was implemented. 10 μM PRI-724 significantly hampered scratch closure at 24 h post treatment. After 24 h, the scratch size remained at ~310 μm over the course of the experiment while the scratch was almost entirely closed up in the control after 72 h (Figure 5B).

4. Discussion

In this study we provide a comprehensive and robust machine learning-based host factor identification strategy (Xiao et al., 2015) using HDF screens performed for SARS-CoV-2 (Daniloski et al., 2021; Wang R. et al., 2021; Wei et al., 2021; Zhu et al., 2021) and a drug screen (Ellinger et al., 2021) along with validation of the predicted hit in the wet-lab. HDF screens of human or African green monkey cells

infected with SARS-CoV-2 showed limited overlap in their hits. A better consistency was observed when regarding cellular pathways enriched in observed HDF of the screens. We reason this observation by variances among the different screens due to diverse experimental/technical settings or/and observed biological entities, such as different host cell types or different viral strains. To some extent, such differences may be cleaned out, when regarding a gene set representing a cellular pathway which is essential for the virus. It may contain parts of signaling cascades supporting each other making knocked out genes replaceable in specific conditions. Such interdependencies are highly relevant for discovering drug targets and serve as a promising research venue for future investigations. We took another path here and let machines learn which genes are indispensable, independent of the experimental settings in an automated way. Our machines returned lists of predicted HDF that reconstruct HDF and non-HDF class labels well from unseen data when cross-validated. This proves that the machines can capture the data structure in the gold standard lists well. We followed up on investigating the predictions for their biological content, and particularly their enrichments in known cellular processes and functions. Interestingly, we found a major portion of enriched gene sets being related to development and morphogenesis, followed by gene sets being related to neural processes. This is an intriguing observation in itself, suggesting that cellular processes for development and morphogenesis in particular are hijacked by the virus to facilitate reprogramming of the host cell from cellular proliferation to the proliferation of virus particles.

Focusing on gene sets related to development and morphogenesis, we selected β-catenin due to its top ranking connectivity and centrality scores in a constructed protein–protein interaction network, and it was highly expressed during viral infection in our observed cells. β-catenin has been intensively studied, particularly in oncology. It is known for its highly ambivalent roles. In its classical role, it migrates to the nucleus and acts as a cofactor for p300 or CBP (Teo and Kahn, 2010). Investigating hematopoietic stem cells (Rebel et al., 2002) and an embryonal murine stem cell line (F9; Kawasaki et al., 1998; Ugai et al., 1999), it was shown that the choice of the interaction partner CBP or p300 matters for the critical decision between proliferation/(pluri) potency and initiation of differentiation. When interrogating publicly available small compound databases, we found PRI-724 and ICG-001 as β-catenin inhibitors. More specifically, they inhibit the interaction of β-catenin with CBP. PRI-724 is a second-generation structural derivative of ICG-001. We sought to explore the effect of both of these CBP/β-catenin inhibitors. PRI-724 performed better than ICG-001 (Supplementary Figure S1) and showed an inhibitory effect on pathogenic viruses such as SARS-CoV-2 variants, SARS-CoV-1, MERS-CoV, and IAV, in two different cell culture models. The inhibition doses were cell-type dependent which might be due to varying proliferation rates of the cells and different basal levels of β-catenin and CBP.

Viruses manipulate the cell cycle to create resources and cellular conditions that are advantageous for viral replication and assembly. Since PRI-724 is also known to affect cell cycle distribution, we hypothesized that the PRI-724-dependent reduction in virus replication and viral progeny production may be due to cell cycle deregulation. Sui et al. observed that SARS-CoV-2 induces dose dependent S and G2/M arrest at the early phase of infection, particularly, when higher MOIs were used, to facilitate virus replication (Sui et al., 2023), and, in turn, at a higher MOI (0.1), a significant decline of cells in S and G2/M phase at the late phase of infection (24 and 48 hpi). In our experiments we did not observe

