


RESEARCH

Open Access

Genome-scale CRISPR screens identify host factors that promote human coronavirus infection



Marco Grodzki¹, Andrew P. Bluhm^{2,3}, Moritz Schaefer⁴, Abderrahmane Tagmount⁴, Max Russo^{4,5}, Amin Sobh⁶, Roya Rafiee⁷, Chris D. Vulpe⁴, Stephanie M. Karst^{1*}  and Michael H. Norris^{2,3*}

Abstract

Background: The COVID-19 pandemic has resulted in 275 million infections and 5.4 million deaths as of December 2021. While effective vaccines are being administered globally, there is still a great need for antiviral therapies as antigenically novel SARS-CoV-2 variants continue to emerge across the globe. Viruses require host factors at every step in their life cycle, representing a rich pool of candidate targets for antiviral drug design.

Methods: To identify host factors that promote SARS-CoV-2 infection with potential for broad-spectrum activity across the coronavirus family, we performed genome-scale CRISPR knockout screens in two cell lines (Vero E6 and HEK293T ectopically expressing ACE2) with SARS-CoV-2 and the common cold-causing human coronavirus OC43. Gene knockdown, CRISPR knockout, and small molecule testing in Vero, HEK293, and human small airway epithelial cells were used to verify our findings.

Results: While we identified multiple genes and functional pathways that have been previously reported to promote human coronavirus replication, we also identified a substantial number of novel genes and pathways. The website <https://sarscrisprscreens.epi.ufl.edu/> was created to allow visualization and comparison of SARS-CoV2 CRISPR screens in a uniformly analyzed way. Of note, host factors involved in cell cycle regulation were enriched in our screens as were several key components of the programmed mRNA decay pathway. The role of EDC4 and XRN1 in coronavirus replication in human small airway epithelial cells was verified. Finally, we identified novel candidate antiviral compounds targeting a number of factors revealed by our screens.

Conclusions: Overall, our studies substantiate and expand the growing body of literature focused on understanding key human coronavirus-host cell interactions and exploit that knowledge for rational antiviral drug development.

Background

The COVID-19 pandemic is arguably the most consequential infectious disease outbreak in modern times. The causative agent of COVID-19, SARS-CoV-2, spread quickly across the planet resulting in 245 million infections and nearly 5

million deaths at the time of this writing. Multiple COVID-19 vaccines recently demonstrated high efficacy, received FDA approval, and are being administered to people across the globe [1–4]. While the importance of this scientific achievement cannot be overstated, there is still a great need for novel antiviral therapies for use in vulnerable immunocompromised individuals, in regions where vaccine access is limited, and in the event antigenically distinct SARS-CoV-2 variants, such as delta and omicron, continue to arise and threaten vaccine efficacy [5]. Moreover, considering that SARS-CoV-2 is the third novel human coronavirus (HCoV)

* Correspondence: skarst@ufl.edu; mhnorris@ufl.edu

¹Department of Molecular Genetics and Microbiology, College of Medicine, University of Florida, Gainesville, FL, USA

²Department of Geography, College of Liberal Arts and Sciences, University of Florida, Gainesville, FL, USA

Full list of author information is available at the end of the article



© The Author(s). 2022 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

to emerge and cause serious disease in the human population in the past two decades following SARS-CoV and MERS-CoV, potent and broad-spectrum antivirals will leave us better prepared to deal with future pandemics. Broad-spectrum antivirals could also reduce morbidity associated with common cold-causing HCoVs including OC43, NL63, 229E, and HKU1.

Antivirals segregate into two basic categories, virus-targeting and host-targeting, both of which require an understanding of the molecular mechanisms used by viruses to replicate in host cells. Coronaviruses replicate via a well-established series of molecular events [6, 7]. Host factors are required at every step in this life cycle and represent candidate druggable targets (i.e., host-targeting antivirals) with the potential for broad-spectrum activity against multiple viruses within a given virus family and even across virus families [8, 9]. Accordingly, we performed CRISPR-based genome-wide knockout screens for both SARS-CoV-2 and OC43 infections to identify host factors that promote HCoV replication. Considering the power of genome-wide screens in the identification of host factors required for viral replication and the enormous global impact of the ongoing COVID-19 pandemic, it is not surprising that other research groups also applied this approach to SARS-CoV-2. Six genome-wide CRISPR screens for the identification of host factors promoting SARS-CoV-2 replication are published [10–15]. Despite the redundancy in the overall approach, there was experimental variability across screens in the selection of cell lines and infection conditions. Together with the sheer magnitude of critical host-virus interactions required for successful viral infection, individual screens are likely to capture only a subset of these host factors. Consistent with this, while specific genes and pathways were identified across published studies, each study also provided unique findings which expand our understanding of host-HCoV interactions.

In this study, we report a global analysis of host-HCoV interactions gleaned from genome-wide screens performed for two HCoVs and in two different cell lines. We also performed a comprehensive comparative analysis of all published genome-wide SARS-CoV-2 screens to date [11, 12, 14, 15]. Multiple genes and functional pathways identified in our screens were previously reported to promote SARS-CoV-2 replication, validating the rigor of our approach and providing further support for the role of specific host factors. Yet we also identified a substantial number of novel genes and pathways not previously reported to promote HCoV replication. We validated the importance of a subset of genes identified in these screens in HCoV replication in multiple cell types, including a respiratory cell line. Notably, several of the novel host factors identified in our study provide unique insight into SARS-CoV-2 replication processes

that could be targeted with antiviral drugs. Host factors involved in cell cycle regulation were enriched in our screens and we show that compounds (abemaciclib, AZ1 protease inhibitor, harmine, nintedanib, and UC2288) targeting these host factors inhibit in vitro HCoV replication. We also identified multiple host factors involved in endocytosis and TBK1 that plays a key role in innate immune responses. Inhibitors of these processes/factors (promethazine and amlexanox, respectively) also displayed antiviral activity. Together, our study has provided significant insight into host-HCoV interactions and identified novel candidate antiviral compounds.

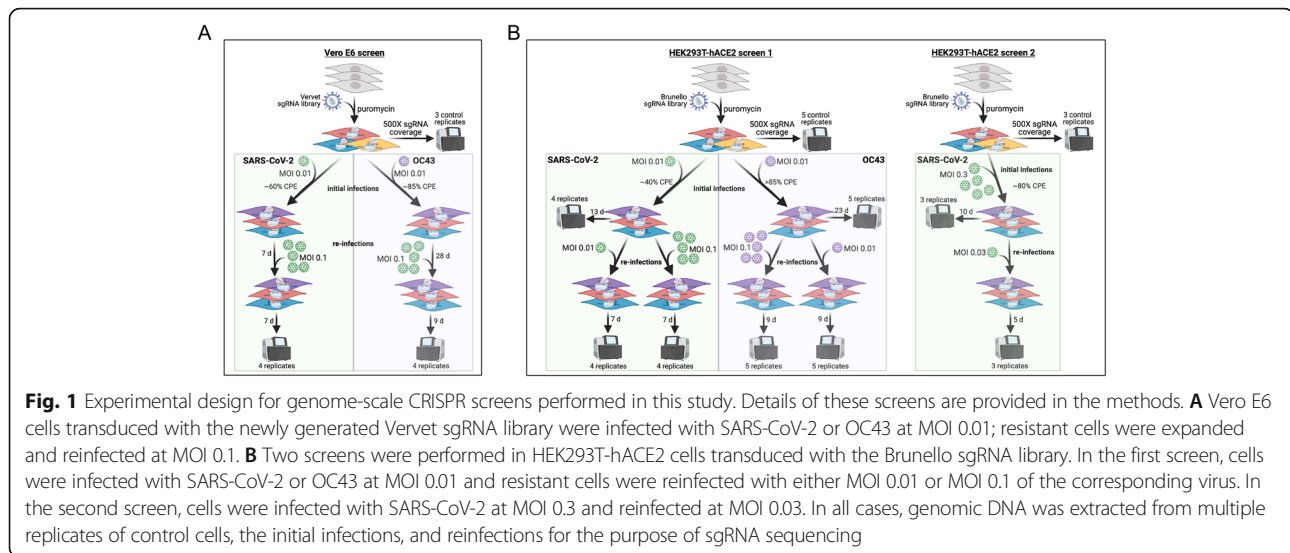
Methods

Study design

The objectives of this study were to identify host factors that promote HCoV replication and to determine whether these host factors can be targeted with commercially available drugs to block viral infection in vitro (Additional file 1: Fig. S1). To achieve this goal, we performed genome-wide CRISPR knockout screens in Vero E6 cells using a newly generated and validated Vervet sgRNA library (Additional file 1: Fig. S2), as described in the “Generation of the Vervet-specific genome-wide sgRNA library” section of Additional file 1: Supplementary Methods (Fig. 1A) and in HEK293T-hACE2 cells using the commercially available human Brunello sgRNA library, as described in the “Genome-wide CRISPR sgRNA screens” section of the “Methods” section (Fig. 1B). In brief, cells transduced with sgRNA libraries were infected with HCoVs, SARS-CoV-2 and OC43, using various MOIs. Cells surviving infection were expanded and reinfected, and the sgRNAs enriched in resistant clones were determined. Genes targeted by enriched sgRNAs were compared between replicates, infection conditions, HCoVs, cell lines, and previously published screens [11, 12, 14, 15] to identify common and unique host factors as well as putative pan-HCoV host factors. Genes of interest were selected for validation and targeted for knockdown using shRNAs followed by infection with HCoVs. Commercially available inhibitors to other important genes identified in our study were evaluated for their capacity to prevent virus-induced toxicity and viral replication in vitro. EC₅₀ and IC₅₀ values for efficacious compounds were determined.

Virus stock generation, viral titer determination, and cell lines

Vero E6 cells were obtained from ATCC. HEK293T cells expressing hACE2 receptor were obtained from GeneCopoeia. hTERT immortalized human small airway epithelial cells (SAEC) were obtained from ATCC. The human ACE2 gene was transduced into SAEC to generate stably expressing SAEC-hACE2 cells. SARS-CoV-2



strain UF-1 (GenBank accession number MT295464.1) was originally isolated from a COVID-19 patient at the University of Florida Health Shands Hospital via nasal swab [16] and manipulated in a Biosafety Level 3 (BSL3) laboratory at the Emerging Pathogens Institute with appropriate approvals obtained under the University of Florida Institutional Biosafety Committee protocol# BIO5594. The HCoV OC43 strain was a kind gift from Dr. John Lednicky (University of Florida). SARS-CoV-2 and OC43 were propagated in Vero E6 cells (grown in Dulbecco's modified Eagle medium (DMEM; Gibco) supplemented with 10% heat-inactivated fetal bovine serum (FBS; Atlanta Biologicals) and Pen-Strep (100 U/ml penicillin, 100 µg/ml streptomycin; Gibco) at 37 °C and 5% CO₂. Virus stocks were prepared by infecting Vero E6 cells at MOI 0.01, centrifuging culture supernatants collected at 3 dpi for 5 min at 1000 × g, and filtering through a 0.44-µm PVDF filter (Millipore) followed by a 0.22-µm PVDF filter (Restek). The virus stocks were aliquoted and stored at -80 °C. Virus stocks or supernatants from infected SAEC-hACE2 cells were titered using a standard TCID₅₀ assay. In brief, Vero E6 cells were seeded at 2 × 10⁴ cells per well in a 96-well plate (Corning) and allowed to attach overnight. Virus stocks or supernatants from infected cells were serially diluted onto cells, with a total of 8 replicates per dilution. Monolayers were visualized in the BSL3 using an EVOS XL Core microscope (Thermo Fisher Scientific) and scored positive or negative for cytopathic effect (CPE) at 7 dpi.

Viral genome copy number enumeration

For SARS-CoV-2, supernatants and cells were harvested into AVL buffer from the QIAamp Viral RNA Kit (Qiagen) and RNA was purified according to the

manufacturer's recommendations. The samples underwent reverse transcription and cDNA synthesis using the iTaq Universal SYBR Green One-Step Kit (Bio-Rad) and primers targeting the nucleocapsid (N) gene of SARS-CoV-2 (NproteinF- GCCTCTTCTCGTTCCTCATCAC, NproteinR-AGCAGCATCACCGCCATTG). qPCR was carried out on a Bio-Rad CFX96, and viral genome copy numbers were extrapolated using CT values from a standard curve generated using a control plasmid containing the N protein gene (Integrated DNA Technologies). For OC43, RNA from infected cells was purified using the RNeasy Mini Kit (Qiagen) according to the manufacturer's recommendations and amplified using the Applied Biosystems AgPath-ID One-Step RT-PCR Kit (Thermo Fisher Scientific) and primers and probe targeting the N gene of OC43 [17]. GAPDH levels were determined for each sample for normalization purposes using previously described primers [18]. All samples were run in triplicate for each primer pair and normalized viral genome copy numbers were calculated using the comparative cycle threshold method.

