

A Model for Network-Based Identification and Pharmacological Targeting of Aberrant, Replication-Permissive Transcriptional Programs Induced by Viral Infection

Pasquale Laise

DarwinHealth, Inc

Megan Stanifer

Gideon Bosker

Xiaoyun Sun

Sergio Triana

Patricio Doldan

Federico La Manna

Marta De Menna

University of Bern

Ronald Realubit

Sergey Pampou

Charles Karan

Theodore Alexandrov

Marianna Kruihof-de Julio

University of Bern <https://orcid.org/0000-0002-6085-7706>

Andrea Califano

Steeve Boulant

Mariano Alvarez (✉ malvarez@darwinhealth.com)

DarwinHealth, Inc <https://orcid.org/0000-0002-7503-2491>

Article

Keywords:

Posted Date: February 4th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1287631/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

1 **A model for network-based identification and pharmacological**
2 **targeting of aberrant, replication-permissive transcriptional programs**
3 **induced by viral infection**
4

5 Pasquale Laise^{1,2,*}, Megan L. Stanifer^{3,*}, Gideon Bosker¹, Xiaoyun Sun¹, Sergio Triana^{4,5}, Patricio
6 Doldan^{6,7}, Federico La Manna^{8,9,10,11}, Marta De Menna^{8,9,10,11}, Ronald B. Realubit², Sergey
7 Pampou², Charles Karan², Theodore Alexandrov^{4,12,13}, Marianna Kruithof-de Julio^{8,9,10,11}, Andrea
8 Califano^{2,14,15,16,17,+}, Steeve Boulant^{6,7,+} and Mariano J. Alvarez^{1,2,+}

9 ¹ DarwinHealth Inc, New York, NY, USA

10 ² Department of Systems Biology, Columbia University Irving Medical Center, New York, NY, USA

11 ³ Department of Infectious Diseases, Molecular Virology, Heidelberg University Hospital,
12 Germany

13 ⁴ Structural and Computational Biology Unit, European Molecular Biology Laboratory, Heidelberg,
14 Germany

15 ⁵ Collaboration for joint PhD degree between EMBL and Heidelberg University, Faculty of
16 Biosciences, Germany

17 ⁶ Department of Infectious Diseases, Virology, Heidelberg University Hospital, Germany

18 ⁷ Research Group "Cellular Polarity and Viral Infection", German Cancer Research Center
19 (DKFZ), Heidelberg, Germany

20 ⁸ Department for BioMedical Research, Urology Research Laboratory, University of Bern, Bern,
21 Switzerland

22 ⁹ Translational Organoid Models, Department for BioMedical Research, University of Bern, Bern,
23 Switzerland

24 ¹⁰ Bern Center for Precision Medicine, University of Bern and Inselspital, Bern, Switzerland

25 ¹¹ Department of Urology, Inselspital, Bern University Hospital, Bern, Switzerland

26 ¹² Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego,
27 La Jolla, CA, USA

28 ¹³ Molecular Medicine Partnership Unit (MMPU), European Molecular Biology Laboratory,
29 Heidelberg, Germany

30 ¹⁴ Herbert Irving Comprehensive Cancer Center, Columbia University Irving Medical Center, New
31 York, NY, USA

32 ¹⁵ Department of Medicine, Columbia University Irving Medical Center, New York, NY, USA

33 ¹⁶ Department of Biochemistry & Molecular Biophysics, Columbia University Irving Medical
34 Center, New York, NY, USA

35 ¹⁷ Department of Biomedical Informatics, Columbia University Irving Medical Center, New York,
36 NY, USA

37

38 * These authors contributed equally.

39 ⁺ To whom correspondence should be addressed:

40 Andrea Califano: ac2248@cumc.columbia.edu

41 Steeve Boulant: s.boulant@dkfz-heidelberg.de

42 Mariano Alvarez: malvarez@darwinhealth.com

43

44

45 **ABSTRACT**

46

47 Precise characterization and targeting of host cell transcriptional machinery hijacked by viral
48 infection remains challenging. Here, we show that SARS-CoV-2 hijacks the host cell
49 transcriptional machinery to induce a phenotypic state amenable to its replication. Specifically,
50 analysis of Master Regulator (MR) proteins representing mechanistic determinants of the gene
51 expression signature induced by SARS-CoV-2 in infected cells revealed coordinated inactivation
52 of MRs enriched in physical interactions with SARS-CoV-2 proteins, suggesting their mechanistic
53 role in maintaining a host cell state refractory to virus replication. To test their functional relevance,
54 we measured SARS-CoV-2 replication in epithelial cells treated with drugs predicted to activate
55 the entire repertoire of repressed MRs, based on their experimentally elucidated, context-specific
56 mechanism of action. Overall, >80% of drugs predicted to be effective by this methodology
57 induced significant reduction of SARS-CoV-2 replication, without affecting cell viability. This model
58 for host-directed pharmacological therapy is fully generalizable and can be deployed to identify
59 drugs targeting host cell-based MR signatures induced by virtually any pathogen.

60

61 INTRODUCTION

62 Several approaches have been employed to identify specific host cell pathways and proteins
63 whose individual interaction with viral proteins is either required to mediate SARS-CoV-2 infection
64 or that represents key modulators of virulence ¹⁻⁶. In contrast, a paucity of effort has been devoted
65 to elucidating the host cell transcriptional control mechanisms and programs hijacked by viruses,
66 including identification of the Master Regulator (MR) proteins that mediate the infection-mediated
67 reprogramming of the host cell transcriptional state. More importantly, there has been no
68 experimental evaluation of the role of such host MR proteins in the virus life cycle nor their
69 amenability to pharmacological targeting.

70 Here, we show that host MR proteins, representing viral infection-mediated determinants of the
71 transcriptional regulatory programs hijacked by viruses, are required for establishing a host-cell
72 phenotypic state amenable to virus replication. Specifically, we leveraged an established systems
73 biology-based methodology, originally developed in the field of oncology ⁷, to identify MR proteins
74 that mechanistically control the transcriptional state of virus infected cells. We then prioritized
75 drugs capable of inverting the activity of MR proteins—thus decommissioning the regulatory
76 programs induced by viral infection to maintain a pro-infective cell state—using another oncology-
77 based approach described in ⁸. We propose that extension and translation of these cancer-based
78 methodologies to study viral infection can identify host cell MR proteins representing key
79 mechanistic determinants of virus-mediated host cell reprogramming, as well as the drugs that
80 can abrogate this transition.

81 As we have previously shown, MRs can be accurately and systematically identified by assessing
82 the enrichment of their transcriptional targets in differentially expressed genes, using the Virtual
83 Inference of Protein activity by Enriched Regulon analysis (VIPER) ⁹. While many approaches
84 can be used to identify the tissue-specific targets of a regulatory protein, the Algorithm for the

85 Accurate Reconstruction of Cellular Networks (ARACNe) ¹⁰ is among the few that have been
86 extensively experimentally tested, with validation rates exceeding 70% ¹⁰⁻¹². We have shown that
87 VIPER can accurately measure the activity levels of >70% of regulatory proteins, including in
88 single cells, where we have shown that metaVIPER ¹³—a VIPER extension specifically designed
89 for single-cell analyses—can virtually eliminate the gene dropout issue due to low single cell
90 profiling depth ^{14,15}; and, notably, outperforms antibody-based measurements ¹⁴. Hereafter, for
91 simplicity, we will refer to the transcriptional activity inferred by VIPER or mtaVIPER, as protein
92 activity. The combination of these two algorithms has been highly effective in elucidating protein-
93 based mechanisms that were virtually undetectable by gene expression-based methods alone
94 ^{7,14,16,17} (see methods for additional details). Moreover, once MR protein activity levels are
95 quantified by VIPER analysis, the CLIA-certified OncoTreat algorithm ⁸ can accurately and
96 efficiently identify small molecule inhibitors that can invert their activity (MR-inverter drugs),
97 thereby abrogating the regulatory programs they control. The OncoTreat algorithm leverages
98 large-scale gene expression profiles of MR-matched cell lines perturbed with a comprehensive
99 repertoire of clinically relevant drugs, including Food and Drug Administration (FDA)-approved
100 and late-stage experimental agents, and has led to several clinical trials evaluating drug therapy
101 for cancer (NCT02066532, NCT02632071, and NCT03211988, among others).

102 Given the urgency and unmet needs mandated by the COVID-19 pandemic, we proceeded to test
103 the applicability of this model to SARS-CoV-2 infection. Specifically, we asked whether this
104 methodology could be used to identify host cell MR proteins representing the mechanistic
105 determinants of the transcriptional programs hijacked by the virus to support efficient replication
106 and, by extension, whether we can identify drugs capable of inverting their activity, thereby
107 making host cells more resistant to hijacking and viral replication. The methodology can be
108 trivially generalized to other pathogens, conditional only on the availability of appropriate infection
109 gene expression signatures.

110 VIPER-inferred MRs from multiple SARS-CoV-2 infection models consistently showed that the
111 host MR proteins that were *significantly activated* following SARS-CoV-2 infection controlled
112 innate immune response programs. This suggests that the transcriptional programs supporting
113 optimal viral replication and infectivity, during the hijack phase, may be controlled by host MRs
114 that were *significantly inactivated* following infection. Supporting this hypothesis, we found the
115 inactivated MRs to be highly enriched in interactions with SARS-CoV-2 proteins and in genes
116 reported as essential antiviral factors by CRISPR screens ^{2,4,6}. To further test this hypothesis, we
117 adapted the OncoTreat algorithm ⁸ to prioritize compounds based on their ability to activate the
118 entire set of virus-inactivated MR proteins, and evaluated their effect on SARS-CoV-2 replication
119 in infected epithelial cell cultures. Prioritization of 154 FDA-approved drugs—primarily for use in
120 oncology—was highly effective, with >80% of the predictions effectively reducing SARS-CoV-2
121 replication in colon epithelial cells, with no significant reduction of cell viability.

122 Based on these findings, we conclude that SARS-CoV-2-induced transition of the host cell
123 phenotypic state is required for its optimal replication. Moreover, we provide a model for
124 systematically dissecting the MR proteins that mechanistically facilitate this transition and for
125 identifying MR-inverting drugs that, by blocking this phenotypic transition, can induce a host cell
126 regulatory state of “viral contraception.” This model, which we call, “ViroTreat,” could be used to
127 identify therapeutic options in the COVID-19 setting and can be easily generalized to virtually any
128 viral pathogen-mediated host cell hijacking that is essential for the infective cycle.

129 **RESULTS**

130 **SARS-CoV-2-induced MR signature**

131 To elucidate the MR proteins mediating SARS-CoV2-induced host-cell phenotypic transition, we
132 analyzed publicly available single cell (scRNASeq) profiles of SARS-CoV-2 infected epithelial
133 cells (Supplementary Table 1), including epithelial cell lines from both lung adenocarcinoma

134 (Calu-3 and H1299)¹⁸, and gastrointestinal organoid models from the ileum and colon¹⁹. Single
135 cell RNASeq analysis allows highly effective identification of individual virus-infected cells, which
136 would otherwise represent only a minority of cells in culture. Moreover, single cell-based gene
137 expression signatures—computed by comparing confirmed infected cells to non-infected
138 controls—are less affected by contamination and dilution effects typical of bulk RNASeq
139 signatures representing a mixture of infected and non-infected cells (Supplementary Fig. 1 and
140 Methods).

141 Single cell analysis revealed highly conserved differential protein activity signatures, as defined
142 by the top 50 most differentially active candidate MRs, by analogy to tumor MRs⁷. We will refer
143 to this repertoire of virus-induced MRs as the Viral CheckPoint. The analysis identified a highly
144 conserved MR core induced by SARS-CoV-2 infection, within each available cellular model,
145 across all post-infection time-points for which data was available ($p < 10^{-40}$, by 2-tailed aREA test,
146 Fig. 1a and Supplementary Fig. 2a).

147 When comparing equivalent time-points, we observed significant conservation of the differentially
148 active protein signature across lineage-related cell models (e.g., Calu-3 vs. H1299, at 12h, $p <$
149 10^{-40} , Supplementary Fig. 2a). Interestingly, the virus-mediated MR signature was highly
150 conserved even across unrelated lineages, when equivalent time-points were considered (e.g.,
151 H1299 vs. colon non-transformed organoid at 24h, $p < 0.01$, Supplementary Fig. 2a). Taken
152 together, these findings suggest the existence of a highly reproducible, SARS-CoV-2-mediated
153 MR activity signature in epithelial cells, regardless of organ context (lung vs. gastrointestinal (GI)).
154 Interestingly, however, *inactivated* MRs were significantly more conserved than *activated* MRs,
155 both across models and lineages ($p < 10^{-6}$, 2-tailed paired U-test, Supplementary Fig. 2b,c),
156 suggesting a potentially distinct biological role for the activated vs. inactivated components of the
157 SARS-CoV-2 MR core.

158 The MR activity signatures detected by single cell analyses were also recapitulated by bulk-tissue
159 analysis of SARS-CoV-2-infected epithelial cells (ST1), albeit at a slightly lower statistical
160 significance, as we expected. These findings applied to both transformed models, including lung
161 (Calu-3, H1299, and A549) and colon (Caco-2) adenocarcinoma, and normal human bronchial
162 epithelial (NHBE) primary cells, as well as to more physiologic models, including lung organoids.
163 As should be expected, MR conservation was more significant for models characterized by high
164 infection rates (Supplementary Fig. 2a), likely due to signature dilution/contamination by a high
165 proportion of non-infected cells in other models.

166 **MRs govern distinct biological functions**

167 Gene Set Enrichment Analysis (GSEA)²⁰ demonstrated a critical dichotomy of biological hallmark
168 programs enriched in activated vs. inactivated MRs (Fig. 1b). Specifically, biological hallmarks
169 enriched in *activated* MRs included inflammatory response, epithelial-to-mesenchymal transition
170 (EMT) and interferon response. Indeed, among the top aberrantly activated MRs, we identified
171 MX1, a protein induced by interferon I and II²¹, the interferon regulator IRF9, and additional
172 transcriptional regulators that mediate cellular response to interferons, such as STAT1 and
173 STAT2²² (Fig. 1a).

174 In contrast, our model shows that biological hallmarks enriched in *inactivated* MRs were strongly
175 related to virus-mediated host-cell hijacking programs, such as PI3K signaling, unfolded protein
176 response, DNA repair, and metabolic-related processes^{23,24} (Fig. 1b). Consistent with this
177 observation, the most significantly inactivated MRs included several ribosomal subunit members
178 (such as RPS27A, RPS3, RPL3, RPS6, RPS14), as well as proteins involved in cell cycle arrest
179 (UBA52)²⁵, translational regulation, and cellular metabolism (GABPB1)²⁶ (Fig. 1a).