any significant difference 24 hpi. This is most likely due to lower MOI (0.01) used in our study and is hence in line with the late phase observations from Sui et al. We detected that PRI-724-dependent G1 population decrease and S population increase is independent of the infection and thus it can effectively reduce virus replication and production for different SARS-CoV-2 variants, SARS-CoV-1, MERS-CoV and IAV. Regarding the drug toxicity, the total number of cells was lower after PRI-724 treatment (Supplementary Figure S4) and flow cytometry analysis showed a slight increase in sub-G1 population which was significantly pronounced following SARS-CoV-2 infection (Figure 5A). The cell distribution pattern we found differs from the previously published work showing an increase in G0/G1 cell population (Arensman et al., 2014; Gabata et al., 2020) while using ICG-001 and C-82 (active compound of PRI-724) when applied to non-infected cells. This observation seems to be cell line or compound type dependent, because another study using colon carcinoma HCT116 cells also found a decrease in G1 and an increase in S cell population following PRI-724 treatment (Kleszcz et al., 2020). In summary, SARS-CoV-2 can reprogram the cell cycle in host cells. Our investigation showed that blocking the interaction between β-catenin and CBP with PRI-724 significantly reduced cytopathic effects, viral RNA replication, and infectious virus production. The treatment caused cell cycle deregulation in the host cell, which appeared to be unrelated to the infection. This suggests that our treatment effectively disrupts the virus-induced reprogramming of the cell cycle, which is essential for viruses to complete their life cycles. PRI-724 treatment has been also clinically investigated. PRI-724 was shown to reduce hepatitis C virus-induced liver fibrosis in mice (Bojkova et al., 2020). These investigations were followed up in two clinical trials of the same group (phase 1 and IIa, respectively; Kimura et al., 2017, 2022). The authors concluded that PRI-724 treatment was well tolerated by patients with HBV and HCV induced liver-cirrhosis and showed improvements of the pathology of concern in several patients. Besides this, blockade of β-catenin/CBP reversed pulmonary fibrosis (Henderson et al., 2010), which is a central COVID-19 complication (Zou et al., 2021). A recent study confirmed overall higher β-catenin levels in SARS-CoV-2 infected patients and targeting reduced virus shedding *in vitro* (Chatterjee et al., 2022).

A limitation of our study is that the machine learning procedure bases on screening data of the four initial screens which were at hand when we developed the computational method. New screens have been, and certainly more will be published as means to better understand the mechanisms of this wide spread and evolving virus. As a future outlook, we would like to repeat the computational analysis and include new screens that have been published since the initiation of this project. We thereby might identify additional HDF that are relevant for the SARS-CoV-2 life cycle. We would like to emphasize that further investigations with 3D models of primary cells, e.g., human bronchial epithelial cells (HBEpC) and human nasal epithelial cells (HNEpC) are crucial to validate our observations. 3D models have the ability to replicate cellular behavior *in vivo* more effectively than immortalized or cancer cell lines. These cell models can also offer a more accurate representation of host-pathogen interactions *in vitro* compared to 2D systems, resulting in more precise data and a superior prediction of drug efficacy and toxicity.

In conclusion, the machine learning approach brought up interesting connections between viral replication and the reprogramming of host cells in the context of SARS-CoV-2 infection. It may also suit for host factor studies of other viruses and a comparison between different virus entities may lead to intriguing new aspects in virus biology. We observed a strong correlation between the inhibition of CPE, vRNA, and infectious virus production for the tested concentrations in two different cell lines, making PRI-724 a promising HDA against SARS-CoV-2 and SARS-CoV-1 in our models.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding authors.

Author contributions

TT, HE, and RK conceived the study. AV-G, AH, TB, MO, and RK performed data analysis and machine learning. MK and TT designed and performed the experiments. NB and PW performed the experiments on cytotoxicity measurement by fluorescence microscopy. TT, SC, RK, and HE acquired funding. All authors contributed to the article and approved the submitted version.