Genome-wide CRISPR sgRNA screens

The human CRISPR Brunello library (Addgene 73178) [19] was amplified following a previously published protocol [20]. We constructed a Vervet domain-targeted sgRNA library since one was not commercially available. Detailed methods are reported in Additional file 1: Supplementary methods. For both libraries, lentiviruses were produced in HEK293T cells by co-transfection of library plasmids together with the packaging plasmid psPAX2 (Addgene 12260) and envelope plasmid pMD2.G (Addgene 12259). CRISPR screens were carried out in two cell lines (outlined in Fig. 1): AGM Vero E6 cells were transduced with the newly generated Vervet

sgRNA library (sgRNAs are listed in Additional file 2) and human HEK293T-hACE2 cells (Genecopoeia) were transduced with the Brunello sgRNA library. For each screen, 1.2×10^8 cells were transduced with lentivirus-packaged sgRNA library at MOI 0.3 in the presence of 8 $\mu\text{g/ml}$ polybrene (Sigma) to achieve ~ 500 -fold overrepresentation of each sgRNA. After 48 h, 0.6 $\mu\text{g/ml}$ puromycin (Gibco) was added to eliminate non-transduced cells and cultures were expanded in Matrigel-coated (Corning) T300 flasks. Control replicates were collected at this time to determine input library composition and additional replicates were infected with SARS-CoV-2 or OC43 at the indicated MOIs. Cells surviving initial infections were collected when they had expanded to confluency. A portion of each replicate was stored in DNA/RNA Shield (Zymo Research) at -80°C for genomic DNA extraction, and the remaining cells were reseeded and reinfected at the indicated MOIs. Cells surviving reinfections were also harvested at confluency for genomic DNA extraction.

Genomic DNA was extracted from each sample (detailed extraction methods are described in Additional file 1: Supplementary methods and Table S1). The sgRNA regions were then amplified and indexed for Illumina sequencing using a one-step PCR method and primers specific to the LentiCRISPRv2-based Vervet and Brunello libraries. Primers and indices used for the generation of amplicon libraries are listed in Additional file 1: Table S2. Brunello and Vervet DNA samples were amplified in ten 100- μl reactions using the NEBNext High-Fidelity 2X Master Mix Kit (New England Biolabs), 0.5 μM of forward and reverse primers, and 10 μg of DNA template per reaction with the following program: initial denaturation at 98°C for 3 min, 24 cycles of denaturation at 98°C for 10 s, annealing 60°C for 15 s and extension 72°C for 25 s, and final extension at 72°C for 2 min. Then, 256-bp amplicons were quantified on a 2% agarose gel stained with SYBR Safe (Invitrogen), using the Gel Doc quantification software (Bio-Rad). Amplicons were first pooled in an equimolar fashion and then the pools were gel-extracted using the PureLink Quick Gel Extraction Kit (Thermo Fisher Scientific). The sequencing was carried out at the Interdisciplinary Center for Biotechnology Research (ICBR; University of Florida) using a NovaSeq 6000 sequencer (Illumina). The Brunello amplicons were sequenced using the S4 2X150 cycles Kit (Illumina) while the Vervet-AGM amplicons were sequenced with the SP 1X100 cycles Kit (Illumina).

Computational analysis

A custom script was developed for implementation of the data analysis pipeline and is available at GitHub <https://github.com/moritzschaefer/covid19-screens> [21]. Briefly, the FASTX-Toolkit [22] was used to demultiplex

raw FASTQ data which were further processed to generate reads containing only the unique 20-bp sgRNA sequences. The sgRNA sequences from the library were assembled into a Burrows-Wheeler index using the Bowtie build-index function and reads were aligned to the index. The efficiency of alignment was checked and the number of uniquely aligned reads for each library sequence was calculated to create a table of raw counts. Ranking of genes corresponding to perturbations that are enriched in infected cultures was performed using a robust ranking aggregation (a-RRA) algorithm implemented in the Model-based Analysis of Genome-wide CRISPR/Cas9 Knockout (MAGeCK) tool through the test module [23]. Tables with raw counts corresponding to each sgRNA in reference (initial pool) and selected (virus-infected) samples were used as an input for the MAGeCK test. Gene-level ranking was based on FDRs and candidates with FDRs < 0.25 were considered as significant hits. Ranking of genes corresponding to positively selected and negatively selected perturbations was performed using a robust ranking aggregation (a-RRA) algorithm implemented in MAGeCK through the test module [23]. Tables with raw counts corresponding to each sgRNA in reference (initial pool) and selected (exposed to virus) samples were used as an input for the MAGeCK test. Gene-level ranking was based on false discovery rate (FDR) and candidates with FDR < 0.25 were considered as significant hits. Additional details can be found in Additional file 1: Supplementary methods. We submitted FastQ files for all CRISPR sequencing to Gene expression omnibus (GSE177545-<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE177545>) [24] and count files for all samples are available in Additional file 3.

Validation of host factors in promoting HCoV infections

To validate selected host factors for their capacity to promote HCoV infection in vitro, we transduced HEK293T-hACE2 cells with lentivirus-packaged shRNAs targeting *CTSL*, *CCZ1*, or *EDC4* (TRC Human shRNA Library collection [25]) or the empty vector pLKO.1. In addition, we generated hACE-expressing SAEC cells by transduction with pLENTI_hACE2_HygR plasmid-packaged lentiviruses (Addgene 155296) and selection of transduced cells with hygromycin (300 $\mu\text{g/ml}$). Next, SAEC-hACE2 cells were transduced with lentivirus-packaged sgRNAs targeting *EDC4* and *XRNI*, designed with PAVOOC [26] and synthesized by the Molecular Cloning Facility (Department of Biological Science, Florida State University). Transduced cells were expanded under puromycin selection (0.8 $\mu\text{g/ml}$) for at least 7 days to generate stable knockdown or knockout cell lines. For knockout SAEC-hACE2 cells, monoclonal colonies were isolated from pools of cells and tested individually. To

confirm gene knockdown and knockout, cell lysates prepared from knockdown/knockout and control cells were tested by western blotting with antibodies directed to CTSL (Invitrogen, BMS1032), CCZ1 (Santa Cruz Biotechnology, sc-514290), EDC4 (Cell Signaling Technology, 2548S), XRN1 (Cell Signaling Technology, 70205S), and actin as a loading control (Sigma-Aldrich, MAB1501R). Once knockdown or knockout was confirmed, cells were infected with SARS-CoV-2 or OC43 at MOI 0.01 and RNA was extracted at 2 dpi for viral genome copy number enumeration, or cell supernatants were tested at various timepoints for quantification of viral infectious particles by TCID₅₀ as described above. Cell viability of SAEC-hACE2 knockout cells was compared to control cells using the CellTiterGlo 2.0 Cell Viability Assay (Promega) according to the manufacturer's instructions. Knockout cells displayed no gross reduction in viability.

Identification and testing inhibitors of host factors from CRISPR screens

Online databases [27–31] and published literature (<https://pubmed.ncbi.nlm.nih.gov/>) were used to find small molecule inhibitors targeting a subset of top-scoring genes in the CRISPR screens. Amlexanox (AMX) was purchased from InvivoGen. Abemaciclib (ABE), AZ1, olaparib (OPB), and nintedanib (NIN) were purchased from Selleckchem. Harmine (HAR), INDY, chlorpromazine (CPZ), promethazine (PMZ), UC2288 (UC2), and CID 1067700 (CID) were purchased from Millipore Sigma. Drugs were diluted according to the manufacturers' recommendations and single-use aliquots were frozen at –80 °C until the time of assay. Drugs were diluted down to 2X concentrations and mixed with 2X concentrations of virus to generate 1X concentrations, then added to the monolayers. Cells were infected at a MOI of 0.2 as determined by preliminary experiments to generate an ideal dynamic range of the colorimetric CytoTox 96° Non-Radioactive Cytotoxicity Assay (Promega). Infections progressed for 72 h at which time supernatant from treatments and controls were processed for LDH release according to the manufacturer's recommendations. Absorbance at 450 nm was read using an accuSkan FC microplate reader (Fisher Scientific) with SkanIt software (Fisher Scientific). Absorbance values were background subtracted and transformed to percent of virus-infected controls. These percentages were compared to the values obtained from virus-infected cell cytotoxicity values by one-way ANOVA using GraphPad Prism version 9. Assays were carried out in biological duplicate and in three independent experiments. The concentration of drug alone that resulted in 50% of maximum toxicity (cytotoxic concentration 50; CC₅₀) and the concentration of drug that inhibited

50% of the vehicle-treated SARS-CoV-2-induced cytotoxicity (effective concentration 50; EC₅₀) were determined by serially diluting small molecule inhibitors during SARS-CoV-2 infection of Vero E6 cells. EC₅₀ and CC₅₀ values were calculated by transforming inhibitor concentrations to log then using the non-linear fit with variable slope function (GraphPad Prism version 9) to determine best fit variables using the percent of maximum SARS-CoV-2-induced cytotoxicity measurements at each drug concentration performed in technical duplicate. Candidate inhibitors were assayed for their ability to prevent OC43 and SARS-CoV-2 replication in SAEC-hACE2 cells by TCID₅₀. SAEC-hACE2 cells were infected in triplicate at an MOI of 0.01 in the presence of the indicated inhibitors for 3 days. TCID₅₀ were determined by diluting the supernatant of each replicate across octuplet columns of Vero E6 cells. Five days later, TCID₅₀ was read. The data are from two independent repetitions. Phospholipidosis was determined in SAEC-hACE2 cells with the HCS LipidTOX Red Neutral Lipid Stain according to the manufacturer's instructions using a Biotek Cytation 3 plate imager. Cytotoxicity in SAEC-hACE2 cells was determined as above for Vero E6 cells.

Website creation and data repository

To facilitate the access, reusability, and integration of this data, we have created and hosted a website (sarscrisprscreens.epi.ufl.edu) which contains data for previously published HCoV CRISPR screens and our integrated MAGeCK analysis using the VISPR pipeline of the SARS-CoV-2 screens [11, 12, 14, 15]. VISPR is a Web-based interactive tool to visualize CRISPR screening experiments [32]. Our goal is to provide a community resource for facile integrated analysis of current and future CRISPR screens. Further details on how to submit new data are provided on the website.

Statistical analysis

For the genome-scale CRISPR analysis, the embedded statistical tools in the MAGeCK/VISPR pipelines were used [23, 33]. Further details are provided in Additional file 1: Supplementary Methods. All other statistical analyses were carried out using GraphPad Prism 9.0. To compare the mean normalized viral genome copy number values in targeted shRNA knockdown experiments (Fig. 6), *P* values were determined using one-way ANOVA test (**P* < 0.05, ***P* < 0.01, ****P* < 0.001), with error bars representing standard errors of mean (*n* = 3 experiments). For testing inhibitory activity of small molecules on SARS-CoV-2 infection of Vero E6 cells (Fig. 7), a one-way ANOVA test was used for comparison of toxicity values for inhibitor-treated infected cells and infected-only control cells (no treatment), with error bars denoting standard deviation for all panels (*n* = 3

experiments). Non-linear regression of data points was used to determine the EC₅₀ and IC₅₀ values for indicated compounds.

Results

Genome-wide CRISPR screens in Vero E6 cells identify host factors required for HCoV infection

In order to identify host factors that promote HCoV infection, we performed genome-wide loss of function CRISPR screens for pathogenic SARS-CoV-2 and common cold-causing OC43 in two susceptible cell lines. Due to the highly cytopathic nature of HCoV infection in the Vero E6 cells derived from African Green Monkey (AGM; *Chlorocebus sabaeus*), we carried out genome-wide screens using a custom Vervet CRISPR knockout library (see supplemental methods) (Fig. 1A). Vero E6 cells were transduced with the Vervet CRISPR library and infected with SARS-CoV-2 or OC43 at a multiplicity of infection (MOI) 0.01. We observed ~60% visible cytopathic effect (CPE) for SARS-CoV-2 and ~85% CPE for OC43.

