

Transcriptomics-based drug repositioning pipeline identifies therapeutic candidates for COVID-19

Brian Le

University of California, San Francisco

Gaia Andreoletti

University of California, San Francisco

Tomiko Oskotsky

University of California, San Francisco

Albert Vallejo-Gracia

Gladstone Institutes

Romel Rosales Ramirez

Icahn School of Medicine at Mount Sinai

Katharine Yu

University of California, San Francisco

Idit Kost

University of California, San Francisco

Kristoffer Leon

Gladstone Institutes

Daniel Bunis

University of California, San Francisco

Christine Li

University of California, San Francisco

G. Renuka Kumar

Gladstone Institutes

Kris White

Icahn School of Medicine at Mount Sinai

Adolfo Garcia-Sastre

Icahn School of Medicine at Mount Sinai

Melanie Ott

Gladstone Institutes

Marina Sirota (✉ Marina.Sirota@ucsf.edu)

University of California, San Francisco

Keywords: SARS-CoV-2, treatment, computational drug repositioning pipeline, Kolmogorov-Smirnov Statistic, Connectivity Map (CMap), Calu-3, 293T-ACE2

Posted Date: March 30th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-333578/v1>

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

[Read Full License](#)

1 **Transcriptomics-based drug repositioning pipeline identifies**
2 **therapeutic candidates for COVID-19**

3
4 **Authors:** Brian L. Le^{1,2+}, Gaia Andreoletti^{1,2+}, Tomiko Oskotsky^{1,2+}, Albert Vallejo-
5 Gracia³, Romel Rosales^{4,5}, Katharine Yu^{1,2,6}, Idit Kost^{1,2}, Kristoffer E. Leon³, Daniel G.
6 Bunis^{1,2,6}, Christine Li^{1,2,7}, G. Renuka Kumar³, Kris M. White^{4,5}, Adolfo García-
7 Sastre^{4,5,8,9}, Melanie Ott^{3,10}, Marina Sirota^{1,2*}

8
9 **Affiliations:**

10 ¹ Department of Pediatrics, UCSF, SF, CA, USA

11 ² Bakar Computational Health Sciences Institute, UCSF, SF, CA, USA

12 ³ Gladstone Institute of Virology, Gladstone Institutes, SF, CA, USA

13 ⁴ Department of Microbiology, Icahn School of Medicine at Mount Sinai, New York, NY,
14 USA

15 ⁵ Global Health and Emerging Pathogens Institute, Icahn School of Medicine at Mount
16 Sinai, New York, NY, USA

17 ⁶ Biomedical Sciences Graduate Program, UCSF, SF, CA, USA

18 ⁷ Shanghai American School, Shanghai, China

19 ⁸ Department of Medicine, Division of Infectious Diseases, Icahn School of Medicine at
20 Mount Sinai, New York, NY, USA

21 ⁹ The Tisch Cancer Institute, Icahn School of Medicine at Mount Sinai, New York, NY,
22 USA

23 ¹⁰ Department of Medicine, UCSF, SF, CA, USA

24
25 + These authors contributed equally to this manuscript.

26
27 *Correspondence should be addressed to: marina.sirota@ucsf.edu
28

29

30

31

32

33

34

35

36

37

38 **Abstract**

39 The novel SARS-CoV-2 virus emerged in December 2019 and has few effective
40 treatments. We applied a computational drug repositioning pipeline to SARS-CoV-2
41 differential gene expression signatures derived from publicly available data. We utilized
42 three independent published studies to acquire or generate lists of differentially
43 expressed genes between control and SARS-CoV-2-infected samples. Using a rank-
44 based pattern matching strategy based on the Kolmogorov-Smirnov Statistic, the
45 signatures were queried against drug profiles from Connectivity Map (CMap). We
46 validated sixteen of our top predicted hits in live SARS-CoV-2 antiviral assays in either
47 Calu-3 or 293T-ACE2 cells. Validation experiments in human cell lines showed that 11
48 of the 16 compounds tested to date (including clofazimine, haloperidol and others) had
49 measurable antiviral activity against SARS-CoV-2. These initial results are encouraging
50 as we continue to work towards a further analysis of these predicted drugs as potential
51 therapeutics for the treatment of COVID-19.