Funding

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2023.1193320/full#supplementary-material

References

Acencio, M. L., and Lemke, N. (2009). Towards the prediction of essential genes by integration of network topology, cellular localization and biological process information. *BMC Bioinformatics.* 10:290. doi: 10.1186/1471-2105-10-290

Akiyama, T. (2000). Wnt/beta-catenin signaling. *Cytokine Growth Factor Rev.* 11, 273–282. doi: 10.1016/S1359-6101(00)00011-3

Alipanahi, B., Delong, A., Weirauch, M. T., and Frey, B. J. (2015). Predicting the sequence specificities of DNA- and RNA-binding proteins by deep learning. *Nat. Biotechnol.* 33, 831–838. doi: 10.1038/nbt.3300

Almagro Armenteros, J. J., Sønderby, C. K., Sønderby, S. K., Nielsen, H., and Winther, O. (2017). DeepLoc: prediction of protein subcellular localization using deep learning. *Bioinformatics* 33, 3387–3395. doi: 10.1093/bioinformatics/btx431

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402. doi: 10.1093/nar/25.17.3389

Arensman, M. D., Telesca, D., Lay, A. R., Kershaw, K. M., Wu, N., Donahue, T. R., et al. (2014). The CREB-binding protein inhibitor ICG-001 suppresses pancreatic cancer growth. *Mol. Cancer Ther.* 13, 2303–2314. doi: 10.1158/1535-7163.MCT-13-1005

Armstrong, J. F., Faccenda, E., Harding, S. D., Pawson, A. J., Southan, C., Sharman, J. L., et al. (2020). The IUPHAR/BPS guide to PHARMACOLOGY in 2020: extending immunopharmacology content and introducing the IUPHAR/MMV guide to MALARIA PHARMACOLOGY. *Nucleic Acids Res.* 48, D1006–D1021. doi: 10.1093/nar/gkz951

Aromolaran, O., Beder, T., Oswald, M., Oyelade, J., Adebiyi, E., and Koenig, R. (2020). Essential gene prediction in *Drosophila melanogaster* using machine learning approaches based on sequence and functional features. *Comput. Struct. Biotechnol. J.* 18, 612–621. doi: 10.1016/j.csbj.2020.02.022

Beder, T., Aromolaran, O., Dönitz, J., Tapanelli, S., Adedeji, E. O., Adebiyi, E., et al. (2021). Identifying essential genes across eukaryotes by machine learning. *NAR Genomics Bioinformatics* 3:110. doi: 10.1093/nargab/lqab110

Beigel, J. H., Tomashek, K. M., Dodd, L. E., Mehta, A. K., Zingman, B. S., Kalil, A. C., et al. (2020). Remdesivir for the treatment of Covid-19—final report. *N. Engl. J. Med.* 383, 1813–1826. doi: 10.1056/NEJMoa2007764

Bojkova, D., Klann, K., Koch, B., Widera, M., Krause, D., Ciesek, S., et al. (2020). Proteomics of SARS-CoV-2-infected host cells reveals therapy targets. *Nature* 583, 469–472. doi: 10.1038/s41586-020-2332-7

Bordonaro, M., and Lazarova, D. L. (2016). Determination of the role of CBP- and p300-mediated Wnt signaling on colonic cells. *JMIR Res Protoc.* 5:e66. doi: 10.2196/resprot.5495

Charif, D., Thioulouse, J., Lobry, J. R., and Perrière, G. (2005). Online synonymous codon usage analyses with the ade4 and seqinR packages. *Bioinformatics* 21, 545–547. doi: 10.1093/bioinformatics/bti037

Chatterjee, S., Keshry, S. S., Ghosh, S., Ray, A., and Chattopadhyay, S. (2022). Versatile β-catenin is crucial for SARS-CoV-2 infection. *Microbiol Spectr.* 10:e0167022. doi: 10.1128/spectrum.01670-22

Daniloski, Z., Jordan, T. X., Wessels, H. H., Hoagland, D. A., Kasela, S., Legut, M., et al. (2021). Identification of required host factors for SARS-CoV-2 infection in human cells. *Cells* 184, 92–105.e16. doi: 10.1016/j.cell.2020.10.030