Resistant clones were expanded, reinfected with the corresponding virus at MOI 0.1, and re-expanded. Genomic DNA was extracted from surviving cells, sgRNAs amplified, and sequenced. We carried out MAGeCK analysis to identify genes targeted by significantly enriched sgRNAs which are labeled in the volcano plots in Fig. 2A, B. Full data sets are available in Additional files 4 and 5. To facilitate the access and reusability of data sets generated from genome-scale CRISPR screens in HCoV-infected cells, we have also hosted a website (sarscrisprscreens.epi.ufl.edu) with the complete set of MAGeCK results for each of the screens described in this study and data from previously published screens reanalyzed herein [8, 12–15]. The website was designed to facilitate integration of upcoming screens and we hope for contributions to drive this as a community project.

We identified multiple candidate host factors previously demonstrated to play a functional role in SARS-CoV-2 and OC43 infections. For example, *ACE2* was identified in the SARS-CoV-2 screen [34]. Furthermore,

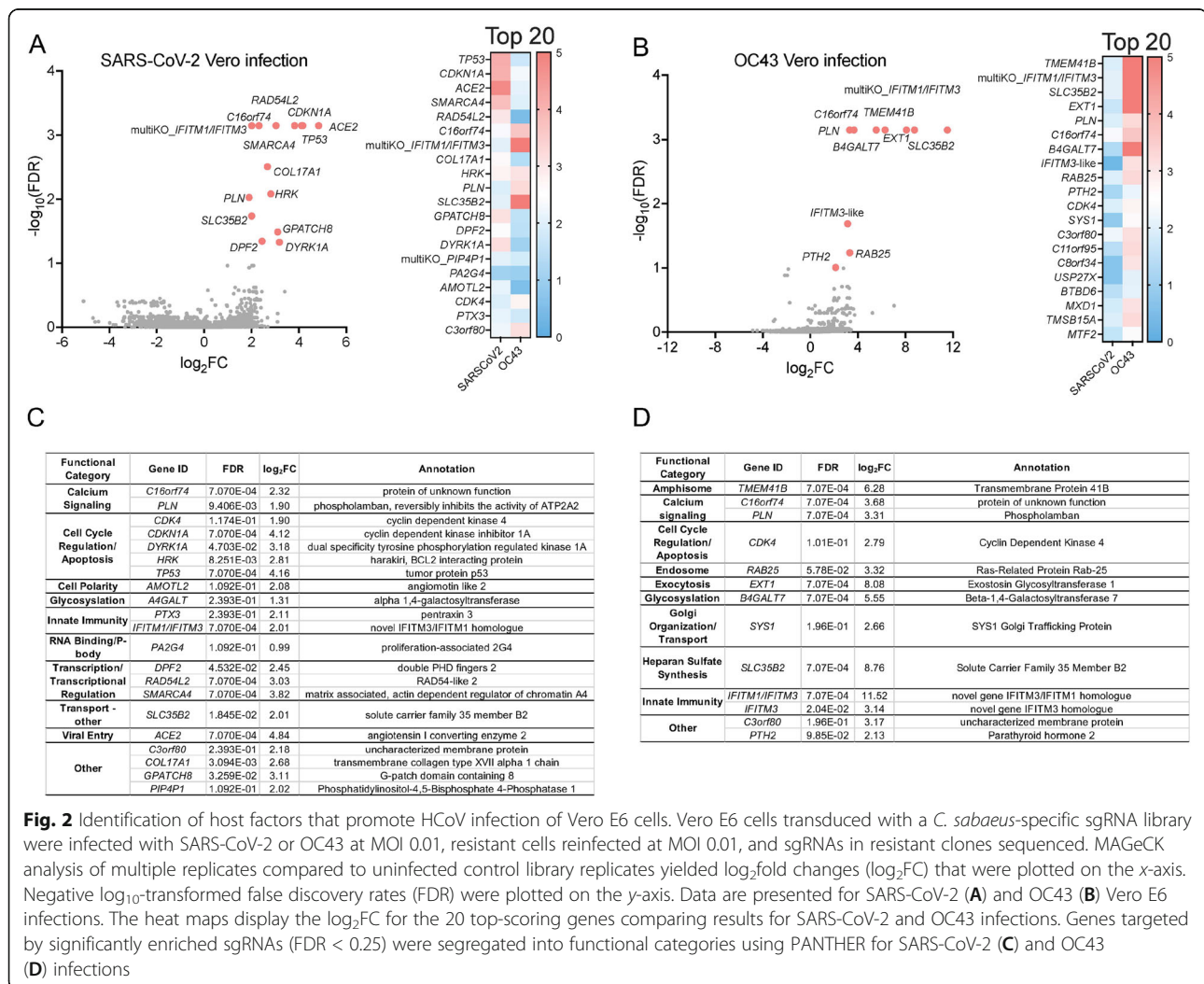


Fig. 2 Identification of host factors that promote HCoV infection of Vero E6 cells. Vero E6 cells transduced with a *C. sabaeus*-specific sgRNA library were infected with SARS-CoV-2 or OC43 at MOI 0.01, resistant cells reinfected at MOI 0.01, and sgRNAs in resistant clones sequenced. MAGeCK analysis of multiple replicates compared to uninfected control library replicates yielded log₂fold changes (log₂FC) that were plotted on the x-axis. Negative log₁₀-transformed false discovery rates (FDR) were plotted on the y-axis. Data are presented for SARS-CoV-2 (A) and OC43 (B) Vero E6 infections. The heat maps display the log₂FC for the 20 top-scoring genes comparing results for SARS-CoV-2 and OC43 infections. Genes targeted by significantly enriched sgRNAs (FDR < 0.25) were segregated into functional categories using PANTHER for SARS-CoV-2 (C) and OC43 (D) infections

TMEM41B was a top-scoring gene in the OC43 screen, supporting recent work by Schneider et al. demonstrating that *TMEM41B* is a pan-HCoV host factor [12]. We also identified interferon (IFN)-induced transmembrane (IFITM) proteins that have been reported to regulate HCoV infection [35–37]. Genes targeted by significantly enriched sgRNAs were next segregated into functional categories listed in tables in Fig. 2C, D. Of note, *CDK4*, a master regulator of the cell cycle, was identified as a key host factor for both viruses. Disruption of additional genes encoding regulators of cell cycle progression, including *CDK1NA*, *DYRK1A*, *HRK*, and *P53*, similarly increased cellular resistance to SARS-CoV-2 infection.

During the completion of our studies, Wei et al. reported similar SARS-CoV-2 screens in Vero E6 cells performed with an independent sgRNA library based on an earlier *C. sabaues* genome assembly [13]. In order to compare our data sets to those of Wei et al., we downloaded raw data from their study and analyzed them using MAGeCK-VISPR [13, 33] (Additional file 1: Supplementary methods and Additional file 3). There were 6 targeted genes identified in common between studies: *ACE2*, *DPF2*, *DYRK1A*, *RAD54L2*, *SMARCA4*, and *TP53*.

Genome-wide CRISPR screens in HEK293T-hACE2 cells identify host factors required for HCoV infection

We similarly performed CRISPR screens in human HEK293T cells ectopically expressing the human ACE2 receptor (HEK293T-hACE2) transduced with the Brunello sgRNA library [19] (Fig. 1B). Transduced cells were infected with SARS-CoV-2 or OC43 at MOI 0.01. SARS-CoV-2-infected cultures developed ~40% CPE and OC43-infected cultures developed >85% CPE. Resistant cell populations propagated to confluence were reinfected with the corresponding virus at either MOI 0.01 or MOI 0.1 and re-expanded. Genomic DNA was extracted and sgRNAs from both the initial and secondary infections were sequenced.

The genes targeted by the most highly enriched sgRNAs in each of the SARS-CoV-2 infections are indicated in Fig. 3A–C. The full data set is available in Additional file 6. *EDC4*, a gene encoding a scaffold protein that functions in programmed mRNA decay, was the overall top-scoring gene. Interestingly, *XRN1* encodes another key player in this pathway and was also top-scoring. We further categorized the genes encoding candidate host factors (FDR < 0.1) into functional categories depicted in heat maps in Fig. 3D. Consistent with other published screens, we identified multiple components of the endocytic pathway including *CCZ1*, *DNM2*, and *WASL*. Other functional categories in which multiple genes were identified include cell adhesion, cell cycle, integrator complex, lysosome, mTOR regulation, and

ubiquitination/proteolysis. We carried out an independent SARS-CoV-2 screen using the higher MOI of 0.3 for initial infection which resulted in ~80% CPE, and MOI 0.03 for secondary infection. Genes targeted by the most significantly enriched sgRNAs in this study are presented in Fig. 3E, F and are segregated into functional categories depicted in heat maps in Fig. 3G. The full data set is available in Additional file 7. *ACE2* was a top-scoring gene in this screen. Functional categories with multiple targeted genes include amphisome, autophagy, endosome, exocytosis, lysosome, peroxisome, transcription/transcriptional regulation, and ion transporters. *C18orf8*, *CCZ1*, *CDH2*, and *TMEM251* were identified in both the low- and high-MOI SARS-CoV-2 screens.

The genes targeted by the most highly enriched sgRNAs in the OC43 HEK293T-hACE2 screens are indicated in Fig. 4A–C and segregated into functional categories in Fig. 4D. The full data set is available in Additional file 8. As expected, based on prior work, genes encoding IFITM proteins were identified as proviral factors for OC43 [35]. *TMEM41B* was a top-scoring gene along with the functionally related *VMP1*, as were *CCZ1*, *CCZ1B*, *SLC35B2*, and *WDR81* which have all been reported in other recent OC43 genome-wide screens [11, 12]. When comparing the SARS-CoV-2 and OC43 HEK293T-hACE2 datasets, there were 6 genes in common targeted by significantly enriched sgRNAs (*C18orf8*, *CCZ1*, *CCZ1B*, *RAB7A*, *WDR81*, and *WDR91*). Notably, all of the corresponding gene products function in vesicle-mediated transport.

During the completion of our studies, similar SARS-CoV-2 screens in human Huh-7.5 [11, 12, 15] and A549 [10, 14] cells were published. In order to compare our data sets to those of other groups, we downloaded raw data from four published studies [11, 12, 14, 15], analyzed them using a common analysis framework (MAGeCK) and stringency (FDR < 0.25), and compared the results to our data sets (Additional files 9 and 10). Using this stringency, no genes were identified in all five studies, 1 gene was identified in four studies (*ACE2*), 6 genes were identified in three studies (*VPS35*, *CTSL*, *DNM2*, *CCZ1B*, *TMEM106B*, and *VAC14*), and 25 genes were identified in two studies (*ALG5*, *ARVCF*, *ATP6V1A*, *ATP6V1G1*, *B3GAT3*, *CNOT4*, *EPT1*, *EXOC2*, *EXT1*, *EXTL3*, *GDI2*, *LUC7L2*, *MBTPS2*, *PIK3C3*, *RAB7A*, *RNH1*, *SCAF4*, *SCAP*, *SLC30A1*, *SLC33A1*, *SNX27*, *TMEM41B*, *TMEM251*, *WDR81*, and *WDR91*) (Fig. 5A, B). It should be noted that these genes were top-scoring across studies performed in different human cell lines, suggesting they are broadly important in SARS-CoV-2 replication. Shared pathways include vesicle-mediated transport (*CCZ1B*, *DNM2*, *EXOC2*, *GDI2*, *PIK3C3*, *RAB7A*, *SNX27*, *VAC14*, *VPS35*, *WDR81*, *WDR91*), vacuolar ATPases important in