180 **VIPER-inferred MRs are enriched in SARS-CoV-2-interacting proteins**

181 To assess whether activated vs. inactivated MRs in our model may represent a more effective
182 target for drug-mediated reversal, we proceeded to assess whether either class was enriched in
183 host proteins previously identified as cognate binding partners of SARS-CoV-2 proteins. For this
184 analysis, we leveraged a collection of 332 host proteins previously reported to be involved in
185 protein-protein interactions (PPIs) with 26 of the 29 proteins encoded by the SARS-CoV-2
186 genome, as determined by mass-spec analysis of pull-down assays ². Of these interactions, 90
187 were with proteins included in the 5,734 we analyzed by VIPER. GSEA ²⁰ revealed statistically
188 significant enrichment of these 90 proteins in SARS-CoV-2 *inactivated* but not *activated* MRs,
189 across all the evaluated single-cell protein activity signatures ($p < 10^{-3}$, 2-tailed GSEA,
190 Supplementary Fig. 3). This suggests that host cell proteins that physically interact with SARS-
191 CoV-2 proteins are mostly inactivated in response to the infection.

192 **VIPER-inferred MRs are enriched in viral infection-essential genes**

193 To further confirm the functional duality of the inferred MRs, we also assessed their enrichment
194 in genes previously reported as essential to the virus infectious cycle. Specifically, we evaluated
195 their enrichment in genes identified by functional CRISPR screens from two different studies,
196 including using SARS-CoV-2 infected Vero ⁶ and Huh-7.5 ⁴ cells. Consistent with our original
197 observation and definition of the SARS-CoV-2 induced MR signature, the 50 most inactivated
198 candidate MRs—as determined by integrating results from both lung and GI models—were
199 significantly enriched in infection-essential genes identified in both CRISPR screen ($p < 10^{-4}$ and
200 $p < 10^{-3}$, respectively), as well as in the integrated set (Supplementary Fig. 4a-c, $p < 10^{-4}$). In
201 contrast, the 50 most activated MRs were not significantly enriched in infection essential genes
202 (Supplementary Fig. 4d-f).

203 **ViroTreat prioritization of FDA-approved drugs**

204 To test the dependence of SARS-CoV-2 replication on inactivation of the MR proteins—termed
205 Viral Checkpoint for analogy to tumor Checkpoints⁷—, we adapted the OncoTreat algorithm⁸ to
206 identify small molecule compounds capable of activating such MRs (ViroTreat, Fig. 2). We
207 hypothesize that such drug-induced effects would keep the host cell phenotype in a “viral
208 contraception” regulatory state that effectively reduces viral replication rate.

209 We have shown that drug Mechanism of Action (MoA)—as represented by the proteins that are
210 differentially activated/inactivated—is an effective predictor of drug activity *in vivo* and in explants
211^{27,28}. This is assessed by VIPER analysis of MR-matched cell lines following perturbation with a
212 large repertoire of drugs, at the highest sublethal concentration (IC₂₀), as assessed by dose
213 response curves. The PanACEA database (PANcancer Analysis of Chemical Entity Activity)²⁹
214 comprises drug perturbation RNA-seq profiles representing 25 cell lines and an average of 350
215 drugs per cell line. Among these, the LoVo and NCI-H1973 cell lines were identified as whose
216 lineage matched the GI epithelial and lung epithelial cell models used of SARS-CoV-2 infection
217 assays, respectively. However, while LoVo (human colon cell line) showed statistically significant
218 MR protein conservation ($p < 10^{-5}$ by OncoMatch analysis²⁷), when compared with the colon
219 adenocarcinoma cell line susceptible to SARS-CoV-2 infection (Caco-2³⁰, Supplementary Fig.
220 5a,b), such conservation was not observed between NCI-H1793 cells and any of the three lung
221 cell lines susceptible to SARS-CoV-2 infection (Calu-3, ACE2-A549 and H1299, Supplementary
222 Fig. 5c-h). Based on these results and considering availability of a compatible cell line as a
223 relevant validation model to experimentally assess ViroTreat-predicted drugs, for this model we
224 focused our validation efforts on the GI context.

225 VIPER was used to elucidate the MoA of 154 FDA-approved oncology drugs, where MoA is
226 defined as the repertoire of proteins differentially activated/inactivated at 24h following drug
227 perturbation. While this was done specifically in colon epithelial cells for this study, the analysis
228 can be easily extended to assess drug MoA in other cellular contexts. Specifically, the RNA-seq

229 profiles used in this analysis were generated at 24h (by PLATE-Seq assays³¹), following
230 treatment of a colon adenocarcinoma cell line (LoVo) with a library of FDA-approved drugs and
231 vehicle control (DMSO). To avoid assessing cell death or stress mechanisms, rather than drug
232 MoA effects, drugs were titrated at their highest sublethal concentration (i.e., their 48h IC₂₀), as
233 assessed by 10-point dose response curves (see methods for additional details). Resulting
234 profiles were then used to assess the differential activity of regulatory proteins in drug vs. vehicle
235 control-treated cells with the VIPER algorithm⁹. Finally, drugs were prioritized based on their
236 ability to activate the MR proteins inactivated by SARS-CoV-2 infection, as assessed by their
237 enrichment in proteins differentially activated by each drug, using the aREA algorithm^{8,9} (Fig. 2).

238 ViroTreat predictions were averaged across available GI organoid models and across all
239 evaluated time points. Among the 154 FDA-approved drugs profiled in LoVo cells, ViroTreat
240 prioritized 22 (13 orally available and 9 intravenous) at a highly conservative statistical threshold
241 ($p < 10^{-5}$, Bonferroni corrected (BC)), see Fig. 3 and Supplementary Table 2).

242 **ViroTreat-predicted drugs inhibit SARS-CoV-2 replication**

243 To provide proof-of-concept validation for the ViroTreat predictions in our model, we first assessed
244 drug-mediated inhibition of SARS-CoV-2 replication by ViroTreat-predicted vs. control drugs in
245 the colon adenocarcinoma cell line (Caco-2) known to support SARS-CoV-2 infection³⁰.

246 For this assay, we considered all 13 ViroTreat-inferred orally-available drugs, as a more clinically
247 relevant group, and the top 5 most significant intravenous (IV) drugs. As candidate negative
248 controls, we selected 12 drugs—including 8 orally available agents and 4 IV drugs—not inferred
249 as statistically significant by ViroTreat ($p \geq 0.01$, Fig. 3 and Supplementary Table 2). Caco-2 cells
250 were pre-treated for 24h prior to SARS-CoV-2 infection. Drug concentration was maintained
251 through the entire infection time course and the relative virus replication levels and cell viability
252 were assessed by immunofluorescence staining 24h post-infection (see methods and Fig. 4a).

253 For each drug, the viability-normalized effect on SARS-CoV-2 replication (antiviral effect) was
254 quantified as the log-ratio between viral replication and cell viability reduction relative to vehicle-
255 treated (DMSO) controls (Supplementary Fig. 6). Since multiple concentrations were tested, the
256 lowest concentration corresponding to a significant antiviral effect was reported (Supplementary
257 Table 2). As a proof-of-concept for the ability of this model to identify drugs capable of reducing
258 replication of SARS-CoV-2, we considered drugs to be validated only if their antiviral effect was
259 statistically significant ($FDR < 0.05$) and they induced a decrease in virus replication of at
260 least 20%. This additional condition was used to further increase the stringency when considering
261 the antiviral effect of a drug (see Methods).

262 Of 18 drugs predicted to activate the MR proteins inactivated by SARS-CoV-2 infection, 15 (83%)
263 showed statistically significant antiviral effect. In contrast, none of the 12 drugs selected as
264 potential negative controls showed significant antiviral effect (Fig. 4b and Supplementary Table
265 2), demonstrating a significant enrichment of ViroTreat results in drugs with antiviral activity ($p <$
266 10^{-5} , 1-tailed Fisher's exact test (FET)). Consistently, the Receiver Operating Characteristic
267 (ROC) had an Area Under the Curve $AUC = 0.907$ (95% Confidence Interval: 0.77–0.91), which
268 is highly statistically significant ($p < 10^{-4}$, Fig. 4c), demonstrating the predictive power of ViroTreat
269 in this proof-of-concept.

270 To further assess the pathogen-specific nature of ViroTreat predictions, we tested the ability of
271 the 8 ViroTreat-inferred drugs showing the strongest inhibition of SARS-CoV-2 replication, to
272 inhibit rotavirus replication in Caco-2 cells. Interestingly, none of these drugs significantly impaired
273 rotavirus replication (Supplementary Fig. 7 and Supplementary Table 2), showing that ViroTreat-
274 inferred antiviral effects cannot be attributed to generalized impairment of host cellular functions
275 universally required for viral replication, but rather to activation of host-cell MRs required for the
276 maintenance of a host-cell phenotypic state specifically refractory to SARS-CoV-2 replication.

277 To also assess whether the antiviral activity of ViroTreat-predicted oncology drugs in Caco-2 cells
278 might possibly be attributed to their antineoplastic effects in a cancer cell context, we evaluated
279 the antiviral properties of the top 8 drugs in non-transformed, human GI organoid-derived 2D
280 primary cell cultures. When tested in this more physiologic context, 7 of the 8 assayed drugs,
281 including idarubicin, bosutinib, cyclosporine, bicalutamide, vorinostat, amiodarone and
282 osimertinib, demonstrated significant antiviral effect against SARS-CoV-2 based on our original
283 criteria (FDR < 0.05 and decrease in SARS-CoV-2 replication of at least 20%, Fig. 4d and
284 Supplementary Fig. 7). Except for bicalutamide, which exerted its antiviral effect at a 125-fold
285 higher concentration, all drugs were tested at concentrations comparable to their 48h IC₂₀ in LoVo
286 cells, representing the highest sub-toxic concentration usable for optimal MoA elucidation. These
287 findings suggest that ViroTreat can apply the molecular characterization of a drug's MoA, as
288 obtained by the measured effect of the drug on protein activity levels in tissue lineage-matched,
289 neoplastic cell line models, to prioritize and repurpose drugs with potential antiviral activity in both
290 infected tumor models as well as non-transformed human organoid-derived 2D primary cell
291 cultures.

292 Finally, to test the tissue lineage context-specificity of ViroTreat predictions, we assessed the
293 antiviral effect of the 8 ViroTreat predicted drugs for the GI context showing the strongest inhibition
294 of SARS-CoV-2 infection in Caco-2, in lung adenocarcinoma cell line models (Calu-3 and ACE2-
295 A549). Interestingly, only cyclosporine and osimertinib showed a significant antiviral effect (FDR
296 < 0.05 and ≥ 20% virus replication decrease), while amiodarone, apremilast, bicalutamide,
297 bosutinib, exemestane, and pimozone did not (Supplementary Fig. 8 and Supplementary Table
298 2). These results highlight the relevance of lineage context-specificity when prioritizing drugs with
299 ViroTreat.

300 **DISCUSSION**

301 We report here a model characterizing the regulatory biology of virus-host interaction, in which
302 viral infection induces a phenotypic transition in the host cell toward a state that is promotive of
303 viral replication. We applied Master Regulator (MR) inference analysis ^{9,16} to systematically
304 dissect the transcriptional regulators (MR proteins) hijacked by the virus (Viral CheckPoint) and
305 demonstrated, using a model of SARS-CoV-2 infection in gastrointestinal epithelial cells, that
306 pharmacologically blocking this transition is sufficient to maintain the host cell in a state of
307 “transcriptional contraception” that is adverse to virus replication. We adapted the OncoTreat
308 framework, originally developed to prioritize drugs for cancer ⁸, to identify drugs with concerted
309 activity on the Viral Checkpoint.

310

311 We propose that the approach employed in this model, which we call ViroTreat, can be used as
312 a mechanism-based framework for repurposing drugs, based on their ability to reprogram host
313 cells to a state refractory to virus hijacking. In contrast to previous host cell-centric approaches
314 aimed at targeting single host cell proteins that directly interact with the viral proteome, the
315 ViroTreat model was designed to target the entire MR protein module, whose concerted
316 regulatory activity is responsible for implementing and maintaining a virus replication-permissive
317 transcriptional state in the host cell. Thus, ViroTreat expands the one disease/one target/one
318 drug paradigm to targeting an entire protein module (i.e, *Viral Checkpoint*) based on the accurate
319 assessment of each drug’s proteome-wide MoA, as dissected from perturbational profile data.
320 Such a holistic approach to matching disease dependencies to drug MoA overcomes the inherent
321 limitations of drug repurposing efforts that focus on inhibitors of individual proteins or single
322 pathways to thwart viral replication as part of a host cell-targeting strategy.

323 Viral Checkpoint MR identification requires availability of gene expression signatures of virus-
324 infected cells. Therefore., to avoid model-specific confounding effects and to identify a more
325 universal and reproducible MR signature of viral infection, we performed MR analysis in multiple,
326 complementary cellular models, including both transformed cell lines and normal 3D-organoid
327 cultures representing both airway and GI epithelium lineages. In addition, to avoid confounding
328 effects from a heterogeneous combination of infected and non-infected cells—representing the
329 majority of the cell population—MR analysis was also performed at the single cell level, using
330 SARS-CoV-2 genome mapped reads to unequivocally identify infected cells. Finally, we avoided
331 confounding effects from single cell transcriptional state heterogeneity by comparing each
332 infected cell to a small pool of the closest non-infected cells, based on MR analysis, as controls.
333 Finally, to achieve cell context-specific elucidation of drug MoA, we analyzed drug perturbations
334 in cell lines that recapitulate the biology of the infected cells, based on conservation of their most
335 differentially active/inactive MRs, as previously described ²⁷.

336 The ViroTreat framework prioritizes drugs from a predefined library used to generate
337 perturbational assays. For this proof-of-concept, we maximized the translational potential of drug
338 predictions, by focusing our analysis on FDA-approved drugs used primarily in an oncology
339 setting; with particular emphasis on orally available drugs. However, the approach can be easily
340 extended to explore a much larger library of pharmacological compounds. Moreover, the
341 database of drug context-specific MoA can be generated independently and prior to the
342 identification, isolation and characterization of a viral pathogen of interest, making it readily
343 available for current as well as future pandemics.

344 In addition, while most studies have focused on drugs that act as high affinity inhibitors of target
345 proteins ^{2-6,32,33}, to our knowledge, this is the first study to focus on pharmacologic agents
346 predicted to activate, rather than inhibit, an entire protein module of Master Regulator proteins
347 whose inactivation by the virus was found to be necessary for viral hijack and replication. By

348 inducing drug-mediated reversion of the Viral Checkpoint activity, we successfully reprogrammed
349 host cells to a regulatory state of “viral contraception,” thereby significantly buffering the virus’s
350 ability to hijack the host cell machinery required for its infective cycle.