Dejnirattisai, W., Zhou, D., Supasa, P., Liu, C., Mentzer, A. J., Ginn, H. M., et al. (2021). Antibody evasion by the P.1 strain of SARS-CoV-2. *Cells* 184, 2939–2954.e9. doi: 10.1016/j.cell.2021.03.055

Dietz, C., and Berthold, M. R. (2016). KNIME for open-source bioimage analysis: a tutorial. *Adv. Anat. Embryol. Cell Biol.* 219, 179–197. doi: 10.1007/978-3-319-28549-8_7

Drosten, C., Günther, S., Preiser, W., van der Werf, S., Brodt, H. R., Becker, S., et al. (2003). Identification of a novel coronavirus in patients with severe acute respiratory syndrome. *N. Engl. J. Med.* 348, 1967–1976. doi: 10.1056/NEJMoa030747

Edinger TOPohl, M. O., and Stertz, S. (2014). Entry of influenza a virus: host factors and antiviral targets. *J. Gen. Virol.* 95, 263–277. doi: 10.1099/vir.0.059477-0

Ellinger, B., Bojkova, D., Zaliani, A., Cinatl, J., Claussen, C., Westhaus, S., et al. (2021). A SARS-CoV-2 cytopathicity dataset generated by high-content screening of a large drug repurposing collection. *Sci Data* 8:70. doi: 10.1038/s41597-021-00848-4

Emami, K. H., Nguyen, C., Ma, H., Kim, D. H., Jeong, K. W., Eguchi, M., et al. (2004). A small molecule inhibitor of beta-catenin/CREB-binding protein transcription [corrected]. *Proc. Natl. Acad. Sci. U. S. A.* 101, 12682–12687. doi: 10.1073/pnas.0404875101

Flores-Vega, V. R., Monroy-Molina, J. V., Jiménez-Hernández, L. E., Torres, A. G., Santos-Preciado, J. I., and Rosales-Reyes, R. (2022). SARS-CoV-2: evolution and emergence of new viral variants. *Viruses* 14:653. doi: 10.3390/v14040653

Gabata, R., Harada, K., Mizutani, Y., Ouchi, H., Yoshimura, K., Sato, Y., et al. (2020). Anti-tumor activity of the small molecule inhibitor PRI-724 against β-catenin-activated hepatocellular carcinoma. *Anticancer Res* 40, 5211–5219. doi: 10.21873/anticanres.14524

Gandhi, S., Klein, J., Robertson, A. J., Peña-Hernández, M. A., Lin, M. J., Roychoudhury, P., et al. (2022). De novo emergence of a remdesivir resistance mutation during treatment of persistent SARS-CoV-2 infection in an immunocompromised patient: a case report. *Nat. Commun.* 13:1547. doi: 10.1038/s41467-022-29104-y

Gilson, M. K., Liu, T., Baitaluk, M., Nicola, G., Hwang, L., and Chong, J. (2016). BindingDB in 2015: a public database for medicinal chemistry, computational chemistry and systems pharmacology. *Nucleic Acids Res.* 44, D1045–D1053. doi: 10.1093/nar/gkv1072

Hammond, J., Leister-Tebbe, H., Gardner, A., Abreu, P., Bao, W., Wisemandle, W., et al. (2022). Oral Nirmatrelvir for high-risk, nonhospitalized adults with Covid-19. *N. Engl. J. Med.* 386, 1397–1408. doi: 10.1056/NEJMoa2118542

Henderson, W. R., Chi, E. Y., Ye, X., Nguyen, C., Tzang, T. Y., Zhou, B., et al. (2010). Inhibition of Wnt/beta-catenin/CREB binding protein (CBP) signaling reverses pulmonary fibrosis. *Proc. Natl. Acad. Sci. U. S. A.* 107, 14309–14314. doi: 10.1073/pnas.1001520107