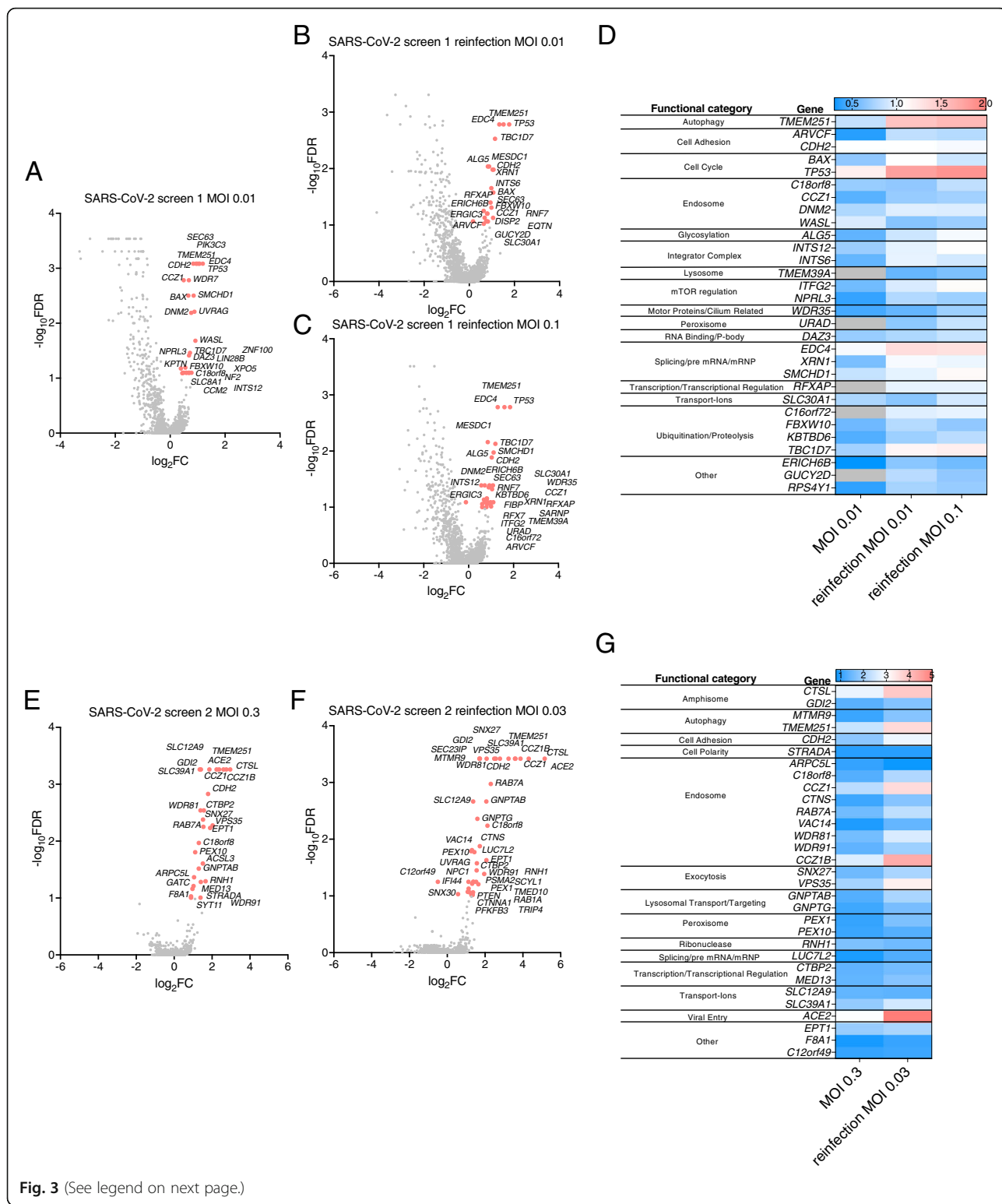


Fig. 3 (See legend on next page.)

(See figure on previous page.)

Fig. 3 Identification of host factors that promote SARS-CoV-2 infection of HEK293T-hACE2 cells. HEK293T-hACE2 cells transduced with the Brunello sgRNA library were infected with SARS-CoV-2 at MOI 0.01 and sgRNAs in resistant clones sequenced. Resistant clones were reinfected with SARS-CoV-2 at MOI 0.01 or MOI 0.1 and sgRNAs in resistant clones sequenced. For all three infections, MAGECK analysis of multiple replicates compared to uninfected control library replicates yielded log₂fold changes (log₂FC) that were plotted on the x-axis. Negative log₁₀-transformed FDR were plotted on the y-axis. Data are presented for the initial infection (A), MOI 0.01 reinfection (B), and MOI 0.1 reinfection (C). D The heat map displays the log₂FC for top-scoring genes (FDR < 0.1) across the three infections. The entire experiment was repeated at MOI 0.3, with sgRNAs sequenced from resistant clones in the initial infection (E) and reinfection (F). G The heat map displays the log₂FC for top-scoring genes (FDR < 0.25) across the two infections

organelle acidification (*ATP6V1A*, *ATP6V1E*, *ATP6V1G1*), and heparan sulfate biosynthesis genes (*EXT1*, *EXTL3*, *B3GAT3*). We identified 53 genes targeted by enriched sgRNAs in our study that were not identified in published studies (Fig. 5C), including *EDC4* and *XRNI*.

Validation of a subset of gene candidates that promote HCoV replication

To confirm that unique genes identified in our screens promote HCoV replication, HEK293T-hACE2 cells were engineered to stably express gene-specific shRNAs targeting *CCZ1* or *EDC4*. *CTSL* knockdown was tested as a

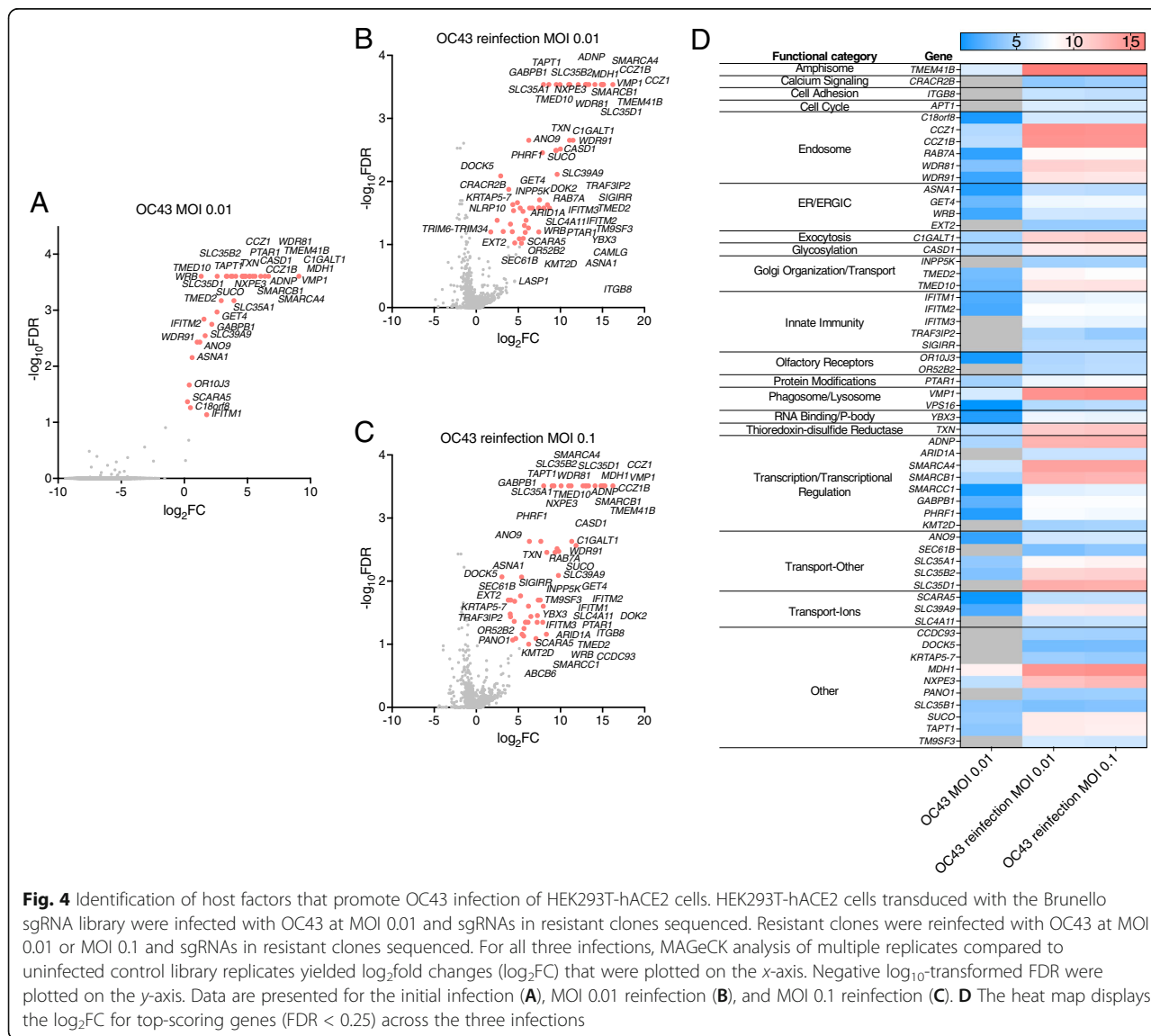
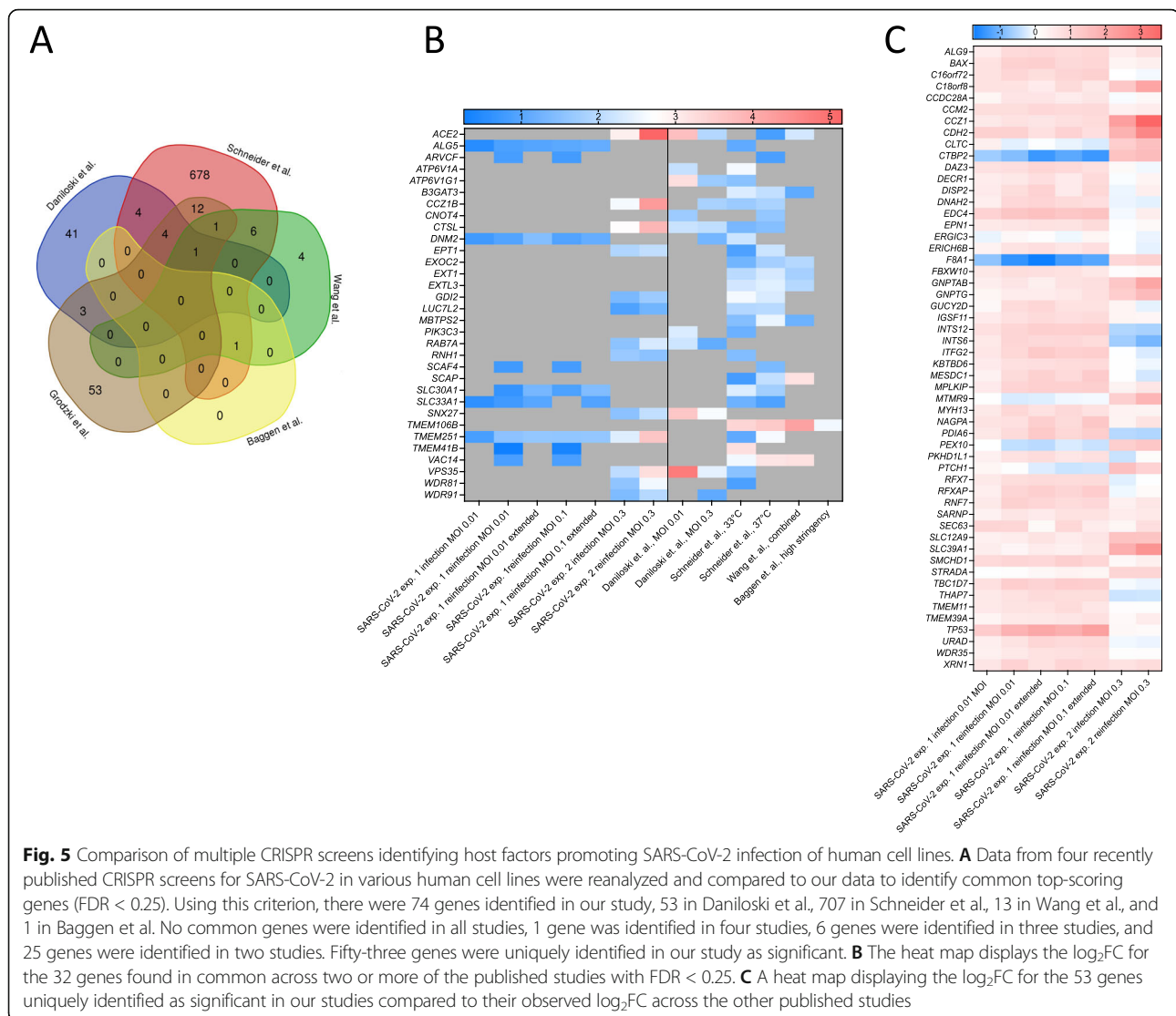