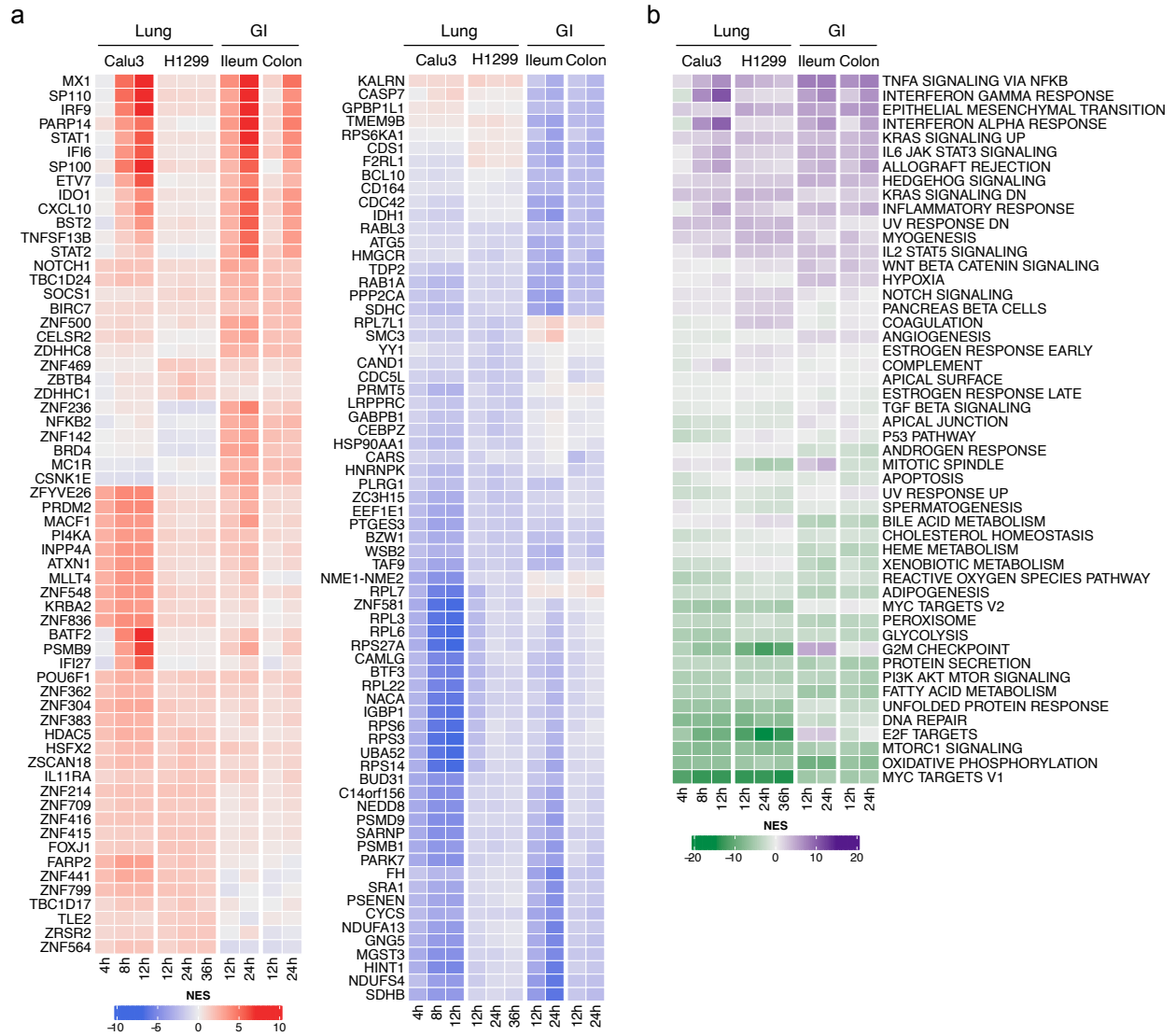
351 Critically, ViroTreat predicted SARS-CoV-2-specific antiviral activity of drugs that have recently
352 emerged as potential host cell-targeting antivirals, in completely unbiased fashion. Among these,
353 cyclosporine ³⁴, amiodarone ³⁵, pimozide ³⁶, mitoxantrone ³⁷, osimertinib ³⁸, bosutinib ³⁹, and
354 bicalutamide ⁴⁰. Moreover, three of the ViroTreat-predicted drugs—cyclosporine (NCT04492891),
355 amiodarone (NCT04351763), and bicalutamide (NCT04509999)—are being evaluated in clinical
356 trials for their safety and efficacy in persons with SARS-CoV-2 infection, suggesting that host cell
357 targeting provides a viable strategy to complement viral-protein targeting drugs.

358 Among the methodological limitations, the most critical one is the need to obtain physiologic
359 models to identify appropriate infection signatures, generate relevant drug perturbational profiles,
360 and validate predicted drugs. In addition, there are also challenges in assessing the optimal
361 concentration at which each compound should be profiled.

362 From a translational perspective, in the setting of both the current and future pandemics, as well
363 as for recurrent epidemics such as those caused by influenza and other viral pathogens, the Viral
364 Checkpoint framework can leverage bulk and single-cell profiles from infected cells to quickly
365 identify the precise set of MR proteins responsible for creating a virus infection-friendly
366 environment in the host cell. Once identified, independent of the specific viral pathogen, potential
367 therapeutic agents can be efficiently prioritized by the ViroTreat model, using readily available—
368 and relatively inexpensive—perturbational databases to elucidate context-specific, proteome-
369 wide drug MoA. Host cell-directed therapies shown to be effective in cell line and organoid models
370 based on such predictions can then undergo rapid validation in more physiologic contexts, prior

371 to testing in human trials designed to evaluate their safety and therapeutic value in the clinical
372 setting.

373



374

375 **Figure 1. Changes in host cell protein activity in response to SARS-CoV-2 virus infection.**

376 **a.** Left, heatmap showing the activity of the top 10 most activated proteins in response to SARS-

377 CoV-2 infection in each of the models and time-points profiled at the single-cell level. Right,

378 heatmap showing the activity of the top 10 most inactivated proteins in response to SARS-CoV-

379 2 infection in each of the models and time-points profiled at the single-cell level. **b.** Heatmap

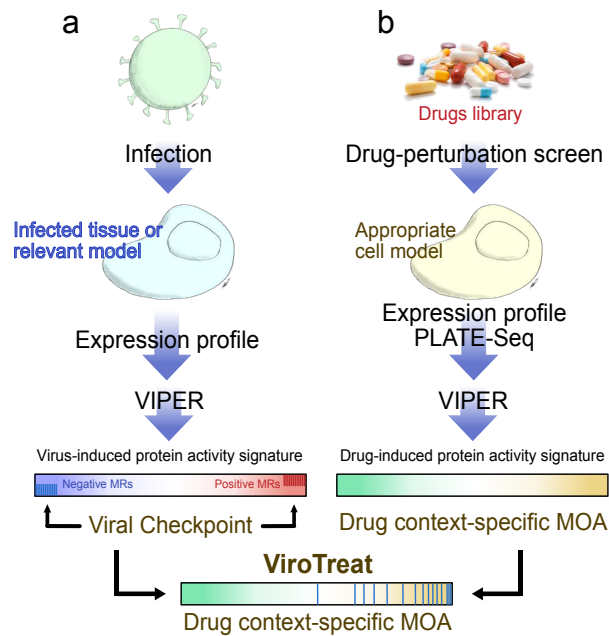
380 showing the enrichment of biological hallmarks in the SARS-CoV-2-induced protein activity

381 signatures. Shown is the Normalized Enrichment Score (NES) estimated by the aREA algorithm,

382 with purple color indicating enrichment in the over activated proteins and green color indicating

383 enrichment in the inactivated proteins.

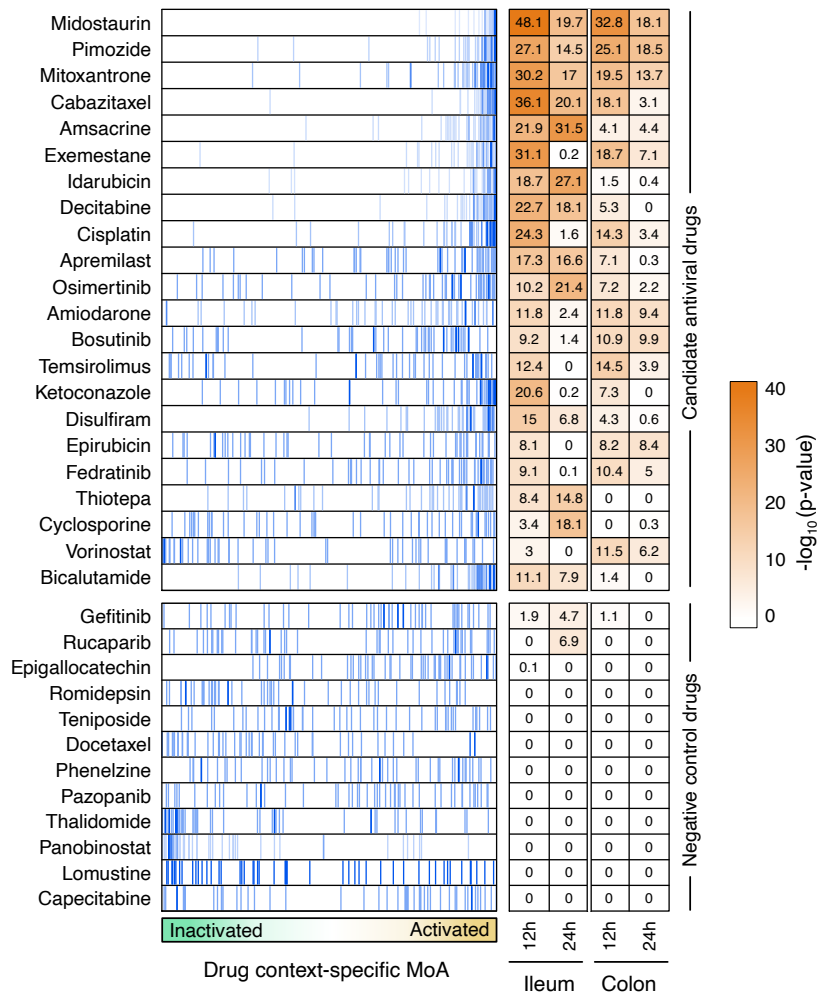
384



385

386 **Figure 2. Schematic representation of the ViroTreat algorithm.** **a.** Virus-induced MR
 387 proteins—the Viral Checkpoint—dissected by VIPER analysis of a gene expression signature,
 388 obtained by comparing an infected tissue or relevant model with non-infected mock controls. **b.**
 389 Context-specific drug MoA database, generated by perturbing an appropriate cell model with
 390 therapeutically relevant drug concentrations, followed by VIPER analysis of the drug-induced
 391 gene expression signatures to infer the drug-induced protein activity signature. ViroTreat
 392 prioritizes drugs able to activate the Viral Checkpoint’s negative MR proteins by quantifying the
 393 enrichment of such proteins on the drugs’ context-specific MoA.

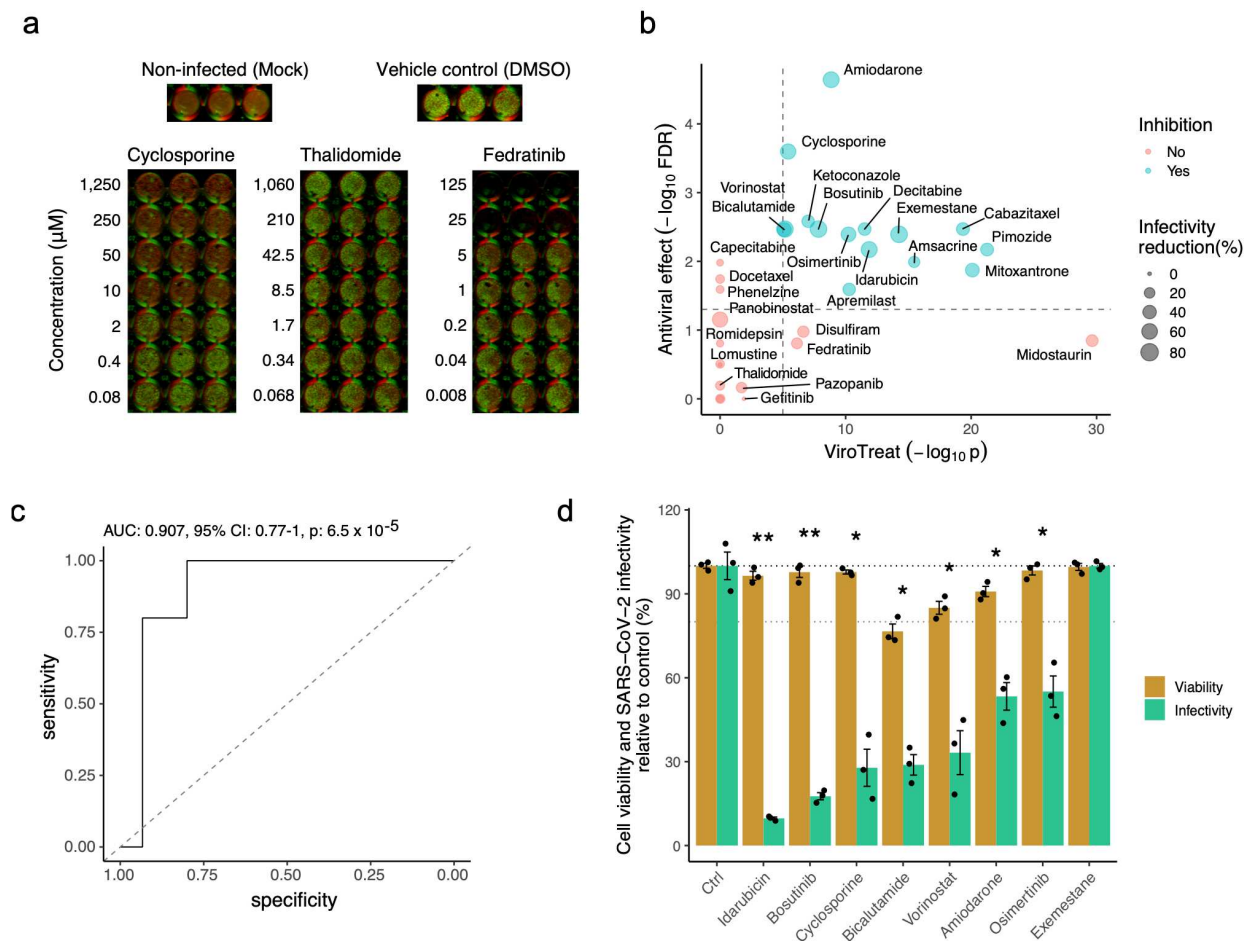
394



395

396 **Figure 3. ViroTreat results for the GI models.** Shown are the enrichment plot for the top 50
 397 most inactivated (blue vertical lines) proteins, in response to SARS-CoV-2 infection (the negative
 398 component of the viral Checkpoint) of the ileum organoid for 12h, on the protein activity signature
 399 induced by the drug perturbations—drug context-specific MoA, represented by the green-orange
 400 color scale in the x-axis—of LoVo colon adenocarcinoma cells. The heatmap shows the
 401 Bonferroni's corrected $-\log_{10}(\text{p-value})$ estimated by ViroTreat. Shown are all the 22 candidate
 402 drugs (ViroTreat $p < 10^{-5}$) and 12 drugs selected as negative controls (ViroTreat $p > 0.01$) in both
 403 ileum and colon-derived organoids at 12 and 24 hours post-infection.