Hershberg, R., and Petrov, D. A. (2009). General rules for optimal codon choice. *PLoS Genet.* 5:e1000556. doi: 10.1371/journal.pgen.1000556

Hogan, J. I., Duerr, R., Dimartino, D., Marier, C., Hochman, S. E., Mehta, S., et al. (2023). Remdesivir resistance in transplant recipients with persistent coronavirus disease 2019. *Clin Infect Dis Off Publ Infect Dis Soc Am.* 76, 342–345. doi: 10.1093/cid/ciac769

Hörhold, F., Eisel, D., Oswald, M., Kolte, A., Röll, D., Osen, W., et al. (2020). Reprogramming of macrophages employing gene regulatory and metabolic network models. *PLoS Comput. Biol.* 16:e1007657. doi: 10.1371/journal.pcbi.1007657

Iketani, S., Liu, L., Guo, Y., Liu, L., Chan, J. F. W., Huang, Y., et al. (2022). Antibody evasion properties of SARS-CoV-2 omicron sublineages. *Nature* 604, 553–556. doi: 10.1038/s41586-022-04594-4

Jayk Bernal, A., Gomes da Silva, M. M., Musungaie, D. B., Kovalchuk, E., Gonzalez, A., Delos Reyes, V., et al. (2022). Molnupiravir for Oral treatment of Covid-19 in nonhospitalized patients. *N. Engl. J. Med.* 386, 509–520. doi: 10.1056/NEJMoa2116044

Kawasaki, H., Eckner, R., Yao, T. P., Taira, K., Chiu, R., Livingston, D. M., et al. (1998). Distinct roles of the co-activators p300 and CBP in retinoic-acid-induced F9-cell differentiation. *Nature* 393, 284–289. doi: 10.1038/30538

Kimura, K., Ikoma, A., Shibakawa, M., Shimoda, S., Harada, K., Saio, M., et al. (2017). Safety, tolerability, and preliminary efficacy of the anti-fibrotic small molecule PRI-724, a CBP/β-catenin inhibitor, in patients with hepatitis C virus-related cirrhosis: a single-center, open-label, dose escalation phase 1 trial. *EBioMedicine* 23, 79–87. doi: 10.1016/j.ebiom.2017.08.016

Kimura, K., Kanto, T., Shimoda, S., Harada, K., Kimura, M., Nishikawa, K., et al. (2022). Safety, tolerability, and anti-fibrotic efficacy of the CBP/β-catenin inhibitor PRI-724 in patients with hepatitis C and B virus-induced liver cirrhosis: an investigator-initiated, open-label, non-randomised, multicentre, phase 1/2a study. *EBioMedicine* 80:104069. doi: 10.1016/j.ebiom.2022.104069

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.e4. doi: 10.1016/j.molcel.2020.08.006

Kleszcz, R., Krajka-Kuźniak, V., and Paluszczak, J. (2020). Porcupine and CBP/β-catenin are the most suitable targets for the inhibition of canonical Wnt signaling in colorectal carcinoma cell lines*. *Postępy Hig Med Dośw.* 74, 224–235. doi: 10.5604/01.3001.0014.2497

Lee, C. H., Hung, H. W., Hung, P. H., and Shieh, Y. S. (2010). Epidermal growth factor receptor regulates beta-catenin location, stability, and transcriptional activity in oral cancer. *Mol. Cancer* 9:64. doi: 10.1186/1476-4598-9-64

Liu, J., Xiao, Q., Xiao, J., Niu, C., Li, Y., Zhang, X., et al. (2022). Wnt/β-catenin signalling: function, biological mechanisms, and therapeutic opportunities. *Signal Transduct. Target. Ther.* 7:3. doi: 10.1038/s41392-021-00762-6

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the $2-\Delta\Delta CT$ method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262

Mahajan, S., Choudhary, S., Kumar, P., and Tomar, S. (2021). Antiviral strategies targeting host factors and mechanisms obliging +ssRNA viral pathogens. *Bioorg. Med. Chem.* 46:116356. doi: 10.1016/j.bmc.2021.116356