Fig. 4 Identification of host factors that promote OC43 infection of HEK293T-hACE2 cells. HEK293T-hACE2 cells transduced with the Brunello sgRNA library were infected with OC43 at MOI 0.01 and sgRNAs in resistant clones sequenced. Resistant clones were reinfected with OC43 at MOI 0.01 or MOI 0.1 and sgRNAs in resistant clones sequenced. For all three infections, MAGECK analysis of multiple replicates compared to uninfected control library replicates yielded log₂fold changes (log₂FC) that were plotted on the x-axis. Negative log₁₀-transformed FDR were plotted on the y-axis. Data are presented for the initial infection (A), MOI 0.01 reinfection (B), and MOI 0.1 reinfection (C). D The heat map displays the log₂FC for top-scoring genes (FDR < 0.25) across the three infections

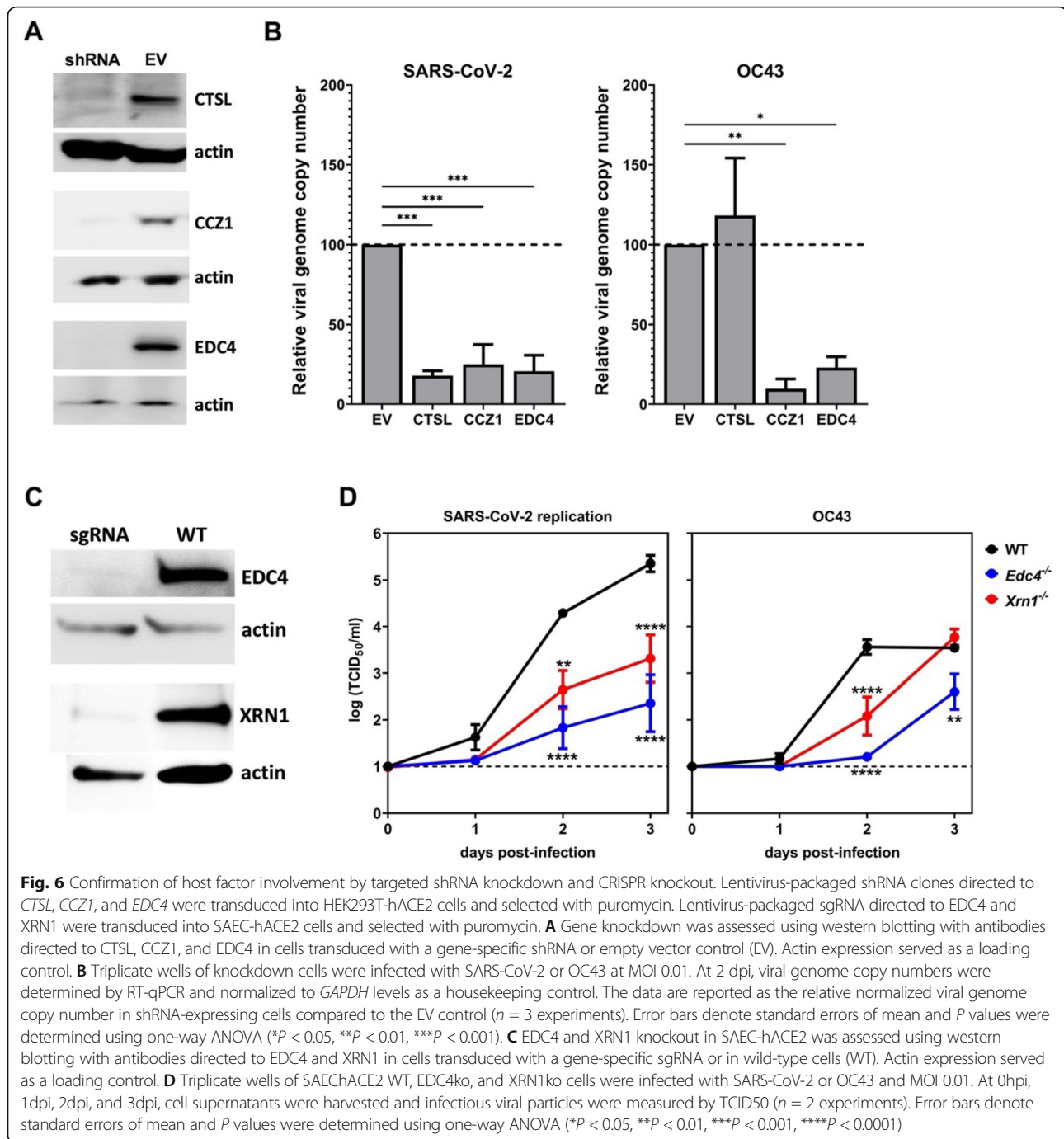


positive control for SARS-CoV-2 [38]. Efficiency of gene knockdown assessed by western blotting was robust for all three genes (Fig. 6A and Additional file 1: Fig. S3A). Knockdown cells were then infected with SARS-CoV-2 or OC43 and viral genome copy number determined at 2 days post-infection (dpi). All three genes were required for optimal SARS-CoV-2 infection while *CCZ1* and *EDC4*, but not *CTSL*, promoted OC43 infection (Fig. 6B). Because *EDC4* was unique to our screens, we next tested whether it plays a role in promoting HCoV replication in respiratory cells. The small airway epithelial cell (SAEC) line was engineered to express human ACE2 (hACE2) and then *EDC4* was targeted for knockout using a CRISPR/Cas9-based approach. As mentioned above, *Xrn1* functions in the same mRNA decay pathway as *Edc4* and was also identified in our screens so we engineered *Xrn1*^{-/-} SAEC-hACE2 expressing cells in parallel. After verifying gene knockout (Fig. 6C and

Additional file 1: Fig. S3B), cells were infected with either SARS-CoV-2 or OC43 and virus titers measured at various time points by TCID50 assay. *EDC4* and *XRN1* were necessary for efficient replication of both viruses (Fig. 6D).

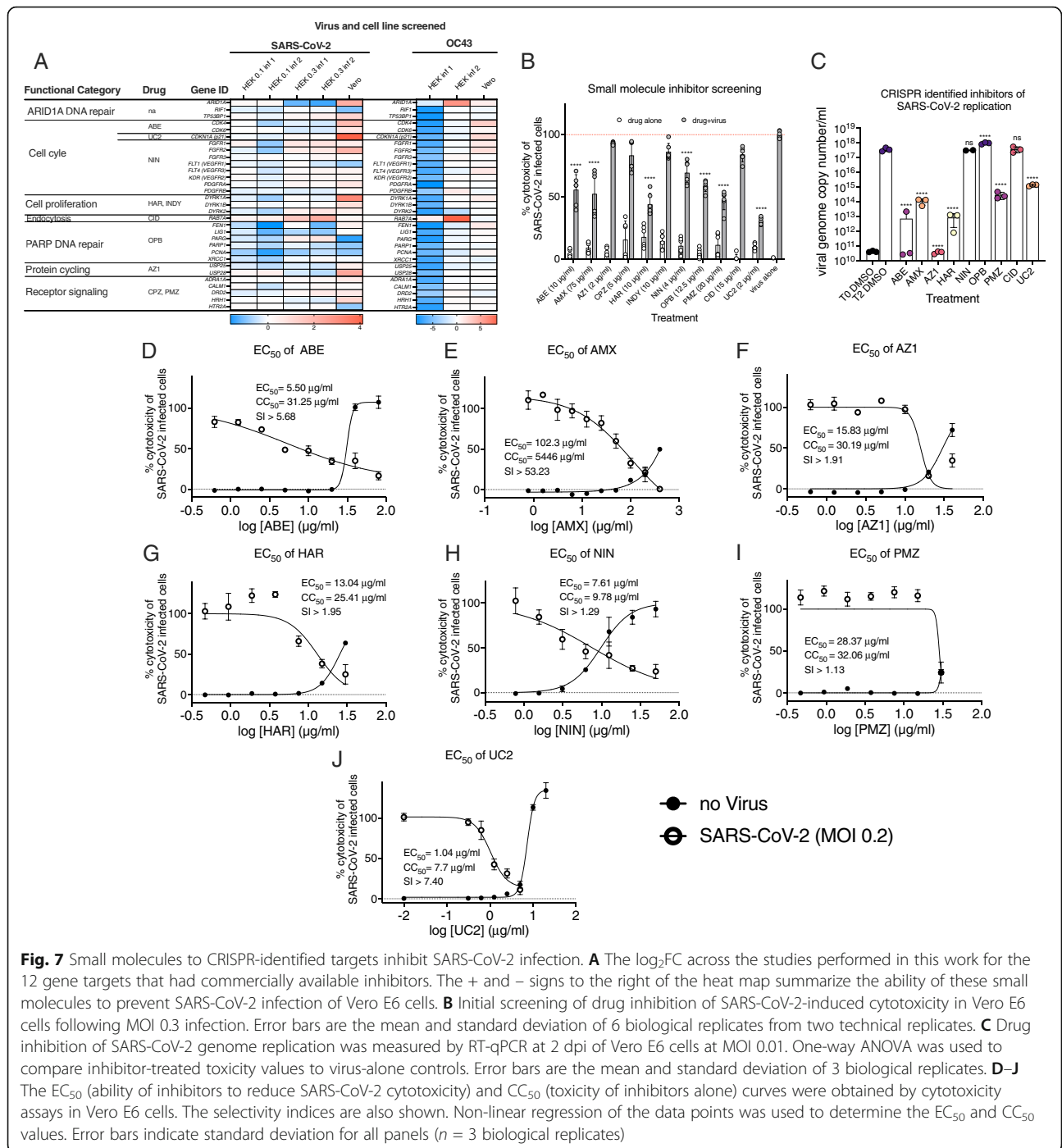
CRISPR screening reveals novel antiviral drugs displaying in vitro efficacy

We next determined whether gene products and pathways identified in our screens could be targeted with commercially available inhibitors to block HCoV infection. Numerous genes involved in cell cycle regulation were identified in our screens. The following inhibitors targeting this class of host factors were tested: abemaciclib (ABE; Cdk4, Cdk6 inhibitor), UC2288 (UC2; CDKN1A/p21 inhibitor), harmine (HAR) and INDY (Dyrk1A/B inhibitors), AZ1 (Usp25/28 inhibitor), olaparib (OPB; ARID1A inhibitors were not available so



inhibition of PARP-mediated DNA repair was investigated to see if DNA damage repair was involved in SARS-CoV-2 replication), and nintedanib (NIN; FGFR1/2/3, VEGFR1/2/3, and PEGFR α/β inhibitor). Host factors involved in endocytosis have been widely reported to regulate HCoV replication and were identified in our and others' CRISPR screens [39] so we also tested several drugs targeting this process including CID1067700 (CID; Rab7a inhibitor), chlorpromazine (CPZ) and

promethazine (PMZ) both suppress clathrin function in cells through signaling receptor inhibition. Finally, we tested amlexanox (AMX) which inhibits TANK binding kinase 1 (TBK1) and its adaptor protein TBK-binding protein 1 (TBKBP1) which has been reported to variously regulate Rab7a activity [40] or induction of IFN response genes [41]. The heat map in Fig. 7A shows the fold enrichment of sgRNAs targeting the genes of interest across the screens performed in this study.