404



405

406 **Figure 4. Experimental validation of ViroTreat predictions.** **a.** Representative
 407 immunofluorescence images of non-infected (Mock) Caco-2 cells, vehicle control (DMSO) treated
 408 and SARS-CoV-2 infected cells, and representative examples of a drug showing significant
 409 antiviral effect (Cyclosporine), of a drug showing non-significant antiviral effect (Thalidomide) and
 410 a drug showing non-significant antiviral effect and cell toxicity (Fedratinib). Drug concentration
 411 (μM) is indicated to the left of the images showing triplicated experiments. Cells were stained with
 412 DNA dye Draq5 (red) and anti-dsRNA antibody (green). **b.** Scatterplot showing the ViroTreat
 413 results (x-axis) compared to the specific antiviral effect (y-axis) experimentally evaluated in Caco-
 414 2 colon adenocarcinoma cells. The vertical and horizontal dashed lines represent the thresholds
 415 for statistical significance for ViroTreat (p-value = 10^{-5} , BC) and specific antiviral effect (FDR =
 416 0.05), respectively. **c.** ROC analysis for the ViroTreat predictions, considering as positive
 417 response a specific antiviral effect at FDR < 0.05 with at least 20% reduction in virus replication.
 418 Estimated AUC, 95% confidence interval (CI) and p-value are indicated in the plots. **d.** Effect of 8
 419 drugs, showing the strongest reduction in SARS-CoV-2 replication in Caco-2 cells, on cell viability

420 and SARS-CoV-2 replication in GI organoid-derived 2D primary cell cultures. Bars indicate the
421 mean \pm SEM. Antiviral effect: * FDR < 0.05, ** FDR < 0.01.

422 **METHODS**

423 **Cells**

424 Vero E6 (ATCC CRL-1586) and Caco-2 (ATCC HTB-37) cells were maintained in DMEM
425 supplemented with 10% fetal bovine serum and 1% penicillin/streptomycin.

426 **GI organoids**

427 Human tissue was received from colon resection from the University Hospital Heidelberg. This
428 study was carried out in accordance with the recommendations of the University Hospital
429 Heidelberg with informed written consent from all subjects in accordance with the Declaration of
430 Helsinki. All samples were received and maintained in an anonymized manner. The protocol was
431 approved by the “Ethics commission of the University Hospital Heidelberg” under the protocol S-
432 443/2017. Stem cells containing crypts were isolated following previously described protocols ⁴¹.
433 Organoids were passaged and maintained in basal and differentiation culture media
434 (Supplementary Table 3) as previously described ⁴¹.

435 **Viruses**

436 SARS-CoV-2 (strain BavPat1) was obtained from the European Virology Archive. The virus was
437 amplified in Vero E6 cells and used at a passage 3 for all experiments as previously described
438 ^{30,42}.

439 **SARS-CoV-2 infection assay**

440 20,000 cells were seeded per well into a 96-well dish 24 hours prior to drug treatment. 100µL of
441 media containing the highest drug concentration was added to the first well. Six serial 1:5 dilutions
442 were made (all samples were performed in triplicate). Drugs were incubated on cells for 24 hours.
443 Prior to infections, fresh drugs were replaced and SARS-CoV-2 at multiplicity of infection (MOI) 3

444 was added to each well. 24 hours post-infection cells were fixed in 4% paraformaldehyde (PFA)
445 for 10 mins at room temperature (RT). PFA was removed and cells were washed twice in 1X PBS
446 and then permeabilized for 10 mins at RT in 0.5% Triton-X. Cells were blocked in a 1:2 dilution of
447 Li-Cor blocking buffer (Li-Cor) for 30 mins at RT. Cells were stained with 1/1000 dilution anti-
448 dsRNA (J2, SCIONS) for 1h at RT as marker of infected cells as previously described³⁰. Cells
449 were washed three times with 0.1% Tween in PBS. Secondary antibody goat anti-mouse IR 800
450 (Thermo) and DNA dye Draq5 (Thermo) were diluted 1/10,000 in blocking buffer and incubated
451 for 1h at RT. Cells were washed three times with 0.1% Tween/PBS. Cells were imaged in 1X PBS
452 on a LICOR imager. Effect of drugs were analyzed by comparing the average fluorescence of
453 mock treated cells to drug treated cells. Draq5 staining was used to determined cell viability.

454 **Rotavirus infection assay**

455 40,000 cells were seeded per well into a collagen-coated 96-well dish 24 hours prior to drug
456 treatment. 100 μ L of media containing the highest drug concentration was added to the first well.
457 Six serial 1:5 dilutions were made (all samples were performed in triplicate). Drugs were incubated
458 on cells for 24 hours. Media was removed and cells were washed 2X with serum-free media and
459 were infected with WT SA11 Rotavirus expressing mKate at MOI 0.1 (calculated in MA104
460 cells) diluted in serum-free media. Rotavirus was previously activated for 30 minutes at 37°C in
461 serum-free media containing 2 μ g/ml trypsin. Infection was allowed to proceed for 1 hour.
462 Following infection, virus was removed and cells were washed 1X with serum-free media. Media
463 containing drugs and 0.5 μ g/ml trypsin were added back to cells to allow for Rotavirus
464 propagation. 24 hours post-infection cells were fixed with 2% PFA for 15 mins and then stained
465 with DAPI. Cells were imaged in 1X PBS on a Cell Discoverer 7 using a 5X objective.
466 Quantifications of infection was carried out by quantifying the number of infected cells (mKate
467 positive cells) in infected and not infected samples using CellProfiler.

468 **SARS-CoV-2 infection of human colon organoids-derived 2D primary cell cultures**

469 Organoids were cultured in 24-well plates in basal medium for 5-7 days following the original
470 protocol of Sato and co-workers ⁴¹. To obtain human colon organoids-derived 2D primary cell
471 cultures, the medium was removed from the 24-well plates, organoids were washed 1X with cold
472 PBS and spun (450g for 5 mins). PBS was removed and organoids were digested with 0.5%
473 Trypsin-EDTA (Life technologies) for 5 mins at 37°C. Digestion was stopped by addition of serum
474 containing medium. Digested-organoids were spun again at 450g for 5 mins and the supernatant
475 was removed and digested organoids were re-suspended in basal media at a ratio of 250 µL
476 media/well (corresponding to approximately 400 organoids per ml). Prior seeding, the 48-well
477 tissue culture plates were coated with 2.5% human collagen in water for 1 h at 37°C. The collagen
478 mixture was removed from the 48-well plate and 250 µL of trypsin-digested organoids
479 (corresponding to about 100 digested organoids) were added to each well. 48 hours post-seeding
480 differentiation media (Supplementary Table 3) was added to cells and 4 days post-differentiation
481 cells were treated with drugs at the indicated concentrations for 2 hours prior to SARS-CoV-2
482 infection. Media containing drugs was removed and 10⁶ focus forming units (FFU) (as determined
483 in Vero cells) of SARS-CoV-2 was added to each well for 1 hour at 37°C. Following 1 hour
484 incubation, virus was removed and fresh differentiation media containing drugs was added to
485 cells. 24 hours post-infection RNA was harvested, and virus replication was monitored by RT-
486 qPCR.

487 **Estimation of the antiviral effect**

488 We define the antiviral effect of a drug as its viability-normalized effect on SARS-CoV-2
489 replication. The antiviral effect was quantified as the log-ratio between virus replication and cell
490 viability reduction relative to vehicle-treated controls. Statistical significance was estimated by
491 Student's t-test for each evaluated drug concentration, and multiple-hypothesis testing due to the

492 multiple evaluated concentrations was corrected using the conservative Bonferroni's method.
493 Multiple hypothesis testing due to multiple evaluated drugs was further corrected by Benjamini-
494 Hochberg False Discovery Rate (FDR).

495 Drugs predicted by ViroTreat were considered validated when showing a significant antiviral effect
496 (FDR < 0.05) and a reduction in virus replication of at least 20%. This additional criterium was
497 used to increase the stringency when evaluating the predictions and the threshold was inferred
498 by fitting a gaussian mixture model (GMM) to the relative replication in response to all evaluated
499 drugs (Supplementary Fig. 9). This analysis identified four groups of drugs—i.e. components of
500 the GMM analysis. The first two groups, based on their mean, showed an average decrease in
501 infectivity of 65% and 30%; the third group showed an average decrease in infectivity close to
502 zero (3.5%); and the fourth group showed an average increased in infectivity of 29%
503 (Supplementary Fig. 9). Based on this analysis, we empirically estimated 20% as a reasonable
504 threshold distinguishing drugs that inhibit viral replication (first and second groups) from drugs
505 that showed no effect or increased replication (third and fourth groups, see Supplementary Fig.
506 9). The GMM analysis was performed using the mixtools package available on CRAN
507 (<https://cran.r-project.org/web/packages/mixtools/index.html>) (Supplementary Fig. 9).

508 **RNA isolation, cDNA, and RT-qPCR**

509 RNA was harvested from cells using RNAeasy RNA extraction kit (Qiagen) as per manufactures
510 instructions. cDNA was made using iSCRIPT reverse transcriptase (BioRad) from 250 ng of total
511 RNA as per manufactures instructions. RT-qPCR was performed using iTaq SYBR green
512 (BioRad) as per manufacturer's instructions, TBP or HPRT1 were used as normalizing genes.
513 See Supplementary Table 4 for primers used.

514 **VIPER analysis of bulk RNA-Seq datasets**

515 The source for all the datasets is listed in Supplementary Table 1. RNA-Seq raw-counts data for
516 Calu-3, H1299 and Caco-2 cell line models were obtained from Gene Expression Omnibus
517 Database (GEO, GSE148729)¹⁸. Raw-counts data for A549 cell line, Normal Human Bronchial
518 Epithelial (NHBE) primary cells, a post-mortem lung tissue sample from a COVID-19 patient and
519 a healthy human lung biopsy were downloaded from GEO (GSE147507)⁴³. Normalized data
520 (Transcript per Kilobase Million, TPM) for lung organoids were downloaded from GEO
521 (GSE160435). Raw-count data was normalized using the variance stabilization transformation
522 (VST) procedure as implemented in the DESeq package from Bioconductor⁴⁴.

523 Differential gene expression signatures for the Wyler's dataset¹⁸ (GSE148729) were computed
524 by comparing the SARS-CoV-2 infected samples against the centroid—i.e. the average
525 expression of each gene—of the closest matched non-infected (mock) samples as identified by
526 unsupervised clustering. Specifically, we first performed K-means cluster analysis of the
527 normalized gene expression profiles. The optimal number of clusters was estimated by silhouette-
528 score analysis as implemented in the “fviz_nbclust” function of the “factoextra” package
529 (<https://cran.r-project.org/web/packages/factoextra/index.html>). Cluster solutions were evaluated
530 from $k=2$ to $k=10$ and the solution with the highest average of silhouette scores was considered
531 as optimal. Based on the optimal cluster solution, we selected as reference for each infected
532 sample the centroid of the mock samples within the same cluster. In cases of clusters constituted
533 by infected samples only, the centroid of the mock samples in the closest cluster were used as
534 reference. Because a two clusters solution was estimated as optimal for all cluster analysis, the
535 other cluster was the trivial closest cluster solution in all cases. Cluster solutions with less than
536 two samples per cluster were considered ineffective. For Calu-3 cell line, we noticed that samples
537 associated to the two series (series-1 and series-2) clustered separately—i.e. samples clustered
538 according to series memberships. To avoid possible batch effects in the analysis, the samples of
539 these two series were re-clustered separately to identify the best matched mock control samples

540 in each series independently. For series-1, the mock samples at 4h and 24h clustered together
541 and were used as reference to compute the differential expression signatures of all the Calu-3
542 SARS-CoV-2 infected samples. For series-2, three mock samples, including one mock sample at
543 4h and two mock samples at 12h clustered together and were used as reference to compute the
544 differential expression signatures for all the Calu-3 SARS-CoV-2 infected samples. Of note, in
545 series-2, one mock sample at 4h (GSM4477923) clustered separately from all the other samples
546 with a silhouette score of zero which indicates no clear cluster assignment. This sample was
547 considered as outlier and excluded from the downstream analysis. For the Caco-2 cell line, the
548 centroid of the 4h mock samples was used as reference to compute the differential expression
549 signatures of the SARS-CoV-2 infected samples at 4h and 12h, while the centroid of 24h mock
550 samples was used as reference to compute the differential expression signatures of the 24h
551 SARS-CoV-2 infected samples. For the H1299 cell line, the centroid of the 4h mock samples was
552 used as reference to compute the differential expression signatures of the SARS-CoV-2 infected
553 samples at 4h and 12h; and the centroid of the 36h mock samples was used as reference to
554 compute the differential expression signatures of the 36h SARS-CoV-2 infected samples.

555 Differential gene expression signatures for the Blanco-Melo's dataset ⁴³ (GSE147507) were
556 computed using the centroid of the matched—i.e. same cell line or primary cells—mock control
557 samples as reference. For the post-mortem human lung sample from a COVID-19 patient, the
558 differential gene expression signature was computed using the healthy human lung biopsy
559 samples as reference.

560 Differential gene expression signatures for the lung organoid sample was computed using as
561 reference its matched mock control sample.

562 The differential activity of 5,734 proteins, including 1,723 transcription factors, 630 co-transcription
563 factors, and 3,381 signaling proteins, was estimated for each of the differential gene expression

564 signatures with the VIPER algorithm ⁹, using matched context-specific models of transcriptional
565 regulation. Lung, colon and rectal adenocarcinoma context-specific models of transcriptional
566 regulation were reverse-engineered, based on 517 lung, 459 colon and 167 rectal
567 adenocarcinoma samples in The Cancer Genome Atlas (TCGA) with the ARACNe algorithm ^{10,45},
568 as discussed in ¹⁶. While, ideally, regulatory networks from non-cancer-related epithelial cells may
569 have been more appropriate, use of cancer-related regulatory networks is justified by the high
570 conservation of protein transcriptional targets in cancer-related and normal cells from the same
571 lineage ¹¹. The regulatory models are available as part of the aracne.networks R package from
572 Bioconductor. Specifically, protein activity signatures in response to SARS-CoV-2 infection of the
573 lung adenocarcinoma cell lines (Calu-3, H1299 and A549), lung organoids and human lung tissue
574 samples were inferred with the VIPER algorithm using the lung adenocarcinoma context-specific
575 network. Protein activity signatures for Caco-2 colorectal carcinoma cell line were estimate with
576 the metaVIPER algorithm ¹³ using the colon and rectal adenocarcinoma context-specific
577 networks.

578 The VIPER-inferred protein activity signatures of infected samples at the same time point in the
579 same cell line were integrated using the Stouffer method ⁴⁶.

580 **VIPER analysis of scRNA-Seq datasets**

581 Single-cell (sc)RNAseq count matrices, based on Unique Molecular Identifiers (UMI), for Calu-3
582 and H1299 lung adenocarcinoma cell lines were downloaded from GEO (GSE148729). Both
583 count matrices were already filtered for low quality cells as described ¹⁸. Count matrices (UMI)
584 from ileum and colon organoids were made available by Boulant lab and are also publicly
585 available on GEO (GSE156760). Count matrices were filtered for low quality cells as described
586 by Triana et al., 2021 ⁴².