Martinez-Font, E., Pérez-Capó, M., Ramos, R., Felipe, I., Garcías, C., Luna, P., et al. (2020). Impact of Wnt/β-catenin inhibition on cell proliferation through CDC25A downregulation in soft tissue sarcomas. *Cancer* 12:2556. doi: 10.3390/cancers12092556

Mendez, D., Gaulton, A., Bento, A. P., Chambers, J., De Veij, M., Félix, E., et al. (2019). ChEMBL: towards direct deposition of bioassay data. *Nucleic Acids Res.* 47, D930–D940. doi: 10.1093/nar/gky1075

Nagy, P. D., and Pogany, J. (2011). The dependence of viral RNA replication on co-opted host factors. *Nat. Rev. Microbiol.* 10, 137–149. doi: 10.1038/nrmicro2692

Oughtred, R., Stark, C., Breitkreutz, B. J., Rust, J., Boucher, L., Chang, C., et al. (2019). The BioGRID interaction database: 2019 update. *Nucleic Acids Res.* 47, D529–D541. doi: 10.1093/nar/gky1079

Pizzato, M., Baraldi, C., Boscato Sopetto, G., Finozzi, D., Gentile, C., Gentile, M. D., et al. (2022). SARS-CoV-2 and the host cell: a tale of interactions. *Front Virol* 1:388. doi: 10.3389/fviro.2021.815388

Planas, D., Veyer, D., Baidaliuk, A., Staropoli, I., Guivel-Benhassine, F., Rajah, M. M., et al. (2021). Reduced sensitivity of SARS-CoV-2 variant Delta to antibody neutralization. *Nature* 596, 276–280. doi: 10.1038/s41586-021-03777-9

Pruitt, K. D., Tatusova, T., and Maglott, D. R. (2005). NCBI reference sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res.* 33, D501–D504. doi: 10.1093/nar/gki025

Raudvere, U., Kolberg, L., Kuzmin, I., Arak, T., Adler, P., Peterson, H., et al. (2019). G:profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res.* 47, W191–W198. doi: 10.1093/nar/gkz369

Rebel, V. I., Kung, A. L., Tanner, E. A., Yang, H., Bronson, R. T., and Livingston, D. M. (2002). Distinct roles for CREB-binding protein and p300 in hematopoietic stem cell self-renewal. *Proc. Natl. Acad. Sci. U. S. A.* 99, 14789–14794. doi: 10.1073/pnas.232568499

Reggiori, F., Monastyrska, I., Verheije, M. H., Calì, T., Ulasli, M., Bianchi, S., et al. (2010). Coronaviruses hijack the LC3-I-positive EDEMosomes, ER-derived vesicles exporting short-lived ERAD regulators, for replication. *Cell Host Microbe* 7, 500–508. doi: 10.1016/j.chom.2010.05.013

Shi, S. T., and Lai, M. M. C. (2005). Viral and cellular proteins involved in coronavirus replication. *Curr. Top. Microbiol. Immunol.* 287, 95–131. doi: 10.1007/3-540-26765-4_4

Smedley, D., Haider, S., Ballester, B., Holland, R., London, D., Thorisson, G., et al. (2009). BioMart--biological queries made easy. *BMC Genomics* 10:22. doi: 10.1186/1471-2164-10-22

Spagnolo, J. F., and Hogue, B. G. (2000). Host protein interactions with the 3′ end of bovine coronavirus RNA and the requirement of the poly(a) tail for coronavirus defective genome replication. *J. Virol.* 74, 5053–5065. doi: 10.1128/JVI.74.11.5053-5065.2000

Spandidos, A., Wang, X., Wang, H., Dragnev, S., Thurber, T., and Seed, B. (2008). A comprehensive collection of experimentally validated primers for polymerase chain reaction quantitation of murine transcript abundance. *BMC Genomics* 9:633. doi: 10.1186/1471-2164-9-633