In an initial experiment of the entire panel of small molecules, inhibitors were added to culture supernatants at the initiation of SARS-CoV-2 infection and evaluated for their capacity to inhibit virus-induced CPE at 3 dpi in Vero E6 cells. The concentrations of inhibitors used, based on available toxicity data, were generally nontoxic in Vero E6 cells (Fig. 7B, white bars). ABE, AMX, HAR, NIN, OPB, PMZ, and UC2 significantly inhibited virus-induced cytotoxicity while AZ1, CPZ, INDY, and CID

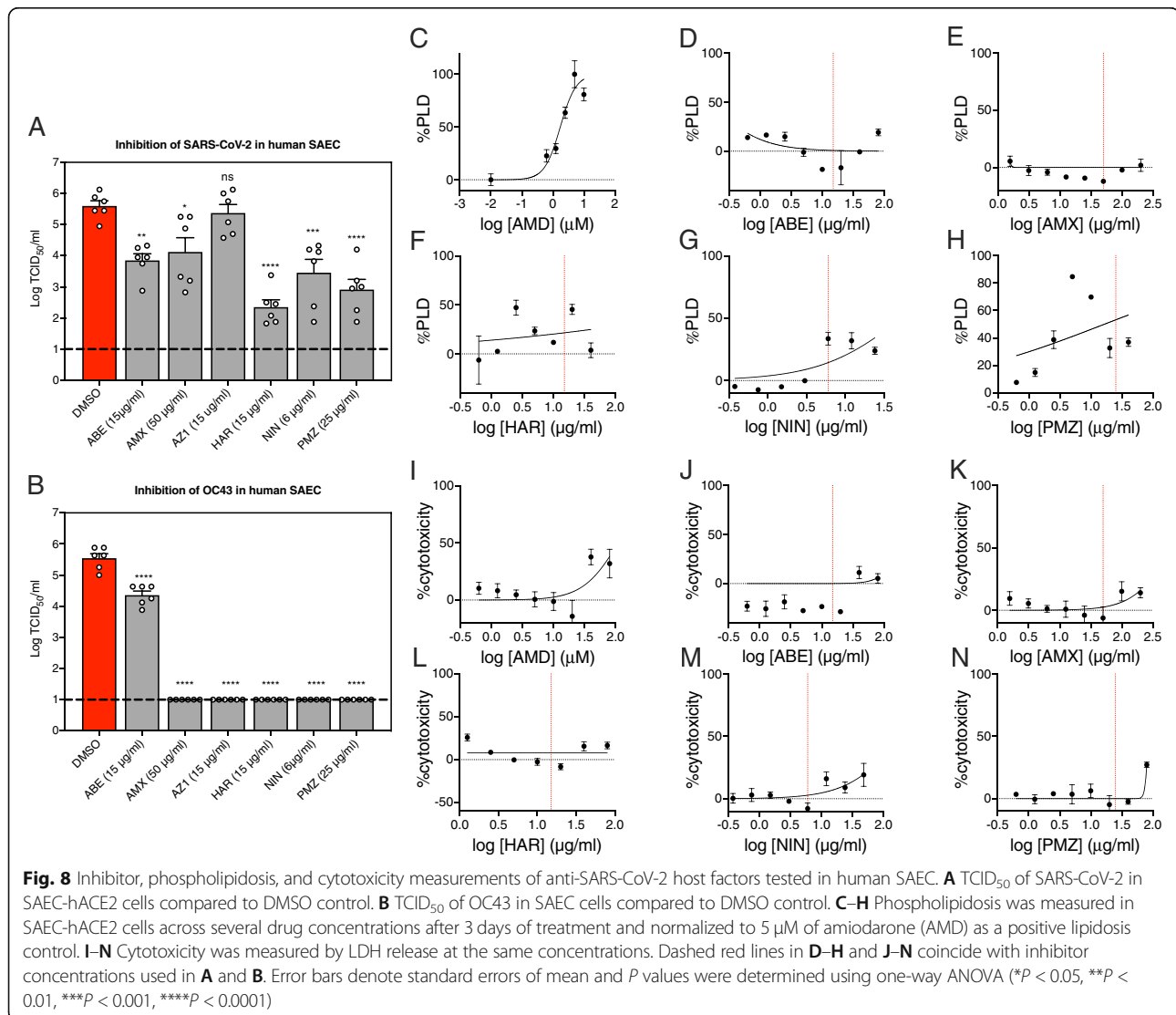
did not (Fig. 7B, gray bars). Our CID results are consistent with prior work which showed reduced CoV egress, but no effect on cell viability or viral replication, in response to CID treatment [7]. Although INDY and HAR both target Dyrk1A, only HAR displayed activity in this assay, potentially due to the lower enzymatic EC_{50} of HAR for Dyrk1A (0.24 μM for INDY vs. 0.08 μM for HAR). As a complementary approach to measure anti-viral activity of these compounds, we quantified viral

genome copies by RT-qPCR in cells treated with each compound at 2 dpi. ABE, AMX, HAR, PMZ, and UC2 significantly decreased viral genome copy number (Fig. 7C), consistent with their ability to protect from virus-induced cytotoxicity. On the other hand, NIN and OPB had no effect on viral genome copy number despite their moderate inhibition of SARS-CoV-2-induced cytotoxicity. Conversely, AZ1 completely inhibited viral genome replication in spite of having no significant effect on cytotoxicity.

For compounds displaying activity in one or both of these assays, we next determined their EC_{50} and CC_{50} against SARS-CoV-2 infection by cytotoxicity measurements in the presence or absence of virus across a series of inhibitor dilutions (Fig. 7D–J). Several of the inhibitors had EC_{50} values below 20 μ M (10.86 μ M for ABE, 14.1 μ M for NIN, and 2.16 μ M for UC2), with the p21 inhibitor UC2 being the most potent. AMX is typically

used as a topical treatment and had a high EC_{50} at 342.96 μ M. AZ1 (37.49 μ M), HAR (61.44 μ M), and PMZ (88.41 μ M) showed intermediate EC_{50} levels. The selectivity indices (SI; ratios of CC_{50} to EC_{50}) of the investigated compounds from highest to lowest are AMX > 53.23, ABE > 5.68, UC2 > 7.40 with the SI of AZ1, HAR, NIN, and PMZ falling below 2. For comparison, the SI of clinically relevant antiviral drugs are as follows: remdesivir is > 129.87, nafamostat > 4.44, and ribavirin > 3.65 [42]. Generally, determining EC_{50} with cytotoxicity measurements results in overestimation of EC_{50} , leading to a conservative estimate of SI.

A subset of inhibitors displaying efficacy in Vero E6 cells was further assessed for their capacity to inhibit SARS-CoV-2 and OC43 replication in SAEC-hACE2 or SAEC, respectively, using $TCID_{50}$ assay as a readout of infectious virus titers. ABE, AMX, AZ1, HAR, NIN, and PMZ inhibited OC43 replication (Fig. 8B) and all



compounds except for AZ1 inhibited SARS-CoV-2 replication (Fig. 8A). Phospholipidosis of cell membranes by drug treatment has been implicated as a confounding issue during in vitro viral inhibition screens [43]. While others have disputed this claim [44], we decided to test our compounds for phospholipidosis induction. In Fig. 8C–H, phospholipidosis was measured in SAEC-hACE2 cells treated with each inhibitor. Compared to the positive phospholipidosis control amiodarone (AMD), induction of phospholipidosis by PMZ was strongest followed by HAR. The phospholipidosis curve for HAR was biphasic, indicating a potential therapeutic window between 2.5 and 20 µg/ml. NIN induced minimal levels of phospholipidosis while ABE and AMX did not induce phospholipidosis. Overall, these findings reveal novel candidates for anti-HCoV drug development.

Discussion

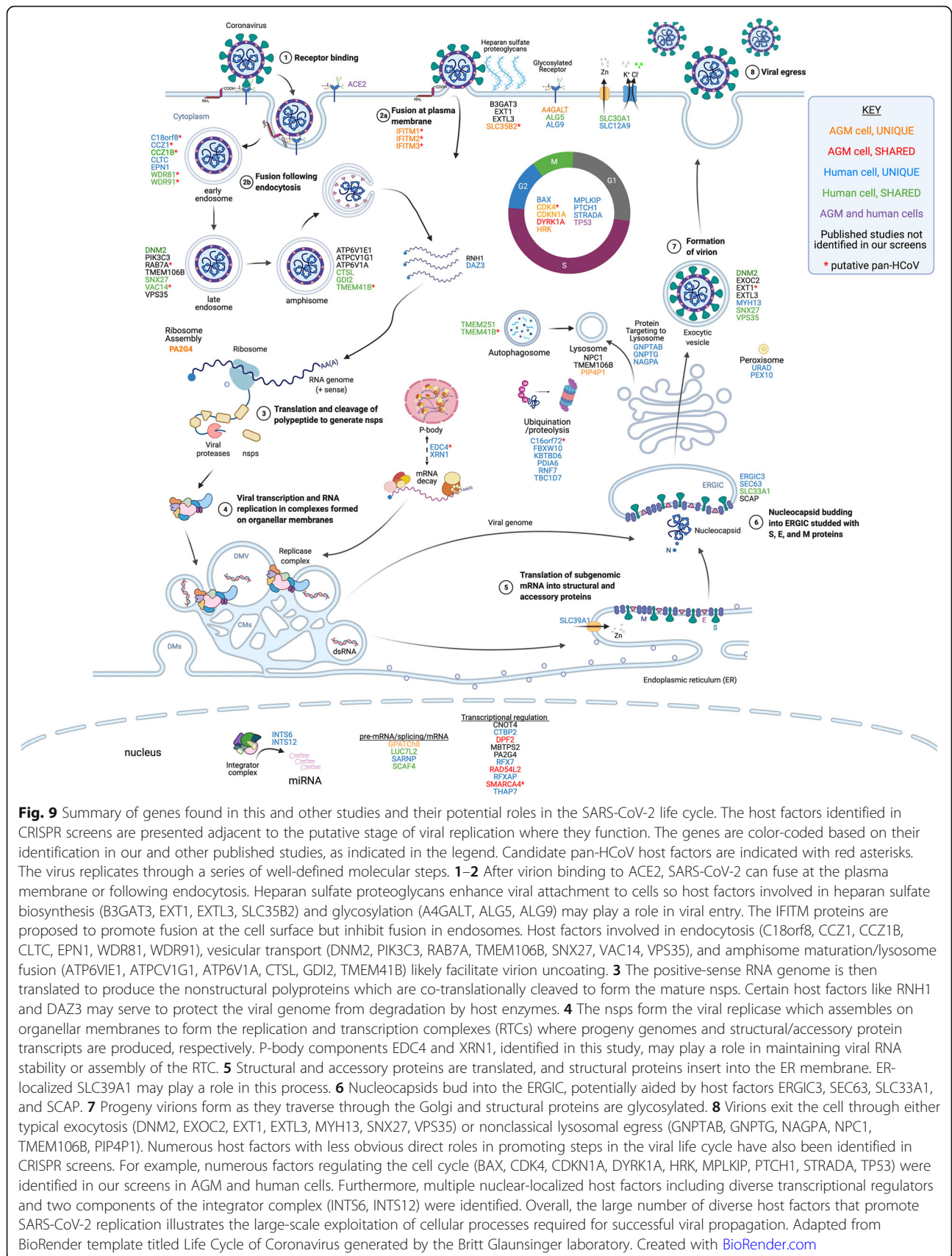
Genome-wide CRISPR knockout screens have been very successful in identifying host factors required for viral infection so it is not surprising that this approach has been applied to the discovery of proviral factors for SARS-CoV-2 infection. Indeed, six recent studies have reported CRISPR screens in cells infected with SARS-CoV-2 [10–15]. As has been observed generally for genome-wide screens, there was limited overlap in the set of genes reported as host factors. In our reanalysis of the data using a common framework, there were only 17 genes identified in two or more published screens performed in human cells (*ACE2*, *ATP6V1A*, *ATP6V1G1*, *B3GAT3*, *CCZ1B*, *CNOT4*, *CTSL*, *DNM2*, *EXOC2*, *EXT1*, *EXTL3*, *MBTPS2*, *PIK3C3*, *SCAP*, *TMEM106B*, *VAC14*, and *VPS35*). This finding is not unexpected considering that the screens were performed in a variety of cell lines and under varying infection conditions. There was more overlap in the functional categories identified across studies, with enrichment of genes involved in glycosaminoglycan biosynthesis, vesicle transport, and ER/Golgi-localized proteins [39]. Due to the limited redundancy of specific host factors identified across published studies and the potential of proviral gene products to be targeted with antiviral drugs, additional genome-wide screens are warranted. To that end, we performed CRISPR screens in AGM Vero E6 cells and human HEK293T-hACE2 cells. We performed screens for both SARS-CoV-2 and the common cold-causing HCoV OC43 to increase the probability of identifying pan-HCoV proviral factors representing strong targets for developing broad-spectrum antivirals. Our study provides additional support for previously identified candidate host factors and reports multiple novel host factors and pathways playing potentially

key roles in viral replication. We summarize the consolidated set of candidate host factors identified for SARS-CoV-2 in our study as well as those identified in two or more studies in Fig. 9.

Host factors promoting SARS-CoV-2 infection of Vero E6 cells (FDR < 0.25) are indicated in Fig. 9 adjacent to the presumptive step in the viral life cycle in which they function. In this cell line, cell cycle regulation was key to SARS-CoV-2 replication. *CDK4* was a top-scoring gene for both SARS-CoV-2 and OC43, suggesting it is broadly required for HCoV replication. CoVs utilize diverse strategies to manipulate the host cell cycle to promote their replication [45]. Identification of specific cell cycle-related host factors required for HCoV replication could provide clues to dissecting viral regulatory mechanisms. We also identified IFITM proteins in both SARS-CoV-2 and OC43 screens in Vero E6 cells, consistent with a prior study reporting that IFITM proteins promote OC43 infection [35]. Interestingly, recent work suggests that IFITM proteins promote HCoV entry when it occurs at the plasma membrane but inhibit HCoV entry when it occurs in the endocytic pathway [36, 37], suggesting that HCoVs enter Vero E6 cells primarily at the plasma membrane instead of using the endosomal pathway. This finding is consistent with the paucity of factors involved in endocytosis identified in these screens, in stark contrast to our and others' results in screens performed in human cell lines.