587 In contrast to bulk RNASeq profiles, single cell RNASeq profiles (scRNASeq) allow effective
588 identification of the individual cells likely to be infected by the virus, which commonly represent a
589 minority of cells in a culture. For this study, therefore, we defined cells to be infected if they present
590 at least one sequenced read mapped to the SARS-CoV-2 genome. Critically, gene expression
591 signatures based on scRNASeq profiles, as computed by comparing *bona fide* infected cells to
592 non-infected controls, are less affected by contamination and dilution effects characteristic of bulk
593 RNASeq-derived signatures, resulting from a variable proportion of infected vs. non-infected cells.

594 To account for confounding effects and gene expression profile heterogeneity associated with
595 mechanisms that are independent of viral infection^{18,42}—such as cell cycle and the use of models
596 derived from cancer cell lines⁴⁷—differential expression signatures between infected and non-
597 infected single cells were computed by comparing each infected cell to its $k = 50$ closest non-
598 infected ones (Supplementary Fig. 1). This approach significantly improved accuracy and
599 reproducibility of differential gene expression signatures, including across different cell lines, by
600 minimizing confounding effects not associated with viral infection. To identify mock controls cells
601 for each individual infected cell we transformed the count matrices to count per million (CPM) and
602 subsequently to VIPER-inferred protein activity signatures. Briefly, gene expression profiles were
603 transformed to differential gene expression signatures using the “scale” method—i.e. z-score
604 transformation—as implemented in the VIPER package⁹. Then, using lung adenocarcinoma
605 context-specific models of transcriptional regulation, we transformed the single-cell gene
606 expression signature matrices for Calu-3 and H1299 cell lines to VIPER-inferred protein activity
607 signature matrices. Similarly, using colon and rectal adenocarcinoma context-specific networks,
608 we transformed the single-cell gene expression signature matrices for ileum and colon organoids
609 to the corresponding metaVIPER-inferred protein activity signature matrices.

610 The phenotypic state similarity between cells of the same dataset was quantified by the euclidean
611 distance, calculated based on the top 100 principal components of the VIPER-inferred protein

612 activity matrix. Briefly, the Singular Value Decomposition (SVD) was used to estimate the matrix
613 of cells by eigenproteins (principal components), and linear regression analysis was used to
614 identify the components (eigenprotein vectors) significantly associated to the viral infection,
615 expressed as the sum of the normalized UMI viral counts—counts mapping to the SARS-CoV-2
616 genome. For ileum and colon, the vectors of viral counts were generated by summing the
617 normalized counts generated by targeted sequencing analysis ⁴². Principal components
618 significantly associated with infection ($p < 0.05$) were removed from the PCA space. Next, we
619 performed a K-Nearest Neighbors (KNN) analysis in the dimensionally reduced PCA space,
620 considering the top 100 infection-independent principal components, to identify the phenotypically
621 closest 50 mock cells for each of the infected cells. The KNN analysis was performed using the
622 FNN package ⁴⁸. The 50 phenotypically closest mock cells were used as reference to compute
623 the SARS-CoV-2-induced differential gene expression signature for each of the infected cells.
624 Specifically, the differential gene expression signature for each infected cell was estimated by
625 subtracting the mean expression of the 50 phenotypically closest mock cells and dividing by their
626 standard deviation. For Calu-3 and H1299 cell lines, we considered as “SARS-CoV-2-infected”
627 all the cells with at least 1 sequencing read mapping to the SARS-CoV-2 genome. For ileum and
628 colon, we considered as “SCOV2-infected”, all cells identified by targeted sequencing ⁴².

629 The differential gene expression signatures of SARS-CoV-2 infected cells were transformed to
630 inferred protein activity signatures by VIPER and metaVIPER algorithms, as described above.

631 Single-cell protein activity signatures of each data set were integrated by arithmetic mean at each
632 available time point for each cell line.

633 **Similarity of VIPER-inferred protein activity signatures**

634 The conservation of MR proteins between VIPER-inferred protein activity signatures was
635 quantified by the reciprocal enrichment of the top 25 most activated, and the top 25 most

636 inactivated proteins in signature S_1 in proteins differentially active in signature S_2 and vice versa
637 ⁴⁹, as implemented by the *viperSimilarity()* function in the viper package from Bioconductor.

638 **Enrichment of biological hallmarks on SARS-CoV-2 infection-induced protein activity**
639 **signatures**

640 Hallmarks gene sets (v.7.2) were downloaded from the molecular signatures database (MSigDB)
641 website (<http://www.gsea-msigdb.org/gsea/msigdb/collections.jsp>). Enrichment of the MSigDB
642 biological hallmarks protein-sets on the SARS-CoV-2 induced, VIPER-inferred protein activity
643 signatures, with the aREA algorithm ⁹.

644 **Enrichment of Viral Checkpoint MRs on infection essential genes identified by CRISPR**
645 **screens**

646 CRISPR screen results (z-score) were downloaded from the supplementary data of Wei et.al ⁶
647 (Vero-E6 cells) and Schneider et. ⁴ (Huh-7.5 cells). Z-scores were integrated across all
648 experimental conditions for each cell line using the Stouffer's method. Enrichment of the top 50
649 most activated, and the top 50 most inactivated proteins in response to SARS-CoV-2 infection,
650 obtained after integrating (average) all 10 single-cell protein activity signatures, on each CRISPR
651 experiment z-score signature, and on their Stouffer's integration, were estimated by GSEA.
652 Normalized Enrichment Score (NES) and p-value were estimated by permuting the genes in the
653 CRISPR signatures 10,000 times uniformly at random. SARS-CoV-2 inactivated MRs essential
654 for infectivity were identified as the genes in the leading-edge for the GSEA of the inactivated
655 MRs on the integrated CRISPR screen signature.

656 **Enrichment of SARS-CoV-2 interacting protein on host proteins differentially active in**
657 **response to SARS-CoV-2 infection**

658 A list of 332 SARS-CoV-2 interacting proteins was obtained from the supplementary materials of
659 Gordon et al.,². 90 of the 332 interacting proteins were represented among the regulatory proteins
660 for which we could infer their activity. Enrichment analysis of this 90 SARS-CoV-2 interacting
661 proteins on the VIPER-inferred protein activity signatures was performed by GSEA. NES and p-
662 values were estimated by permuting the VIPER-inferred protein activity signatures 10,000 times
663 uniformly at random.

664 **ViroTreat analysis**

665 We have previously shown that tumor checkpoints can be pharmacologically switched, either off
666 ^{8,12,17,50,51} or on ¹⁶, leading to their collapse and loss of viability or gain of associated functional
667 properties, respectively. This observation was instrumental for the development and validation of
668 the NY CLIA certified, VIPER-based methodology OncoTreat, for the prioritization of small
669 molecule compounds that can either inactivate or activate a tumor checkpoint on a sample-by-
670 sample basis, with critical applications in precision oncology⁸. Based on the successful outcomes
671 observed with OncoTreat when evaluated in the context of tumor suppression, we sought to
672 develop a novel, analogous algorithm, ViroTreat, to identify small molecule compounds capable
673 of suppressing viral infection by targeting the Viral Checkpoint module. Similar to its use in cancer,
674 ViroTreat systematically assesses and prioritizes a small-molecule compound's ability to reverse
675 the activity of a set of MR proteins based on large-scale drug perturbation assays in cell lines that
676 recapitulate (a) the regulatory model of the target cellular population and (b) the activity of MR
677 proteins. Specifically, perturbational assay data are comprised of RNASeq profiles generated at
678 24h (by PLATE-Seq assays³¹), following treatment of MR-matched cell lines with a library of FDA-
679 approved and late-stage experimental drugs (in Phase 2 and 3 clinical trials) and DMSO as
680 control. These profiles are then used to assess the differential activity of relevant MRs in drug vs.
681 DMSO-treated cells. Finally, enrichment of MR proteins in proteins whose activity has been
682 inverted by the drug is computed by protein set enrichment analysis (PSEA) using the aREA

683 algorithm^{8,52}. The RNASeq profiles used for ViroTreat analysis were generated at 24h following
684 treatment of LoVo cells with a repertoire of 154 FDA-approved oncology drugs. Perturbations
685 were performed at each drug's highest sublethal concentration (48h IC₂₀) or maximum serum
686 concentration (C_{max}) at its Maximum Tolerated Dose (MTD), whichever was lower. This was done
687 to prevent confounding effects, unrelated to the drug MoA, resulting from cell death or stress
688 pathway activation. RNASeq data was generated using PLATE-Seq, a fully automated, 96-well
689 based assay³¹ (Supplementary Table 2).

690 REFERENCES

- 691 1 Bojkova, D. *et al.* Proteomics of SARS-CoV-2-infected host cells reveals therapy targets.
692 *Nature* **583**, 469-472, doi:10.1038/s41586-020-2332-7 (2020).
- 693 2 Gordon, D. E. *et al.* A SARS-CoV-2 protein interaction map reveals targets for drug
694 repurposing. *Nature* **583**, 459-468, doi:10.1038/s41586-020-2286-9 (2020).
- 695 3 Daniloski, Z. *et al.* Identification of Required Host Factors for SARS-CoV-2 Infection in
696 Human Cells. *Cell* **184**, 92-105 e116, doi:10.1016/j.cell.2020.10.030 (2021).
- 697 4 Schneider, W. M. *et al.* Genome-Scale Identification of SARS-CoV-2 and Pan-coronavirus
698 Host Factor Networks. *Cell* **184**, 120-132 e114, doi:10.1016/j.cell.2020.12.006 (2021).
- 699 5 Wang, R. *et al.* Genetic Screens Identify Host Factors for SARS-CoV-2 and Common Cold
700 Coronaviruses. *Cell* **184**, 106-119 e114, doi:10.1016/j.cell.2020.12.004 (2021).
- 701 6 Wei, J. *et al.* Genome-wide CRISPR Screens Reveal Host Factors Critical for SARS-CoV-2
702 Infection. *Cell* **184**, 76-91 e13, doi:10.1016/j.cell.2020.10.028 (2021).
- 703 7 Califano, A. & Alvarez, M. J. The recurrent architecture of tumour initiation, progression
704 and drug sensitivity. *Nat Rev Cancer* **17**, 116-130, doi:10.1038/nrc.2016.124 (2017).
- 705 8 Alvarez, M. J. *et al.* A precision oncology approach to the pharmacological targeting of
706 mechanistic dependencies in neuroendocrine tumors. *Nat Genet* **50**, 979-989,
707 doi:10.1038/s41588-018-0138-4 (2018).
- 708 9 Alvarez, M. J. *et al.* Functional characterization of somatic mutations in cancer using
709 network-based inference of protein activity. *Nat Genet* **48**, 838-847, doi:10.1038/ng.3593
710 (2016).

- 711 10 Basso, K. *et al.* Reverse engineering of regulatory networks in human B cells. *Nat Genet*
712 **37**, 382-390, doi:10.1038/ng1532 (2005).
- 713 11 Lefebvre, C. *et al.* A human B-cell interactome identifies MYB and FOXM1 as master
714 regulators of proliferation in germinal centers. *Mol Syst Biol* **6**, 377,
715 doi:10.1038/msb.2010.31 (2010).
- 716 12 Carro, M. S. *et al.* The transcriptional network for mesenchymal transformation of brain
717 tumours. *Nature* **463**, 318-325, doi:10.1038/nature08712 (2010).
- 718 13 Ding, H. *et al.* Quantitative assessment of protein activity in orphan tissues and single cells
719 using the metaVIPER algorithm. *Nat Commun* **9**, 1471, doi:10.1038/s41467-018-03843-3
720 (2018).
- 721 14 Obradovic, A. *et al.* Single-cell protein activity analysis identifies recurrence-associated
722 renal tumor macrophages. *Cell* **184**, 2988-3005 e2916, doi:10.1016/j.cell.2021.04.038
723 (2021).
- 724 15 Elyada, E. *et al.* Cross-Species Single-Cell Analysis of Pancreatic Ductal Adenocarcinoma
725 Reveals Antigen-Presenting Cancer-Associated Fibroblasts. *Cancer Discov* **9**, 1102-1123,
726 doi:10.1158/2159-8290.CD-19-0094 (2019).
- 727 16 Paull, E. O. *et al.* A modular master regulator landscape controls cancer transcriptional
728 identity. *Cell* **184**, 334-351 e320, doi:10.1016/j.cell.2020.11.045 (2021).
- 729 17 Aytes, A. *et al.* Cross-species regulatory network analysis identifies a synergistic
730 interaction between FOXM1 and CENPF that drives prostate cancer malignancy. *Cancer*
731 *Cell* **25**, 638-651, doi:10.1016/j.ccr.2014.03.017 (2014).
- 732 18 Wyler, E. *et al.* Transcriptomic profiling of SARS-CoV-2 infected human cell lines identifies
733 HSP90 as target for COVID-19 therapy. *iScience* **24**, 102151,
734 doi:10.1016/j.isci.2021.102151 (2021).
- 735 19 Triana, S. *et al.* Single-cell analyses reveal SARS-CoV-2 interference with intrinsic immune
736 response in the human gut. *bioRxiv*, 2020.2010.2021.348854,
737 doi:10.1101/2020.10.21.348854 (2020).
- 738 20 Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for
739 interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* **102**, 15545-15550,
740 doi:10.1073/pnas.0506580102 (2005).
- 741 21 Verhelst, J., Hulpiau, P. & Saelens, X. Mx proteins: antiviral gatekeepers that restrain the
742 uninvited. *Microbiol Mol Biol Rev* **77**, 551-566, doi:10.1128/MMBR.00024-13 (2013).
- 743 22 Ivashkiv, L. B. & Donlin, L. T. Regulation of type I interferon responses. *Nat Rev Immunol*
744 **14**, 36-49, doi:10.1038/nri3581 (2014).