Sui, L., Li, L., Zhao, Y., Zhao, Y., Hao, P., Guo, X., et al. (2023). Host cell cycle checkpoint as antiviral target for SARS-CoV-2 revealed by integrative transcriptome and proteome analyses. *Signal Transduct. Target. Ther.* 8:21. doi: 10.1038/s41392-022-01296-1

Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., et al. (2019). STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* 47, D607–D613. doi: 10.1093/nar/gky1131

Teo, J. L., and Kahn, M. (2010). The Wnt signaling pathway in cellular proliferation and differentiation: a tale of two coactivators. *Adv. Drug Deliv. Rev.* 62, 1149–1155. doi: 10.1016/j.addr.2010.09.012

Thorn, C. F., Klein, T. E., and Altman, R. B. (2013). PharmGKB: the pharmacogenomics Knowledge Base. *Methods Mol Biol* 1015, 311–320. doi: 10.1007/978-1-62703-435-7_20

Toptan, T., Hoehl, S., Westhaus, S., Bojkova, D., Berger, A., Rotter, B., et al. (2020). Optimized qRT-PCR approach for the detection of intra- and extra-cellular SARS-CoV-2 RNAs. *Int. J. Mol. Sci.* 21:4396. doi: 10.3390/ijms21124396

Ugai, H., Uchida, K., Kawasaki, H., and Yokoyama, K. K. (1999). The coactivators p300 and CBP have different functions during the differentiation of F9 cells. *J Mol Med Berl Ger.* 77, 481–494. doi: 10.1007/s001099900021

Ursu, O., Holmes, J., Knockel, J., Bologa, C. G., Yang, J. J., Mathias, S. L., et al. (2017). DrugCentral: online drug compendium. *Nucleic Acids Res.* 45, D932–D939. doi: 10.1093/nar/gkw993

Van Egeren, D., Novokhodko, A., Stoddard, M., Tran, U., Zetter, B., Rogers, M. S., et al. (2021). Controlling long-term SARS-CoV-2 infections can slow viral evolution and reduce the risk of treatment failure. *Sci. Rep.* 11:22630. doi: 10.1038/s41598-021-02148-8

Veleanu, A., Kelch, M. A., Ye, C., Flohr, M., Wilhelm, A., Widera, M., et al. (2022). Molecular analyses of clinical isolates and recombinant SARS-CoV-2 carrying B.1 and B.1.617.2 spike mutations suggest a potential role of non-spike mutations in infection kinetics. *Viruses* 14:2017. doi: 10.3390/v14092017

Vierbuchen, T., and Wernig, M. (2012). Molecular roadblocks for cellular reprogramming. *Mol. Cell* 47, 827–838. doi: 10.1016/j.molcel.2012.09.008

Wagoner, J., Herring, S., Hsiang, T. Y., Ianevski, A., Biering, S. B., Xu, S., et al. (2022). Combinations of host- and virus-targeting antiviral drugs confer synergistic suppression of SARS-CoV-2. *Microbiol Spectr.* 10:e0333122. doi: 10.1128/spectrum.03331-22

Wählby, C., Sintorn, I. M., Erlandsson, F., Borgefors, G., and Bengtsson, E. (2004). Combining intensity, edge and shape information for 2D and 3D segmentation of cell nuclei in tissue sections. *J. Microsc.* 215, 67–76. doi: 10.1111/j.0022-2720.2004.01338.x

Wang, P., Nair, M. S., Liu, L., Iketani, S., Luo, Y., Guo, Y., et al. (2021). Antibody resistance of SARS-CoV-2 variants B.1.351 and B.1.1.7. *Nature* 593, 130–135. doi: 10.1038/s41586-021-03398-2

Wang, R., Simoneau, C. R., Kulsuptrakul, J., Bouhaddou, M., Travisano, K. A., Hayashi, J. M., et al. (2021). Genetic screens identify host factors for SARS-CoV-2 and common cold coronaviruses. *Cells* 184, 106–119.e14. doi: 10.1016/j.cell.2020.12.004