In addition to *CDK4* and IFITM proteins, targeting of *SLC35B2* in both SARS-CoV-2 and OC43 Vero E6 screens increased resistance to infection, suggesting that it is a pan-HCoV host factor in this cell line. *SLC35B2* encodes 3'-phosphoadenosine 5'-phosphosulfate transporter 1 (PAPST1) which plays an important role in heparan sulfate biosynthesis. PAPST1 is required for optimal replication of a variety of viruses including HIV, dengue virus, and bunyaviruses, enabling heparan sulfate-mediated viral entry or sulfating a viral receptor that enables virion binding [46–48]. It is hence logical to predict that it functions in HCoV entry in Vero E6 cells as well. Additional candidate pan-HCoV factors identified in the Vero E6 studies include *PLN* encoding phospholamban and *C16orf74* which are both implicated in maintaining calcium homeostasis [49–51], and *C3orf80* encoding a protein of unknown function. None of these gene products has been previously identified as viral host factors to our knowledge, and their functional roles will require further study.

Host factors promoting SARS-CoV-2 infection of HEK293T-hACE2 cells (FDR < 0.25) are indicated in Fig. 9 adjacent to the presumptive step in the viral life cycle in which they function. The functional categories



with the most top-scoring genes were vesicle transport, cell cycle regulation, autophagy, and ubiquitination/proteolysis. For the OC43 screens, the most abundant functional categories were vesicle transport, transcriptional regulation including the SWI/SNF complex, innate immunity, and transporters. The host factors identified in the HEK293T-hACE2 screens for both SARS-CoV-2 and OC43 (*C18orf8*, *CCZ1*, *CCZ1B*, *RAB7A*, *WDR81*, and *WDR91*) are all involved in vesicle-mediated transport and particularly in endosomal maturation, underscoring the importance of this process for HCoV infection.

When comparing our SARS-CoV-2 data sets in HEK293T-hACE2 cells to published data sets in other human cell lines, there were 21 genes in common with other studies (FDR < 0.25; *ACE2*, *ALG5*, *ARVCF*, *CCZ1B*, *CTSL*, *DNM2*, *EPT1*, *GDI2*, *LUC7L2*, *RAB7A*, *RNH1*, *SCAF4*, *SLC30A1*, *SLC33A1*, *SNX27*, *TMEM41B*, *TMEM251*, *VAC14*, *VPS35*, *WDR81*, and *WDR91*) which highlight key functional pathways required for viral infection, including endocytosis, glycosylation, and exocytosis. Remarkably though, we identified 53 unique genes, underscoring the importance of continued screening to fully elucidate host factors promoting SARS-CoV-2 replication. Certain unique genes function in previously identified pathways such as vesicle transport (e.g., *CCZ1*, *C18orf8*) and ER/Golgi-localized proteins (e.g., *SEC63*, *ERGIC3*). Other unique genes function in processes that have not been previously described as proviral in HCoV infections. For example, *EDC4* was a top-scoring gene in our SARS-CoV-2 screens in HEK293T-hACE2 cells. *EDC4* functions as a scaffold protein for the assembly of the programmed mRNA decay complex. Although it has not been reported to play a role in HCoV infection before, it does promote rotavirus replication complex assembly [52]. Another component of the programmed mRNA decay pathway *XRN1* was also modestly enriched, suggesting that this pathway promotes SARS-CoV-2 replication. Alternatively, *EDC4* and *XRN1* are both P-body components. Many RNA viruses interact with and hijack P-bodies in order to promote viral replication [53] and SARS-CoV-2 has recently been reported to disrupt P-bodies [54] so it is possible that the virus interacts with these host factors to disassemble P-bodies and facilitate viral replication. We also identified three unique genes encoding factors involved in targeting proteins to lysosomes—*GNPTAB*, *GNPTG*, and *NAGPA*. Considering recent progress in understanding the key role played by lysosomes in HCoV egress [7], it is interesting to speculate that HCoVs interact with these proteins to facilitate virion release from the infected cell.

Two approaches were taken to validate the proviral role of a subset of unique host factors identified in our screens. First, shRNA-mediated knockdown of *CCZ1*, *EDC4*, and *XRN1* resulted in reduced SARS-

CoV-2 and OC43 replication. Second, drugs targeting selected host factors displayed antiviral efficacy in vitro against SARS-CoV-2. These include cell cycle inhibitors ABE targeting *Cdk4*, AZ1 targeting *Usp25/28*, HAR targeting *Dyrk1A*, NIN targeting *Fgfr1/2/3*, and UC2 targeting *p21*; the endocytosis inhibitor PMZ targeting *Wdr81*; and the *Tbk1* inhibitor AMX. Chen et al. recently reported similar activity of ABE against SARS-CoV-2 [55], validating our findings. To our knowledge, the discovery that AMX, AZ1, HAR, NIN, PMZ, and UC2 possess antiviral activity against SARS-CoV-2 has not been reported. AMX, PMZ, and NIN are currently available drugs which could potentially be repurposed, while HAR is a natural product being investigated for the treatment of a variety of diseases. While no clinical therapeutics are currently available targeting *p21* or *Usp25/28*, our data suggest that these could be worthwhile targets for drug development. Further study of the potential in vivo utility of these compounds in treating HCoV infections and their mechanism of action is warranted.

Conclusions

Our studies substantiate and expand the growing body of literature focused on understanding key HCoV-host cell interactions. The fairly limited redundancy in proviral factors identified across our study and other published studies using genome-wide CRISPR screens [10–15] highlights the extensive scope of these interactions and suggests that even more host factors remain to be discovered. Cell type differences and variable infection conditions undoubtedly influence the outcomes of screens and could provide novel insight into nuanced viral replication mechanisms. For example, we identified lysosomal proteins as proviral in HEK293T-hACE2 cells but not in Vero E6 cells, raising the possibility that there are cell type-specific differences in the use of lysosomes for HCoV egress [7]. Detailed molecular studies testing hypotheses like this stemming from genome-scale CRISPR screens are a critical next step. Similarly, although we have identified novel compounds displaying antiviral activity against HCoVs in vitro, additional work is needed to determine their mechanism of action at the molecular level and in vivo efficacy before they can be applied in the clinic.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13073-022-01013-1>.

Additional file 1. Supplementary Methods, Figures, and Tables.

Additional file 2. AGM CRISPR KO sgRNAs.

Additional file 3. Raw and normalized counts for each CRISPR screen performed in this study.

Additional file 4. VERO E6 SARS-CoV-2 study analysis.

Additional file 5. VERO E6 OC43 study analysis.

Additional file 6. HEK293T-Ace2 SARS-CoV-2 CRISPR Study 1 analysis.

Additional file 7. HEK293T-Ace2 SARS-CoV-2 CRISPR Study 2 analysis.

Additional file 8. HEK293T-Ace2 OC43 CRISPR study analysis.

Additional file 9. Summaries of HEK293T-Ace2 and VERO E6 screens with SARS-CoV-2 and OC43.

Additional file 10. Reanalysis of published human CRISPR screen data sets.

Acknowledgements

We thank Dr. John Lednický (University of Florida, UF) and the UF Health Cancer Center for providing key reagents.

Authors' contributions

All authors read and approved the final manuscript. Conceptualization: MHN, SMK, and CDV. Methodology: MG, APB, MS, AT, MHN, SMK, and CDV. Investigation: MG, APB, MS, AT, MR, AS, RR, and MHN. Funding acquisition: MHN, SMK, and CDV. Project administration: MHN, SMK, and CDV. Supervision: MHN, SMK, and CDV. Writing the original draft: MG and APB. Writing, review, and editing: MHN, SMK, and CDV.

Funding

This work was supported by a UF Clinical and Translational Science Institute award (S.M.K., C.D.V., and M.H.N.) and the Emerging Pathogens Institute (M.H.N.). An NIH R01AI123144 further supported S.M.K. and M.G.

Availability of data and materials

All data are available in the main text, in the supplementary materials, or in the GEO database with accession GSE177545 <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE177545> [24]. The GEO dataset contains all the raw sequence files for all CRISPR barcode sequencing data generated in this study. The supplement contains read counts for the corresponding CRISPR samples and the data analysis. Custom scripts for the data analysis pipeline are available on github <https://github.com/moritzschaefer/covid19-screens> [21]. In addition, we have created and hosted a website (sarscrisprscreens.epi.ufl.edu) which contains data for previously published HCoV CRISPR screens and our integrated MAGeCK analysis using the VISPR pipeline of the SARS-CoV-2 screens.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

A patent entitled "Methods of Treatment for SARS-CoV-2 Infections (PROV Appl. No. 63/145,763)" for use of compounds identified in this paper as COVID treatments was filed on February 4, 2021, with S.M.K., C.D.V., and M.H.N. listed as inventors. The provisional patent was filed after the data in this work were obtained. The remaining authors declare that they have no competing interests.

Author details

¹Department of Molecular Genetics and Microbiology, College of Medicine, University of Florida, Gainesville, FL, USA. ²Department of Geography, College of Liberal Arts and Sciences, University of Florida, Gainesville, FL, USA. ³Emerging Pathogens Institute, University of Florida, Gainesville, FL, USA. ⁴Department of Physiological Sciences, College of Veterinary Medicine, University of Florida, Gainesville, FL, USA. ⁵Present address: Department of Medical Oncology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA, USA. ⁶University of Florida Health Cancer Center, University of Florida, Gainesville, FL, USA. ⁷Department of Pharmacotherapy and Translational Research, College of Pharmacy, University of Florida, Gainesville, FL, USA.