- 745 23 de Wilde, A. H., Snijder, E. J., Kikkert, M. & van Hemert, M. J. Host Factors in Coronavirus
746 Replication. *Curr Top Microbiol Immunol* **419**, 1-42, doi:10.1007/82_2017_25 (2018).
- 747 24 Walsh, L. A. *et al.* An Integrated Systems Biology Approach Identifies TRIM25 as a Key
748 Determinant of Breast Cancer Metastasis. *Cell Rep* **20**, 1623-1640,
749 doi:10.1016/j.celrep.2017.07.052 (2017).
- 750 25 Kobayashi, M. *et al.* The ubiquitin hybrid gene UBA52 regulates ubiquitination of
751 ribosome and sustains embryonic development. *Sci Rep* **6**, 36780, doi:10.1038/srep36780
752 (2016).
- 753 26 Thaker, S. K., Ch'ng, J. & Christofk, H. R. Viral hijacking of cellular metabolism. *BMC Biol*
754 **17**, 59, doi:10.1186/s12915-019-0678-9 (2019).
- 755 27 Alvarez, M. J. *et al.* Unbiased Assessment of H-STS cells as high-fidelity models for gastro-
756 enteropancreatic neuroendocrine tumor drug mechanism of action analysis. *bioRxiv*,
757 677435, doi:10.1101/677435 (2019).
- 758 28 Alvarez, M. J. *et al.* Reply to 'H-STS, L-STS and KRJ-I are not authentic GEPNET cell lines'.
759 *Nat Genet* **51**, 1427-1428, doi:10.1038/s41588-019-0509-5 (2019).
- 760 29 Douglass, E. F. *et al.* A Community Challenge for Pancancer Drug Mechanism of Action
761 Inference from Perturbational Profile Data. *bioRxiv*, 2020.2012.2021.423514,
762 doi:10.1101/2020.12.21.423514 (2020).
- 763 30 Stanifer, M. L. *et al.* Critical Role of Type III Interferon in Controlling SARS-CoV-2 Infection
764 in Human Intestinal Epithelial Cells. *Cell Rep* **32**, 107863,
765 doi:10.1016/j.celrep.2020.107863 (2020).
- 766 31 Bush, E. C. *et al.* PLATE-Seq for genome-wide regulatory network analysis of high-
767 throughput screens. *Nat Commun* **8**, 105, doi:10.1038/s41467-017-00136-z (2017).
- 768 32 Stukalov, A. *et al.* Multilevel proteomics reveals host perturbations by SARS-CoV-2 and
769 SARS-CoV. *Nature*, doi:10.1038/s41586-021-03493-4 (2021).
- 770 33 Bouhaddou, M. *et al.* The Global Phosphorylation Landscape of SARS-CoV-2 Infection. *Cell*
771 **182**, 685-712 e619, doi:10.1016/j.cell.2020.06.034 (2020).
- 772 34 Poulsen, N. N., von Brunn, A., Hornum, M. & Blomberg Jensen, M. Cyclosporine and
773 COVID-19: Risk or favorable? *Am J Transplant* **20**, 2975-2982, doi:10.1111/ajt.16250
774 (2020).
- 775 35 Sanchis-Gomar, F. *et al.* Amiodarone in the COVID-19 Era: Treatment for Symptomatic
776 Patients Only, or Drug to Prevent Infection? *Am J Cardiovasc Drugs* **20**, 413-418,
777 doi:10.1007/s40256-020-00429-7 (2020).

778 36 Vatansever, E. C. *et al.* Bepridil is potent against SARS-CoV-2 in vitro. *Proc Natl Acad Sci U*
779 *S A* **118**, doi:10.1073/pnas.2012201118 (2021).

780 37 Zhang, Q. *et al.* Heparan sulfate assists SARS-CoV-2 in cell entry and can be targeted by
781 approved drugs in vitro. *Cell Discov* **6**, 80, doi:10.1038/s41421-020-00222-5 (2020).

782 38 Chen, C. Z. *et al.* Identifying SARS-CoV-2 Entry Inhibitors through Drug Repurposing
783 Screens of SARS-S and MERS-S Pseudotyped Particles. *ACS Pharmacol Transl Sci* **3**, 1165-
784 1175, doi:10.1021/acscptsci.0c00112 (2020).

785 39 Yang, L. *et al.* Identification of SARS-CoV-2 entry inhibitors among already approved drugs.
786 *Acta Pharmacol Sin*, doi:10.1038/s41401-020-00556-6 (2020).

787 40 Costa, B. & Vale, N. A Review of Repurposed Cancer Drugs in Clinical Trials for Potential
788 Treatment of COVID-19. *Pharmaceutics* **13**, doi:10.3390/pharmaceutics13060815 (2021).

789 41 Sato, T. *et al.* Long-term expansion of epithelial organoids from human colon, adenoma,
790 adenocarcinoma, and Barrett's epithelium. *Gastroenterology* **141**, 1762-1772,
791 doi:10.1053/j.gastro.2011.07.050 (2011).

792 42 Triana, S. *et al.* Single-cell analyses reveal SARS-CoV-2 interference with intrinsic immune
793 response in the human gut. *Mol Syst Biol* **17**, e10232, doi:10.15252/msb.202110232
794 (2021).

795 43 Blanco-Melo, D. *et al.* Imbalanced Host Response to SARS-CoV-2 Drives Development of
796 COVID-19. *Cell* **181**, 1036-1045 e1039, doi:10.1016/j.cell.2020.04.026 (2020).

797 44 Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome*
798 *Biol* **11**, R106, doi:10.1186/gb-2010-11-10-r106 (2010).

799 45 Lachmann, A., Giorgi, F. M., Lopez, G. & Califano, A. ARACNe-AP: gene network reverse
800 engineering through adaptive partitioning inference of mutual information.
801 *Bioinformatics* **32**, 2233-2235, doi:10.1093/bioinformatics/btw216 (2016).

802 46 Stouffer, S. A., Suchman, E. A., DeVinney, L. C., Star, S. A. & Williams, R. M. J. *The American*
803 *soldier: Adjustment during Army life*. Vol. Vol. I (NJ: Princeton University Press, 1949).

804 47 Laise, P. *et al.* Pancreatic Ductal Adenocarcinoma Comprises Coexisting Regulatory States
805 with both Common and Distinct Dependencies. *bioRxiv*, 2020.2010.2027.357269,
806 doi:10.1101/2020.10.27.357269 (2020).

807 48 Beygelzimer, A. *et al.* *FNN: Fast Nearest Neighbor Search Algorithms and Applications*,
808 <<https://CRAN.R-project.org/package=FNN>> (2019).

809 49 Kruithof-de Julio, M. *et al.* Regulation of extra-embryonic endoderm stem cell
810 differentiation by Nodal and Cripto signaling. *Development* **138**, 3885-3895,
811 doi:10.1242/dev.065656 (2011).

812 50 Rajbhandari, P. *et al.* Cross-Cohort Analysis Identifies a TEAD4-MYCN Positive Feedback
813 Loop as the Core Regulatory Element of High-Risk Neuroblastoma. *Cancer Discov* **8**, 582-
814 599, doi:10.1158/2159-8290.CD-16-0861 (2018).

815 51 Alexa, A., Rahnenfuhrer, J. & Lengauer, T. Improved scoring of functional groups from
816 gene expression data by decorrelating GO graph structure. *Bioinformatics* **22**, 1600-1607,
817 doi:10.1093/bioinformatics/btl140 (2006).

818 52 Shen, Y. *et al.* Systematic, network-based characterization of therapeutic target inhibitors.
819 *PLoS Comput Biol* **13**, e1005599, doi:10.1371/journal.pcbi.1005599 (2017).

820

821 **Acknowledgements** We would like to thank Dr. Vibor Laketa and Dr. Sylvia Olberg for their
822 support through the Infectious Diseases Imaging Platform (IDIP) and Tatiana Alvarez for original
823 artwork. This research was supported by the following NIH grants to A.C.: R35 CA197745
824 (Outstanding Investigator Award); U01 CA217858 (Cancer Target Discovery and Development);
825 S10 OD012351 and S10 OD021764 (Shared Instrument Grants); by grants to S.B.: Deutsche
826 Forschungsgemeinschaft (DFG) project numbers 415089553 (Heisenberg program), 240245660
827 (SFB1129), 278001972 (TRR186), and 272983813 (TRR179), and from the state of Baden
828 Wuerttemberg (AZ: 33.7533.-6-21/5/1) and the Bundesministerium Bildung und Forschung
829 (BMBF) (01KI20198A) and within the Network University Medicine - Organo-Strat COVID-19; by
830 grants to M.L.S.: BMBF (01KI20239B) and DFG project 416072091; by grant to T.A.: ERC
831 Consolidator grant METACELL (grant number 773089); by grant to M.K.J.: BCPM grant to
832 Thomas Geiser (Department of Pulmonary Medicine, University Hospital/DBMR), and support
833 from the Department of Urology of the Bern University Hospital; and to M.J.A.: research support
834 from Karyopharm Therapeutics, Inc.

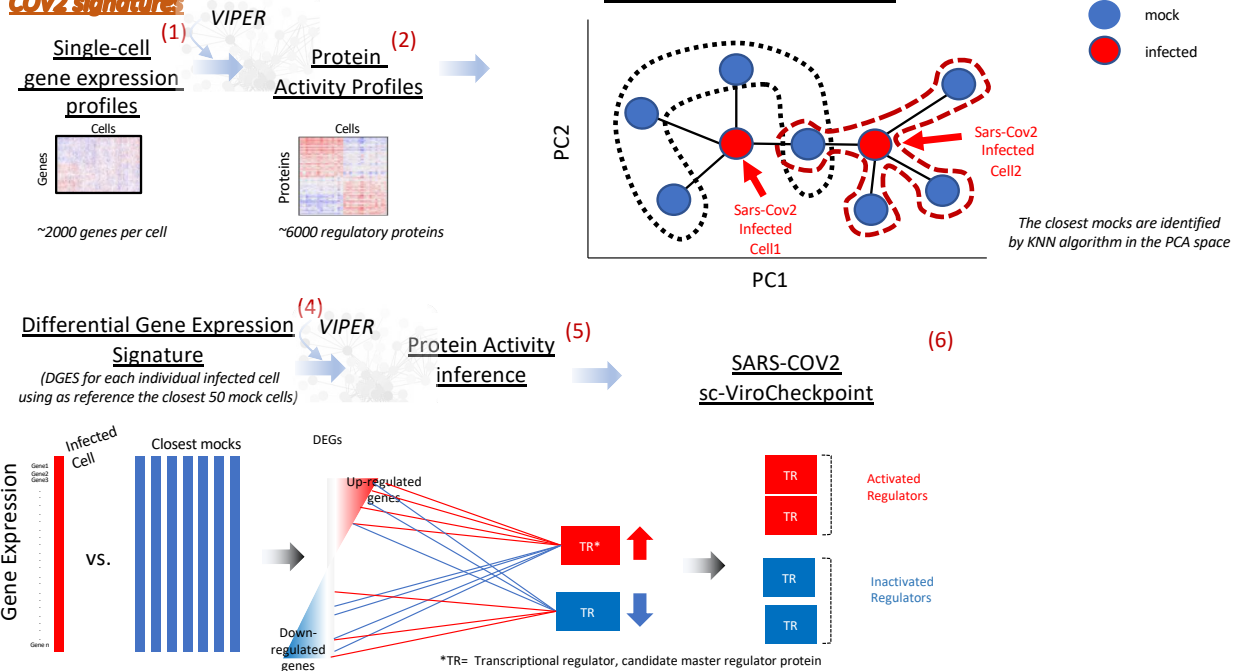
835 **Author Contributions** P.L., G.B., M.K.J., S.B., A.C., and M.J.A. conceived this work. M.L.S.,
836 S.T., P.D., T.A., F.L.M, and M.D.M performed experiments. C.K., R.B.R. and S.P. performed
837 experiments and generated the drugs' perturbational data. P.L., X.S., and M.J.A. performed
838 analysis. G.B., C.K., T.A., M.K.J, A.C., S.B., and M.J.A. supervised experiments and data
839 analysis. P.L., M.L.S., G.B., A.C., S.B., and M.J.A. wrote the manuscript. P.L., M.L.S., G.B.,
840 M.K.J., A.C., S.B., and M.J.A. reviewed the manuscript. All authors approved the final manuscript.

841 **Competing Financial Interest Statement** P.L. is Director of Single-Cell Systems Biology at
842 DarwinHealth, Inc., a company that has licensed some of the algorithms used in this manuscript
843 from Columbia University. G.B. is founder, CEO and equity holder of DarwinHealth, Inc. X.S. is
844 Senior Computational Biologist at DarwinHealth, Inc. A.C. is founder, equity holder, and
845 consultant of DarwinHealth Inc. M.J.A. is CSO and equity holder of DarwinHealth, Inc. Columbia
846 University is also an equity holder in DarwinHealth Inc.

847 **Correspondence and requests for materials** should be addressed to A.C., S.B. or M.J.A.

848

Schematic workflow for the Identification of SARS-COV2 signatures

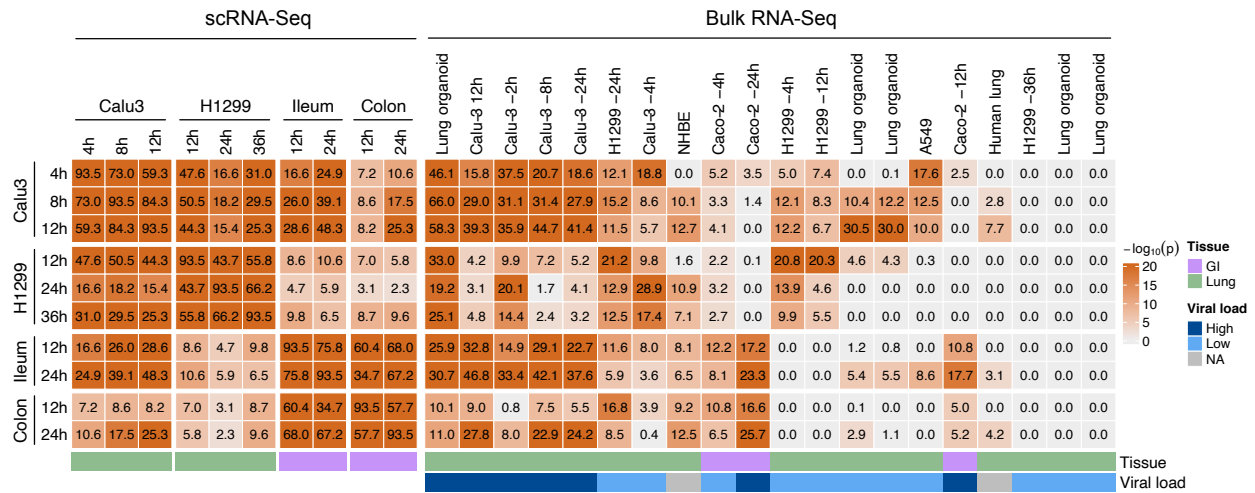


849

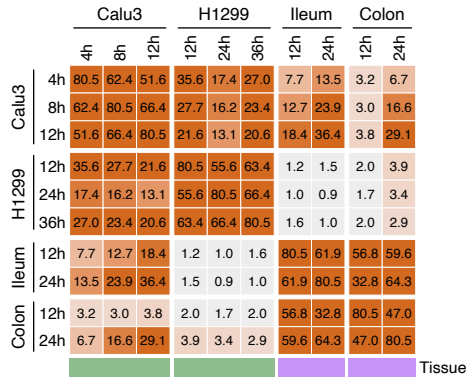
850 **Supplementary Figure 1. Diagram showing the workflow used to compute the protein**
 851 **activity signatures induced by SARS-CoV-2 infection from scRNA-Seq data. Related to**
 852 **Figure 2 and methods.** Normalized single-cell gene expression profiles for all cells of the same
 853 model (i.e. Calu3, H1299, colon and ileum) were transformed to differential gene expression
 854 signatures by applying the z-score procedure. Single-cell differential gene expression signatures
 855 were then transformed to protein activity profiles by applying the VIPER algorithm with context-
 856 specific regulatory networks. A principal component analysis (PCA) was performed on these
 857 VIPER-inferred protein activity profiles. For each infected cell the closest 50 mock cells in the
 858 PCA space were selected as reference to compute a SARS-CoV-2 induced differential gene
 859 expression signature. The VIPER algorithm was then applied to these SARS-CoV-2 induced
 860 differential gene expression signatures to infer SARS-CoV-2 induced protein activity signatures.