Wang, Y., Zhang, S., Li, F., Zhou, Y., Zhang, Y., Wang, Z., et al. (2020). Therapeutic target database 2020: enriched resource for facilitating research and early development of targeted therapeutics. *Nucleic Acids Res.* 48, D1031–D1041. doi: 10.1093/nar/gkz981

Wei, J., Alfajaro, M. M., DeWeirdt, P. C., Hanna, R. E., Lu-Culligan, W. J., Cai, W. L., et al. (2021). Genome-wide CRISPR screens reveal host factors critical for SARS-CoV-2 infection. *Cells* 184, 76–91.e13. doi: 10.1016/j.cell.2020.10.028

Widera, M., Wilhelm, A., Hoehl, S., Pallas, C., Kohmer, N., Wolf, T., et al. (2021b). Limited neutralization of authentic severe acute respiratory syndrome coronavirus 2 variants carrying E484K in vitro. *J Infect Dis* 224, 1109–1114. doi: 10.1093/infdis/jiab355

Widera, M., Wilhelm, A., Toptan, T., Raffel, J. M., Kowarz, E., Roesmann, F., et al. (2021a). Generation of a sleeping beauty transposon-based cellular system for rapid and sensitive screening for compounds and cellular factors limiting SARS-CoV-2 replication. *Front. Microbiol.* 12:701198. doi: 10.3389/fmicb.2021.701198

Wilhelm, A., Toptan, T., Pallas, C., Wolf, T., Goetsch, U., Gottschalk, R., et al. (2021). Antibody-mediated neutralization of authentic SARS-CoV-2 B.1.617 variants harboring L452R and T478K/E484Q. *Viruses* 13:1693. doi: 10.3390/v13091693

Wilhelm, A., Widera, M., Grikscheit, K., Toptan, T., Schenk, B., Pallas, C., et al. (2022). Limited neutralisation of the SARS-CoV-2 omicron subvariants BA.1 and BA.2 by convalescent and vaccine serum and monoclonal antibodies. *EBioMedicine* 82:104158. doi: 10.1016/j.ebiom.2022.104158

Wishart, D. S., Feunang, Y. D., Guo, A. C., Lo, E. J., Marcu, A., Grant, J. R., et al. (2018). DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res.* 46, D1074–D1082. doi: 10.1093/nar/gkx1037

Wong, C. K. H., Au, I. C. H., Lau, K. T. K., Lau, E. H. Y., Cowling, B. J., and Leung, G. M. (2022). Real-world effectiveness of early molnupiravir or nirmatrelvir-ritonavir in hospitalised patients with COVID-19 without supplemental oxygen requirement on admission during Hong Kong's omicron BA.2 wave: a retrospective cohort study. *Lancet Infect. Dis.* 22, 1681–1693. doi: 10.1016/S1473-3099(22)00507-2

Xiao, N., Cao, D. S., Zhu, M. F., and Xu, Q. S. (2015). Protr/ProtrWeb: R package and web server for generating various numerical representation schemes of protein sequences. *Bioinformatics* 31, 1857–1859. doi: 10.1093/bioinformatics/btv042

Zaki, A. M., van Boheemen, S., Bestebroer, T. M., Osterhaus, A. D. M. E., and Fouchier, R. A. M. (2012). Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. *N. Engl. J. Med.* 367, 1814–1820. doi: 10.1056/NEJMoa1211721

Zhu, Y., Feng, F., Hu, G., Wang, Y., Yu, Y., Zhu, Y., et al. (2021). A genome-wide CRISPR screen identifies host factors that regulate SARS-CoV-2 entry. *Nat. Commun.* 12:961. doi: 10.1038/s41467-021-21213-4

Zou, J. N., Sun, L., Wang, B. R., Zou, Y., Xu, S., Ding, Y. J., et al. (2021). The characteristics and evolution of pulmonary fibrosis in COVID-19 patients as assessed by AI-assisted chest HRCT. *PLoS One* 16:e0248957. doi: 10.1371/journal.pone.0248957