52

53

54

55

56

57

58

59

60

61 **Introduction**

62 SARS-CoV-2 has already claimed at least a million lives, has been detected in at
63 least 40 million people, and has likely infected at least another 200 million. The spectrum
64 of disease caused by the virus can be broad ranging from silent infection to lethal disease,
65 with an estimated infection-fatality ratio around 1%¹. SARS-CoV-2 infection has been
66 shown to affect many organs of the body in addition to the lungs². Three epidemiological
67 factors increase the risk of disease severity: increasing age, decade-by-decade, after the
68 age of 50 years; being male; and various underlying medical conditions¹. However, even
69 taking these factors into account, there is immense interindividual clinical variability in
70 each demographic category considered³. Recently, researchers found that more than
71 10% of people who develop severe COVID-19 have misguided
72 antibodies—autoantibodies—that attack the innate immune system. Another 3.5% or
73 more of people who develop severe COVID-19 carry specific genetic mutations that
74 impact innate immunity. Consequently, both groups lack effective innate immune
75 responses that depend on type I interferon, demonstrating a crucial role for type I
76 interferon in protecting cells and the body from COVID-19. Whether the type I interferon
77 has been neutralized by autoantibodies or—because of a faulty gene—is produced in
78 insufficient amounts or induced an inadequate antiviral response, the absence of type I
79 IFN-mediated immune response appears to be a commonality among a subgroup of
80 people who suffer from life-threatening COVID-19 pneumonia³.

81 While numerous efforts are underway to identify potential therapies targeting
82 various aspects of the disease, there is a paucity of clinically proven treatments for
83 COVID-19. There have been efforts to therapeutically target the hyperinflammation

84 associated with severe COVID-19⁴, as well as to utilize previously identified antiviral
85 medications^{5,6}. One of these antivirals, remdesivir, an intravenously administered RNA-
86 dependent RNA polymerase inhibitor, showed positive preliminary results in patients with
87 severe COVID-19⁷. In October 2020, the FDA approved remdesivir for the treatment of
88 COVID-19⁸. Dexamethasone has also been shown to reduce the mortality rate in cases
89 of severe COVID-19⁹.

90 Nevertheless, the lack of treatments and the severity of the current health
91 pandemic warrant the exploration of rapid identification methods of preventive and
92 therapeutic strategies from every angle. The traditional paradigm of drug discovery is
93 generally regarded as protracted and costly, taking approximately 15 years and over \$1
94 billion to develop and bring a novel drug to market¹⁰. The repositioning of drugs already
95 approved for human use mitigates the costs and risks associated with early stages of
96 drug development, and offers shorter routes to approval for therapeutic indications.
97 Successful examples of drug repositioning include the indication of thalidomide for severe
98 erythema nodosum leprosum and retinoic acid for acute promyelocytic leukemia¹¹. The
99 development and availability of large-scale genomic, transcriptomic, and other molecular
100 profiling technologies and publicly available databases, in combination with the
101 deployment of the network concept of drug targets and the power of phenotypic
102 screening, provide an unprecedented opportunity to advance rational drug design.

103 Drug repositioning is being extensively explored for COVID-19. High-throughput
104 screening pipelines have been implemented in order to quickly test drug candidates as
105 they are identified¹²⁻¹⁵. In the past, our group has successfully applied a transcriptomics-
106 based computational drug repositioning pipeline to identify novel therapeutic uses for

107 existing drugs¹⁶. This pipeline leverages transcriptomic data to perform a pattern-
108 matching search between diseases and drugs. The underlying hypothesis is that for a
109 given disease signature consisting of a set of up and down-regulated genes, if there is a
110 drug profile where those same sets of genes are instead down-regulated and up-
111 regulated, respectively, then that drug could be therapeutic for the disease. This method
112 based on the Kolmogorov-Smirnov (KS) test statistic has shown promising results for a
113 variety of different indications, including inflammatory bowel disease¹⁷,
114 dermatomyositis¹⁸, cancer¹⁹⁻²¹, and preterm birth²².

115 In existing work from Xing et al.²³, this pipeline has been used to identify potential
116 drug hits from multiple input disease signatures derived from SARS-CoV or MERS-CoV
117 data. The results were aggregated to obtain a consensus ranking, with 10 drugs selected
118 for *in vitro* testing against SARS-CoV-2 in Vero E6 cell lines, with four drugs (bortezomib,
119 dactolisib, alvocidib and methotrexate) showing viral inhibition²³. However, this pipeline
120 has not yet been applied specifically to SARS-CoV-2 infection.