861

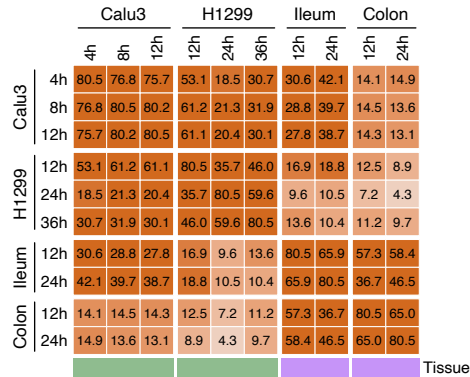
a Top 25 most inactivated and top 25 most activated proteins



b Top 50 most inactivated proteins



c Top 50 most activated proteins



862

863 **Supplementary Figure 2. Conservation of VIPER-inferred Viral Checkpoint. Related to**

864 **Figure 2. a.** Heatmap showing the conservation across single-cell and bulk-tissue samples.

865 Results are expressed as $-\log_{10}(\text{p-value})$, estimated by the reciprocal enrichment of the 25 most

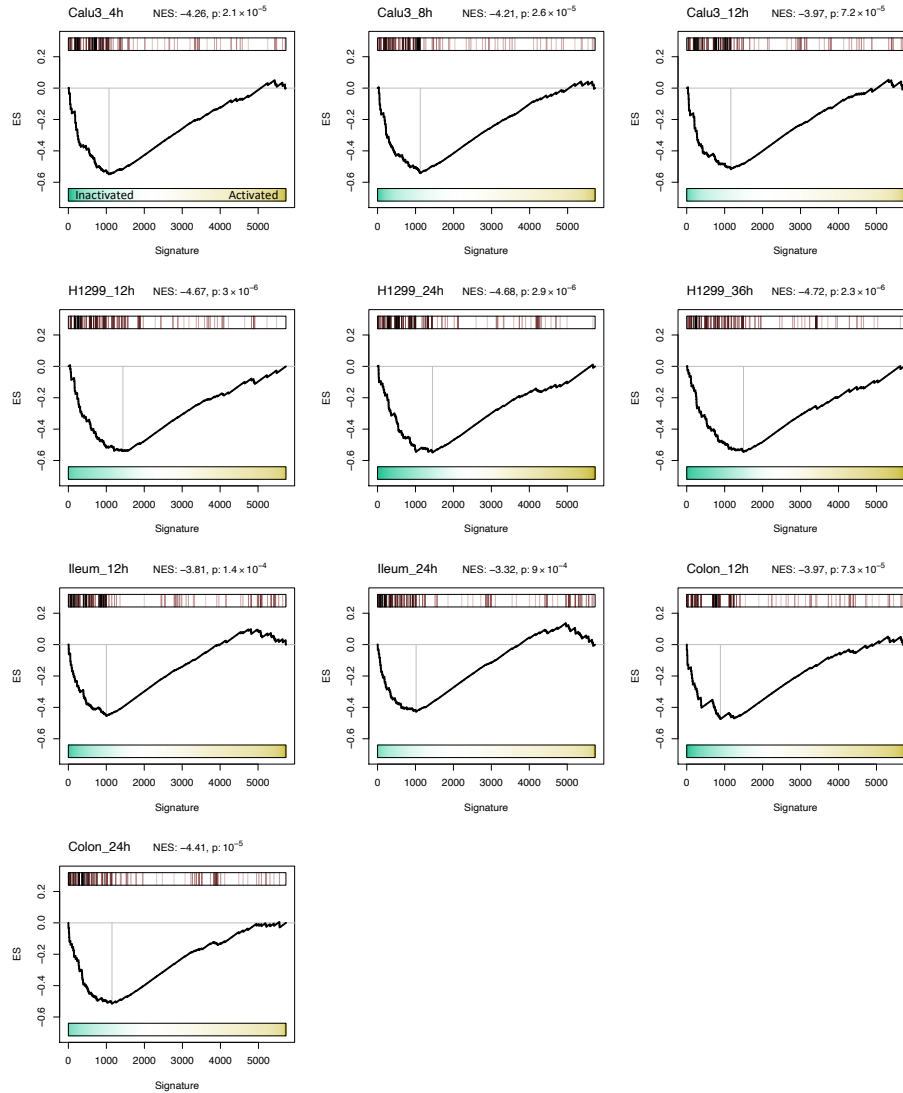
866 activated and 25 most inactivated proteins in each signature using the aREA algorithm as

867 implemented in the viperSimilarity function of the VIPER package. **b-c.** Conservation specifically

868 for the top 50 most activated proteins (b) and most inactivated proteins (c) in response to SARS-

869 CoV-2 infection between time points and models profiled at the single-cell level.

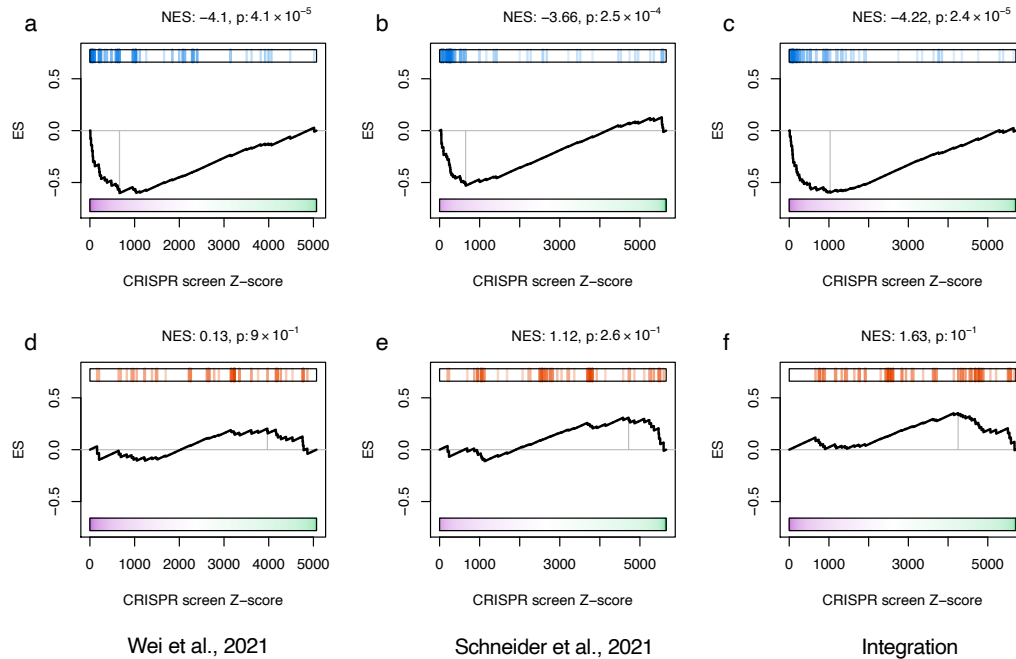
870



871

872 **Supplementary Figure 3. Enrichment of host factors known to physically interact with**
 873 **SARS-CoV-2 proteins on the host proteins differentially active in response to viral**
 874 **infection. Related to Figure 2. GSEA showing the enrichment for the SARS-CoV-2 interacting**
 875 **proteins in the individual SARS-CoV-2 induced protein activity signatures. NES and p-values were**
 876 **estimated by one-tailed test and 1,000 permutations.**

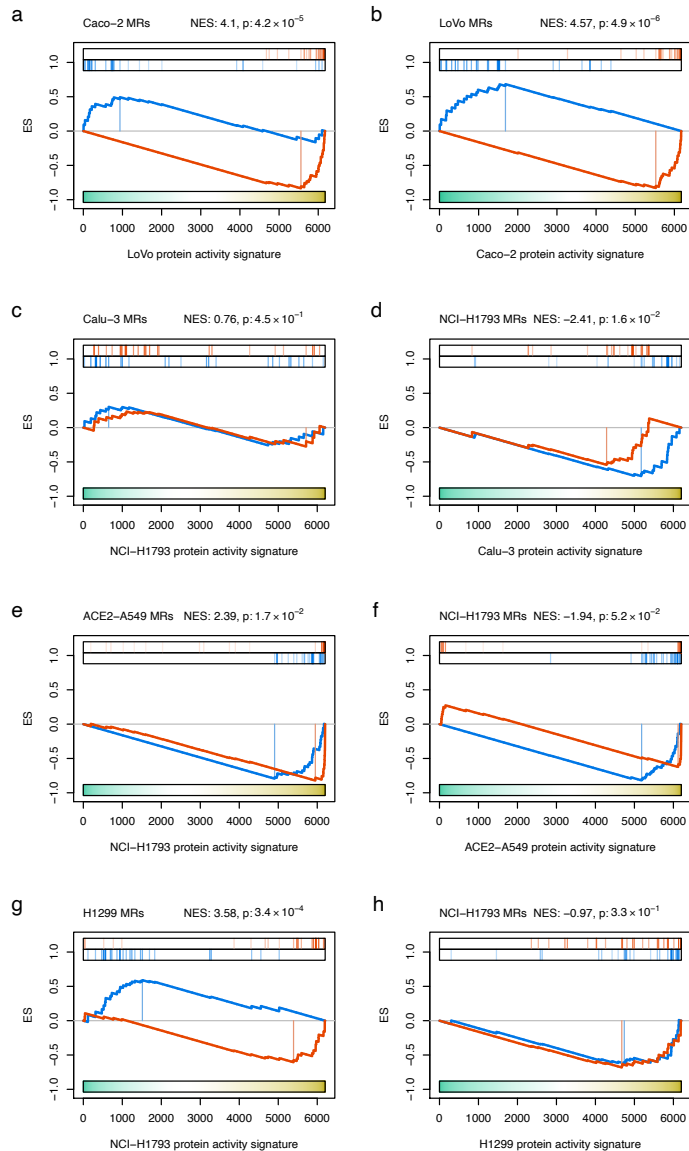
877



878

879 **Supplementary Figure 4. Enrichment of candidate SARS-CoV-2 infection MR proteins on**
 880 **host factors essential for SARS-CoV-2 infectivity. Related to Figure 2.** GSEA showing the
 881 enrichment of the top 50 most inactivated proteins in response to SARS-CoV-2 infection
 882 (inactivated candidate MR proteins) on the antiviral essential genes (a-c), but no enrichment of
 883 the top 50 most activated proteins in response to SARS-CoV-2 infection (activated candidate MR
 884 proteins) on the pro-viral essential genes (d-f), identified by 2 CRISPR screens (a, b, d and e) and
 885 their integration (c and f).

886

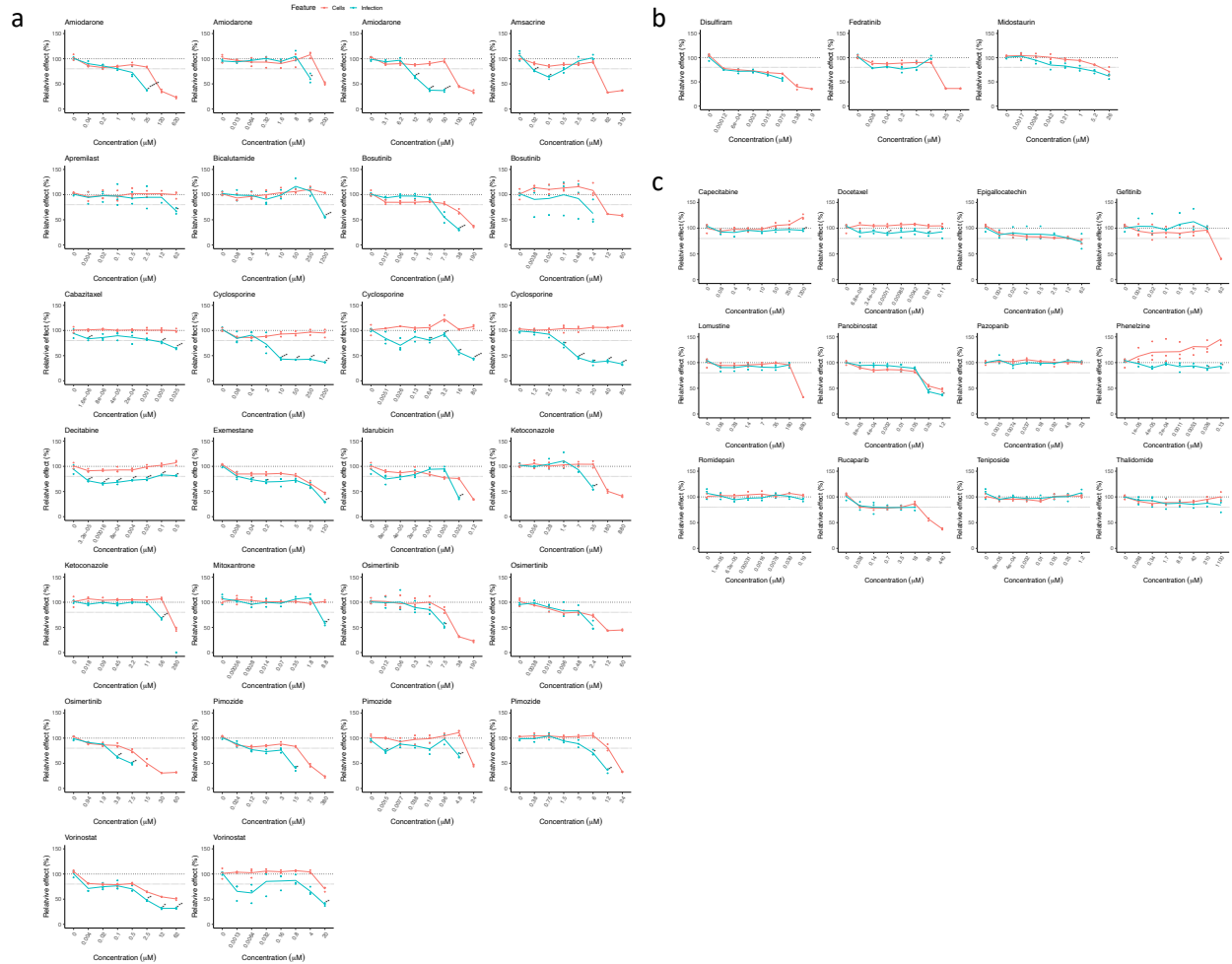


887

888 **Supplementary Figure 5. Conserved activity of MR proteins between cell line models**
 889 **susceptible to SARS-CoV-2 infection (Caco-2, Calu-3, ACE2-A549 and H1299) and the**
 890 **lineage context-matched cell lines included in the drug perturbation PANACEA resource**
 891 **(LoVo and NCI-H1793). Related to Figure 2-3. a. GSEA for the enrichment of the Caco-2 top**
 892 **25 most activated and top 25 most inactivated proteins in the LoVo protein activity signature. b.**
 893 **GSEA for the enrichment of the LoVo top 25 most activated and top 25 most inactivated proteins**
 894 **in the Caco-2 protein activity signature. c. GSEA for the enrichment of the Calu-3 top 25 most**
 895 **activated and top 25 most inactivated proteins in the NCI-H1793 protein activity signature. d.**
 896 **GSEA for the enrichment of the NCI-H1793 top 25 most activated and top 25 most inactivated**
 897 **proteins in the Calu-3 protein activity signature. e. GSEA for the enrichment of the ACE2-A549**

898 top 25 most activated and top 25 most inactivated proteins in the NCI-H1793 protein activity
899 signature. **f.** GSEA for the enrichment of the NCI-H1793 top 25 most activated and top 25 most
900 inactivated proteins in the ACE2-A549protein activity signature. **g.** GSEA for the enrichment of
901 the H1299 top 25 most activated and top 25 most inactivated proteins in the NCI-H1793 protein
902 activity signature. **h.** GSEA for the enrichment of the NCI-H1793 top 25 most activated and top
903 25 most inactivated proteins in the H1299 protein activity signature. Normalized enrichment score
904 (NES) and p-value were estimated by two-tailed test and 1,000 permutations.