Received: 29 October 2021 Accepted: 10 January 2022

Published online: 27 January 2022

References

- Baden LR, El Sahly HM, Essink B, Kotloff K, Frey S, Novak R, et al. Efficacy and safety of the mRNA-1273 SARS-CoV-2 vaccine. *N Engl J Med.* 2021;384:403–16.
- Polack FP, Thomas SJ, Kitchin N, Absalon J, Gurtman A, Lockhart S, et al. Safety and efficacy of the BNT162b2 mRNA Covid-19 vaccine. *N Engl J Med.* 2020;383(27):2603–15. <https://doi.org/10.1056/NEJMoa2034577>.
- Sadoff J, Gray G, Vandebosch A, Cárdenas V, Shukarev G, Grinsztejn B, et al. Safety and efficacy of single-dose Ad26.COV2.S vaccine against Covid-19. *N Engl J Med.* 2021;384:2187–201.
- M. Voysey, S. A. C. Clemens, S. A. Madhi, L. Y. Weckx, P. M. Folegatti, P. K. Aley, et al. Safety and efficacy of the ChAdOx1 nCoV-19 vaccine (AZD1222) against SARS-CoV-2: an interim analysis of four randomised controlled trials in Brazil, South Africa, and the UK. *The Lancet* 397, 99–111 (2021).
- Doria-Rose NA, Shen X, Schmidt SD, O'Dell S, McDanal C, Feng W, et al. Booster of mRNA-1273 vaccine reduces SARS-CoV-2 omicron escape from neutralizing antibodies. *medRxiv*, 2021.12.15.21267805. 2021.
- Hartenian E, Nandakumar D, Lari A, Ly M, Tucker JM, Glaunsinger BA. The molecular virology of coronaviruses. *J Biol Chem.* 2020;295(37):12910–34. <https://doi.org/10.1074/jbc.REV120.013930>.
- Ghosh S, Dellibovi-Ragheb TA, Kerviel A, Pak E, Qiu Q, Fisher M, et al. β -Coronaviruses use lysosomes for egress instead of the biosynthetic secretory pathway. *Cell.* 2020;183:1520–1535.e14.
- Wang Y, Jin F, Wang R, Li F, Wu Y, Kitazato K, et al. HSP90: a promising broad-spectrum antiviral drug target. *Arch Virol.* 2017;162(11):3269–82. <https://doi.org/10.1007/s00705-017-3511-1>.
- Debing Y, Neyts J, Delang L. The future of antivirals: broad-spectrum inhibitors. *Curr Opin Infect Dis.* 2015;28(6):596–602. <https://doi.org/10.1097/QCO.0000000000000212>.
- Zhu Y, Feng F, Hu G, Wang Y, Yu Y, Zhu Y, et al. A genome-wide CRISPR screen identifies host factors that regulate SARS-CoV-2 entry. *Nat Commun.* 2021;12(1):961. <https://doi.org/10.1038/s41467-021-21213-4>.
- Wang R, Simoneau CR, Kulsuptrakul J, Bouhaddou M, Travasano KA, Hayashi JM, et al. Genetic screens identify host factors for SARS-CoV-2 and common cold coronaviruses. *Cell.* 2020;184(1):106–19. <https://doi.org/10.1016/j.cell.2020.12.004>.
- Schneider WM, Luna JM, Hoffmann H-H, Sánchez-Rivera FJ, Leal AA, Ashbrook AW, et al. Genome-scale identification of SARS-CoV-2 and Pan-coronavirus host factor networks. *Cell.* 2021;184:120–132.e14.
- Wei J, Alfajaro MM, DeWeirdt PC, Hanna RE, Lu-Culligan WJ, Cai WL, et al. Genome-wide CRISPR screens reveal host factors critical for SARS-CoV-2 infection. *Cell.* 2020;184:76–91.
- Daniloski Z, Jordan TX, Wessels H-H, Hoagland DA, Kasela S, Legut M, et al. Identification of required host factors for SARS-CoV-2 infection in human cells. *Cell.* 2020;184:92–105.
- Baggen J, Persoons L, Jansen S, Vanstreels E, Jacquemyn M, Jochmans D, et al. Identification of TMEM106B as proviral host factor for SARS-CoV-2. *Nat Genet.* 2021;53(4):435–44. <https://doi.org/10.1038/s41588-021-00805-2>.
- Reznikov LR, Norris MH, Vashisht R, Bluhm AP, Li D, Liao Y-SJ, et al. Identification of antiviral antihistamines for COVID-19 repurposing. *Biochem Biophys Res Commun.* 2021;538:173–9. <https://doi.org/10.1016/j.bbrc.2020.11.095>.
- Hammit LL, Kazungu S, Welch S, Bett A, Onyango CO, Gunson RN, et al. Added value of an oropharyngeal swab in detection of viruses in children hospitalized with lower respiratory tract infection. *J Clin Microbiol.* 2011; 49(6):2318–20. <https://doi.org/10.1128/JCM.02605-10>.
- Nabokina SM, Inoue K, Subramanian VS, Valle JE, Yuasa H, Said HM. Molecular identification and functional characterization of the human colonic thiamine pyrophosphate transporter. *J Biol Chem.* 2014;289(7):4405–16. <https://doi.org/10.1074/jbc.M113.528257>.
- Doench JG, Fusi N, Sullender M, Hegde M, Vaimberg EW, Donovan KF, et al. Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat Biotechnol.* 2016;34(2):184–91. <https://doi.org/10.1038/nbt.3437>.
- Shalem O, Sanjana NE, Hartenian E, Shi X, Scott DA, Mikkelsen TS, et al. Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science.* 2014;343(6166):84–7. <https://doi.org/10.1126/science.1247005>.
- Schäfer M. covid19-screens. GitHub. <https://github.com/moritzschaefer/covid19-screens>. 2021.

22. Gordon A, Hannon GJ. GitHub. FASTX-Toolkit. https://github.com/agordon/fastx_toolkit. 2010.
23. Li W, Xu H, Xiao T, Cong L, Love MI, Zhang F, et al. MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome Biol.* 2014;15(12):554. <https://doi.org/10.1186/s13059-014-0554-4>.
24. M. Grodzki, A. Bluhm, M. Schäfer, A. Tagmount, M. Russo, A. Sobh, R. Raifée, C. Vulpe, S. Karst, M. H. Norris, Genome-wide knockout screen of OC43/SARS-CoV2-infected HEK293-hACE2 and Vero E6 cells. GSE177545. Gene Expression Omnibus 2021. <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE177545> (2021).
25. Moffat J, Grueneberg DA, Yang X, Kim SY, Kloepper AM, Hinkle G, et al. A lentiviral RNAi library for human and mouse genes applied to an arrayed viral high-content screen. *Cell.* 2006;124(6):1283–98. <https://doi.org/10.1016/j.cell.2006.01.040>.
26. Schäfer M, Clevert D-A, Weiss B, Steffen A. PAVOOC: designing CRISPR sgRNAs using 3D protein structures and functional domain annotations. *Bioinformatics.* 2019;35(13):2309–10. <https://doi.org/10.1093/bioinformatics/bty935>.
27. Wishart DS, Feunang YD, Guo AC, Lo EJ, Marcu A, Grant JR, et al. DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res.* 2018;46:D1074–82.
28. Szklarczyk D, Santos A, von Mering C, Jensen LJ, Bork P, Kuhn M. STITCH 5: augmenting protein-chemical interaction networks with tissue and affinity data. *Nucleic Acids Res.* 2016;44(D1):D380–4. <https://doi.org/10.1093/nar/gkv1277>.
29. Zhou Y, Zhang Y, Lian X, Li F, Wang C, Zhu F, et al. Therapeutic target database update 2022: facilitating drug discovery with enriched comparative data of targeted agents. *Nucleic Acids Res.* 2021;50(D1):D1398–407.
30. Whirl-Carrillo M, Huddart R, Gong L, Sangkuhl K, Thorn CF, Whaley R, et al. An evidence-based framework for evaluating pharmacogenomics knowledge for personalized medicine. *Clin Pharmacol Ther.* 2021;110(3):563–72. <https://doi.org/10.1002/cpt.2350>.
31. Stelzer G, Rosen N, Plaschkes I, Zimmerman S, Twik M, Fishilevich S, et al. The GeneCards suite: from gene data mining to disease genome sequence analyses. *Curr Protoc Bioinform.* 2016;54(1):1.30.1–1.30.33. <https://doi.org/10.1002/cpbi.5>.
32. Cui Y, Wang Z, Köster J, Liao X, Peng S, Tang T, et al. VISPR-online: a web-based interactive tool to visualize CRISPR screening experiments. *BMC Bioinformatics.* 2021;22(1):344. <https://doi.org/10.1186/s12859-021-04275-5>.
33. Li W, Köster J, Xu H, Chen C-H, Xiao T, Liu JS, et al. Quality control, modeling, and visualization of CRISPR screens with MAGeCK-VISPR. *Genome Biol.* 2015;16(1):281. <https://doi.org/10.1186/s13059-015-0843-6>.
34. Hoffmann M, Kleine-Weber H, Schroeder S, Krüger N, Herrler T, Erichsen S, et al. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell.* 2020;181:271–280.e8.
35. Zhao X, Guo F, Liu F, Cuconati A, Chang J, Block TM, et al. Interferon induction of IFITM proteins promotes infection by human coronavirus OC43. *Proc Natl Acad Sci U S A.* 2014;111(18):6756–61. <https://doi.org/10.1073/pnas.1320856111>.
36. Shi G, Kenney AD, Kudryashova E, Zani A, Zhang L, Lai KK, et al. Opposing activities of IFITM proteins in SARS-CoV-2 infection. *EMBO J.* 2021;40(3):e106501. <https://doi.org/10.15252/embj.2020106501>.
37. Winstone H, Lista MJ, Reid AC, Bouton C, Pickering S, Galao RP, et al. The polybasic cleavage site in the SARS-CoV-2 spike modulates viral sensitivity to Type I interferon and IFITM2. *J Virol.* 2021;95(9). <https://doi.org/10.1128/JVI.02422-20>.
38. Ou X, Liu Y, Lei X, Li P, Mi D, Ren L, et al. Characterization of spike glycoprotein of SARS-CoV-2 on virus entry and its immune cross-reactivity with SARS-CoV. *Nat Commun.* 2020;11(1):1620. <https://doi.org/10.1038/s41467-020-15562-9>.
39. Bailey AL, Diamond MS. A Crisp(r) New perspective on SARS-CoV-2 biology. *Cell.* 2020;0. <https://doi.org/10.1016/j.cell.2020.12.003>.
40. Heo J-M, Ordureau A, Swarup S, Paulo JA, Shen K, Sabatini DM, et al. RAB7A phosphorylation by TBK1 promotes mitophagy via the PINK-PARKIN pathway. *Sci Adv.* 2018;4:eav0443.
41. Revach O-Y, Liu S, Jenkins RW. Targeting TANK-binding kinase 1 (TBK1) in cancer. *null.* 2020;24:1065–78.
42. Wang M, Cao R, Zhang L, Yang X, Liu J, Xu M, et al. Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (2019-nCoV) in vitro. *Cell Res.* 2020;30(3):269–71. <https://doi.org/10.1038/s41422-020-0282-0>.
43. Tummino Tia A, Rezeli Veronica V, Benoit F, Audrey F, O'Meara Matthew J, Blandine M, et al. Drug-induced phospholipidosis confounds drug repurposing for SARS-CoV-2. *Science.* 2021;373:541–7.
44. Lane TR, Ekins S. Defending antiviral cationic amphiphilic drugs that may cause drug-induced phospholipidosis. *J Chem Inf Model.* 2021;61(9):4125–30. <https://doi.org/10.1021/acs.jcim.1c00903>.
45. Su M, Chen Y, Qi S, Shi D, Feng L, Sun D. A mini-review on cell cycle regulation of coronavirus infection. *Front Vet Sci.* 2020;7. <https://doi.org/10.3389/fvets.2020.586826>.
46. Thamamongood T, Aebischer A, Wagner V, Chang MW, Elling R, Benner C, et al. A genome-wide CRISPR-Cas9 screen reveals the requirement of host cell sulfation for Schmallenberg virus infection. *J Virol.* 2020;94. <https://doi.org/10.1128/JVI.00752-20>.
47. Gao H, Lin Y, He J, Zhou S, Liang M, Huang C, et al. Role of heparan sulfate in the Zika virus entry, replication, and cell death. *Virology.* 2019;529:91–100. <https://doi.org/10.1016/j.virol.2019.01.019>.
48. Park RJ, Wang T, Koundakjian D, Lamothe-Molina P, Monel B, et al. A genome-wide CRISPR screen identifies a restricted set of HIV host dependency factors. *Nat Genet.* 2017;49(2):193–203. <https://doi.org/10.1038/ng.3741>.
49. Minamisawa S, Hoshijima M, Chu G, Ward CA, Frank K, Gu Y, et al. Chronic phospholamban-sarcoplasmic reticulum calcium ATPase interaction is the critical calcium cycling defect in dilated cardiomyopathy. *Cell.* 1999;99(3):313–22. [https://doi.org/10.1016/S0092-8674\(00\)81662-1](https://doi.org/10.1016/S0092-8674(00)81662-1).
50. Kushibiki T, Nakamura T, Tsuda M, Tsuchikawa T, Hontani K, Inoko K, et al. Role of dimerized C16orf74 in aggressive pancreatic cancer: a novel therapeutic target. *Mol Cancer Ther.* 2020;19(1):187–98. <https://doi.org/10.1158/1535-7163.MCT-19-0491>.
51. Nakamura T, Katagiri T, Sato S, Kushibiki T, Hontani K, Tsuchikawa T, et al. Overexpression of C16orf74 is involved in aggressive pancreatic cancers. *Oncotarget.* 2017;8(31):50460–75. <https://doi.org/10.18632/oncotarget.10912>.
52. Dhilon P, Rao CD. Rotavirus induces formation of remodeled stress granules and P bodies and their sequestration in viroplasm to promote progeny virus production. *J Virol.* 2018;92. <https://doi.org/10.1128/JVI.01363-18>.
53. Beckham CJ, Parker R. P Bodies, Stress granules, and viral life cycles. *Cell Host Microbe.* 2008;3(4):206–12. <https://doi.org/10.1016/j.chom.2008.03.004>.
54. Robinson C-A, Kleer M, Mulloy RP, Castle EL, Boudreau BQ, Corcoran JA. Human coronaviruses disassemble processing bodies. *bioRxiv.* 2020.11.08.372995. 2020.
55. Chen CZ, Xu M, Pradhan M, Gorshkov K, Petersen JD, Straus MR, et al. Identifying SARS-CoV-2 entry inhibitors through drug repurposing screens of SARS-S and MERS-S pseudotyped particles. *ACS Pharmacol Transl Sci.* 2020;3(6):1165–75. <https://doi.org/10.1021/acspsci.0c00112>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