121 A variety of different transcriptomic datasets related to SARS-CoV-2 were
122 published in the spring of 2020. In May 2020, Blanco-Melo et al. studied the transcriptomic
123 signature of SARS-CoV-2 in a variety of different systems, including human cell lines and
124 a ferret model²⁴. By infecting human adenocarcinomic alveolar basal epithelial cells with
125 SARS-CoV-2 and comparing to controls, the authors generated a list of 120 differentially
126 expressed genes. They observed two enriched pathways: one composed primarily of
127 type-I interferon-stimulated genes (ISGs) involved in the cellular response to viral
128 infection; and a second composed of chemokines, cytokines, and complement proteins
129 involved in the humoral response. After infecting the cell lines, Blanco-Melo et al. did not

130 detect either ACE2 or TMPRSS2, which are the SARS-CoV-2 receptor and SARS-CoV-
131 2 protease, respectively²⁵. However, supported viral replication was observed, thereby
132 allowing the capture of some of the biological responses to SARS-CoV-2.

133 In May 2020, another study by Lamers et al. examined SARS-CoV-2 infection in
134 human small intestinal organoids grown from primary gut epithelial stem cells²⁶. The
135 organoids were exposed to SARS-CoV-2 and grown in various conditions, including Wnt-
136 high expansion media. Enterocytes were readily infected by the virus, and RNA
137 sequencing revealed upregulation of cytokines and genes related to type I and III
138 interferon responses.

139 A limited amount of transcriptomic data from human samples has also been
140 published. One study detailed the transcriptional signature of bronchoalveolar lavage fluid
141 (of which responding immune cells are often a primary component) of COVID-19 patients
142 compared to controls²⁷. Despite a limited number of samples, the results revealed
143 inflammatory cytokine profiles in the COVID-19 cases, along with enrichments in the
144 activation of apoptosis and the P53 signaling pathways.

145 On the drug side, data are available in the form of differential gene expression
146 profiles from testing on human cells. Publicly-available versions include the Connectivity
147 Map (CMap)²⁸, which contains genome-wide testing on approximately 1,300 drugs,
148 wherein the differential profile for a drug was generated by comparing cultured cells
149 treated with the drug to untreated control cultures.

150 Here, we applied our existing computational drug repositioning pipeline to identify
151 drug profiles with significantly reversed differential gene expression compared to several
152 diverse input signatures for SARS-CoV-2 effects on human cells. By taking into account

153 a broader view of differentially expressed gene sets from both cell line and organoid
154 disease models and human samples, the predictions are complementary to other drug
155 discovery approaches. We identified 102 unique drug hits, from which 25 were identified
156 in at least two of the signatures, several of which have been already investigated in clinical
157 trials. We furthermore explore our findings in the context of other computational drug
158 repurposing efforts for COVID-19. Finally, we tested 16 of our top predicted hits in live
159 SARS-CoV-2 antiviral assays. Four of the top predicted inhibitors were tested for virus
160 inhibition in a human lung cell line, Calu-3, infected with SARS-CoV-2 with quantitation of
161 the secreted virus assessed by RT-qPCR assay. Thirteen predicted inhibitors (including
162 one tested in Calu-3) were incubated with SARS-CoV-2 infected human embryonic kidney
163 293T cells overexpressing ACE2 (293T-ACE2) with viral replication determined using an
164 immunofluorescence-based assay.

165

166

167

168

169

170

171

172

173

174

175

176 **Results**

177 In this study, we applied our drug repositioning pipeline to SARS-CoV-2 differential
178 gene expression signatures derived from publicly available RNA-seq data (Figure 1). The
179 transcriptomic data were generated from distinct types of tissues, so rather than
180 aggregating them together, we predicted therapeutics for each signature and then
181 combined the results. We utilized three independent gene expression signatures (labelled
182 “ALV”, “EXP”, and “BALF”), each of which consisted of lists of differentially expressed
183 genes between SARS-CoV-2 samples and their respective controls. The ALV signature
184 was generated from human adenocarcinomic alveolar basal epithelial cells by comparing
185 SARS-CoV-2 infection to mock-infection conditions²⁴. The EXP signature originated from
186 a study where organoids, grown from human intestinal cells expanded in Wnt-high
187 expansion media, were infected with SARS-CoV-2 and then compared to controls²⁶. The
188 BALF signature was from a contrast of primary human BALF samples from two COVID-
189 19 patients versus three controls²⁷. Each of these signatures was contrasted with drug
190 profiles of differential gene expression from CMap.