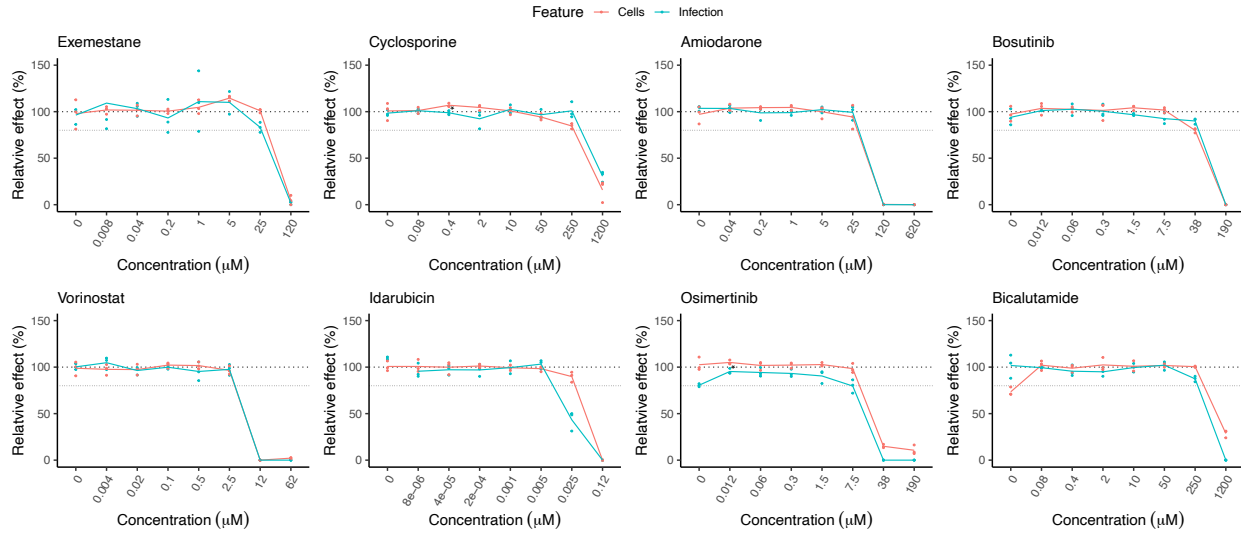
905



906

907 **Supplementary Figure 6. Experimental evaluation of the antiviral effect of FDA-approved**
 908 **drugs in Caco-2 cells. Related to Figure 4 and Supplementary Table 2. a.** 15 of the 18 drugs
 909 predicted by ViroTreat showing significant antiviral effect (FDR < 0.05 and $\geq 20\%$ viral replication
 910 decrease). **b.** 3 of the 18 drugs predicted by ViroTreat showing no significant antiviral effect. **c.**
 911 12 drugs not significant by ViroTreat ($p \geq 0.01$) selected as putative negative controls. The scatter-
 912 plots show the effect of each drug—SARS-CoV-2 replication shown in cyan and cell viability in
 913 red—relative to vehicle control (y-axis), assayed at different concentrations (x-axis) in triplicate.
 914 The lines indicate the average across replicates. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 10^{-4}$,
 915 ***** $p < 10^{-6}$, 1-tailed Student's t-test, BC.

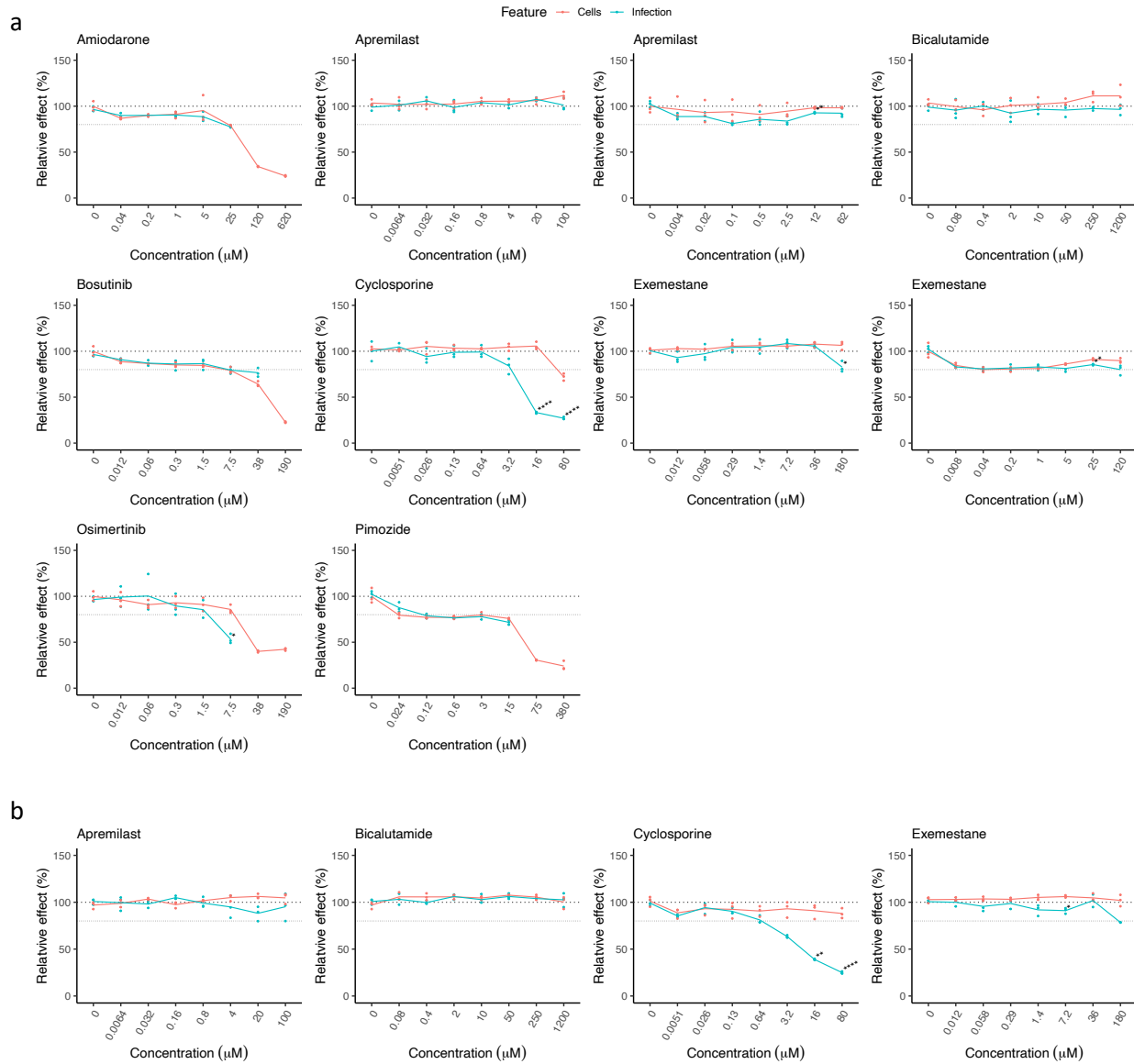
916



917

918 **Supplementary Figure 7. Experimental evaluation of 8 drugs, predicted by ViroTreat and**
 919 **showing the strongest SARS-CoV-2 antiviral effect in Caco-2 cells, for their effect on**
 920 **rotavirus replication. Related to Figure 4 and Supplementary Table 2.** The scatter-plots show
 921 the effect of each drug—rotavirus replication shown in cyan and cell viability in red—relative to
 922 vehicle control (y-axis), assayed at different concentrations (x-axis) in triplicate. The lines indicate
 923 the average across replicates. * $p < 0.05$, 1-tailed Student's t-test, BC.

924

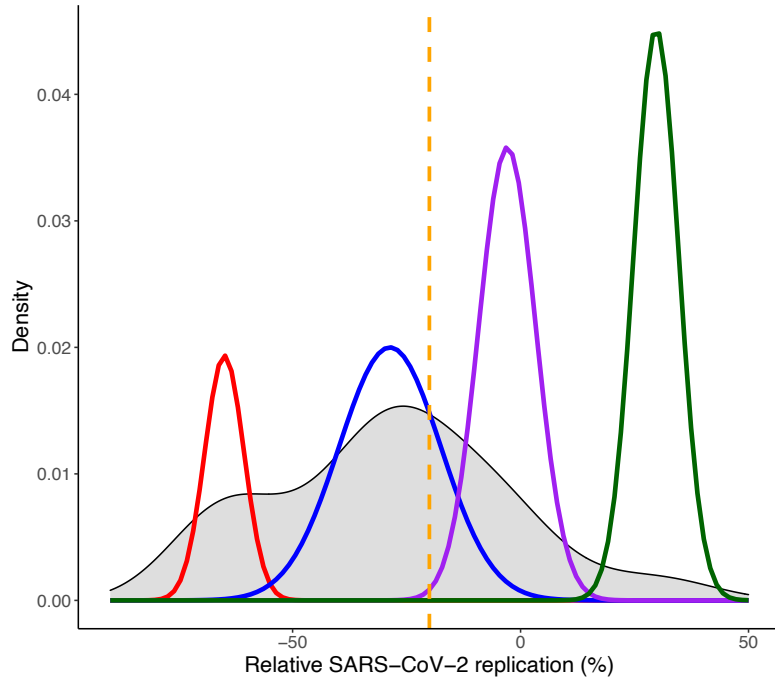


925

926 **Supplementary Figure 8. Experimental evaluation of the antiviral effect of FDA-approved**
 927 **drugs in lung adenocarcinoma cell lines. Related to Figure 4 and Supplementary Table 2.**

928 A set of drugs, predicted by ViroTreat for the GI context and with validated antiviral effect in Caco-
 929 2 cells were evaluated in Calu-3 (a) and A549-ACE2 (b) cells. The scatter-plots show the effect
 930 of each drug—SARS-CoV-2 replication shown in cyan and cell viability in red—relative to vehicle
 931 control (y-axis), assayed at different concentrations (x-axis) in triplicate. The lines indicate the
 932 average across replicates. * $p < 0.05$, ** $p < 0.01$, **** $p < 10^{-4}$, 1-tailed Student's t-test, BC.

933



934

935 **Supplementary Figure 9. Distribution for the relative effect of the evaluated drugs on**
 936 **SARS-CoV-2 replication. Related to Figure 4.** Histogram and Gaussian Mixture Model (GMM)
 937 fitted to the relative effect of the drugs, expressed as percentage, on SARCS-CoV-2 replication
 938 in Caco-2 cells. The dashed orange vertical line represents the threshold of 20% used as
 939 additional criteria when considering the antiviral effect of a drug.

940

941

942

943 **Supplementary Table 1: SARS-CoV-2 host cell RNA-Seq and scRNA-Seq datasets.**

<i>Model</i>	<i>Type of Data</i>	<i>Publication</i>	<i>Source</i>
Calu3	Bulk RNASeq	Wyler et al. ¹	(GEO) GSE148729
H1299	Bulk RNASeq	Wyler et al. ¹	(GEO) GSE148729
Caco2	Bulk RNASeq	Wyler et al. ¹	(GEO) GSE148729
A549	Bulk RNASeq	Blanco Melo et al. ²	(GEO) GSE147507
Lung Organoids	Bulk RNASeq		(GEO) GSE160435
NHBE	Bulk RNASeq	Blanco Melo et al. ²	(GEO) GSE147507
Human lung	Bulk RNASeq	Blanco Melo et al. ²	(GEO) GSE147507
Calu3	scRNASeq	Wyler et al. ¹	(GEO) GSE148729
H1299	scRNASeq	Wyler et al. ¹	(GEO) GSE148729
Ileum	scRNASeq	Triana et al. ³	Boulant Lab
Colon	scRNASeq	Triana et al. ³	Boulant Lab
Vero6	CRISPRcas9	Wei et al. ⁴	Supplementary Data
A549	CRISPRcas9	Daniloski et al. ⁵	Supplementary Data
Huh-7.5	CRISPRcas9	Wang et al. ⁶	Supplementary Data
Huh-7.5	CRISPRcas9	Schneider et al. ⁷	Supplementary Data

944

945 ¹Wyler, E., et al. (2021). *iScience* **24**(3): 102151.

946 ²Blanco-Melo, D., et al. (2020). *Cell* **181**(5): 1036-1045 e1039.

947 ³Triana, S., et al. (2021). *Mol Syst Biol* **17**(4): e10232.

948 ⁴Wei, J., et al. (2021). *Cell* **184**(1): 76-91 e13.

949 ⁵Daniloski, Z., et al. (2021). *Cell* **184**(1): 92-105 e116.

950 ⁶Wang, R., et al. (2021). *Cell* **184**(1): 106-119 e114.

951 ⁷Schneider, W. M., et al. (2021). *Cell* **184**(1): 120-132 e114.

952

953 **Supplementary Table 2:** Drugs library, ViroTreat and focused validation screen results.

954 < See supplementary file Table-S2.xlsx >

955

956 **Supplementary Table 3:** Organoids' culture media.

<i>Compound</i>	<i>Final concentration</i>
Basal media	
Ad DMEM/F12	
+GlutaMAX	
+HEPES	
+P/S	
L-WRN	50% by volume
B27	1:50
N-acetyl-cysteine	1 mM
EGF	50 ng/mL
A83-01	500 nM
IGF-1	100 ng/mL
FGF basic	50 ng/mL
Gastrin	10 mM
Differentiation Media	
Ad DMEM/F12	
+GlutaMAX	
+HEPES	
+P/S	
B27	1:50
N-acetyl-cysteine	1 mM
R-spondin	5% by volume
Noggin	50 ng/mL
EGF	50 ng/mL
Gastrin	10 mM
A83-01	500 nM

957

958 **Supplementary Table 4:** PCR primers.

<i>Gene name</i>	<i>Species</i>	<i>Forward sequence</i>	<i>Reverse sequence</i>
HPRT1	Human	cct ggc gtc gtg att agt gat	aga cgt tca gtc ctg tcc ata a
COV1	SARS-CoV-2	gcc tct tct gtt cct cat cac	aga cag cat cac cgc cat tg

959

960

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [TableS2.xlsx](#)