191 For each of the input signatures, we applied a significance threshold false
192 discovery rate (FDR) < 0.05. We further applied minimum fold change thresholds in order
193 to identify the driving genes. The ALV signature had only 120 genes, with 109 genes
194 shared with the drug profiles; in order to maintain at least 100 genes for the pattern-
195 matching algorithm to work with, we applied no fold-change threshold. For the EXP
196 signature, we applied a $|\log_2FC| > 2$ cutoff, resulting in 125 genes for the expansion
197 signature (108 shared with the drug profiles). For the BALF signature, we processed the
198 raw read count data to calculate differential gene expression values. We applied a

199 $|\log_2FC| > 4$ cutoff, with the BALF data yielding 1,349 protein-coding genes for the lavage
200 fluid signature (941 shared with the drug profiles). As the sample types across the three
201 datasets were very different, we used different fold change thresholds to identify the
202 appropriate gene signatures to be used for drug repurposing. The gene lists for each of
203 these signatures can be found in the supplement (Tables S1, S2, S3).

204 We used GSEA (Gene Set Enrichment Analysis)^{29,30} to annotate enriched (FDR
205 0.05) Hallmark pathways from each of the input signatures (Figure 2A). A number of
206 pathways common to at least two signatures were found. Interferon alpha response and
207 interferon beta response were upregulated in the ALV and EXP signatures. Adipogenesis
208 and cholesterol homeostasis pathways were downregulated in the EXP and BALF
209 signatures. KRAS signaling, and mTORC1 (mammalian target of rapamycin complex 1)
210 signaling were enriched in all three signatures, but not in the same direction, showing the
211 diversity of effects SARS-CoV-2 may have on human cells, and highlighting a need for
212 utilization of diverse profiles as we do in the present study. When we look at the
213 contributing genes within the three signatures (Figure 2B), we found one overlapping
214 upregulated gene - Dickkopf WNT Signaling Pathway Inhibitor 1 (DKK1). We used the
215 publicly available single-cell RNAseq dataset GSE128033³¹ composed of 13 patients (4
216 healthy, 3 presenting with mild COVID-19 symptoms, and 6 presenting with severe
217 COVID-19 symptoms) to further characterize the expression of DKK1 (Figure S1). Data
218 were re-analyzed following the standard Seurat pipeline. From the analyses of the single-
219 cell data, DKK1 is highly expressed in COVID-19 patients compared to controls,
220 specifically in severe patients and it is expressed by epithelial cells.

221 After analyzing the input SARS-CoV-2 signatures, we utilized our repositioning
222 pipeline to identify drugs with reversed profiles from CMap (Figure 1). Significantly
223 reversed drug profiles were identified for each of the signatures using a permutation
224 approach: 30 hits from the ALV signature (Table S4), 15 hits from the EXP signature
225 (Table S5), and 86 hits from the BALF signature (Table S6). When visualizing the gene
226 regulation of the input signatures and their respective top 15 drug hits, the overall reversal
227 pattern can be observed (Figure 2C-E). Interestingly, we found several drugs shared
228 across datasets that significantly reversed the disease signature. For example,
229 haloperidol, highlighted in purple in Figures 2C-E, was shown to reverse the disease gene
230 signature from three datasets, whereas levopropoxyphene, shown in green in Figures
231 2D-E, was observed to reverse the disease gene signature from two of the datasets. In
232 total, our analysis identified 102 unique drug hits (Table S7). Twenty-five drug hits
233 reversed at least two signatures ($p = 0.0334$, random sampling), and four drug hits
234 reversed three signatures ($p = 0.0599$, random sampling) (Table 1, Figure 3A).

235 We further characterized the common hits by examining their interactions with
236 proteins in humans. We used known drug targets from DrugBank³² and predicted
237 additional targets using the similarity ensemble approach (SEA)³³. We visualized the
238 known interactions from DrugBank in a network. Figure 3B shows the connectivity across
239 compounds highlighting both single drug genes (such as *SIGMAR1* for haloperidol) and
240 genes shared across drugs, such as *ADRA2A* and *DRD1* for haloperidol and co-
241 dergocrine mesilate. The proteins with the most known interactions with our list of 25
242 drugs included adrenergic receptors (particularly $\alpha 2$ adrenoreceptors), dopamine
243 receptors, and serotonin receptors.

244 To confirm the validity of our approach, the inhibitory effects of 16 of our drug hits
245 which significantly reversed multiple SARS-CoV-2 profiles were assessed in live **pCC50**
246 antiviral assays. Next, we wanted to test the potential of the predicted compounds to
247 inhibit viral activity robustly using different human cell lines - Calu-3 and 293T-ACE2. The
248 respective selection of 16 and 13 compounds for testing was based on side effect profiles
249 and compound availability.

250 The inhibitory effects of haloperidol, clofazimine, valproic acid, and fluticasone
251 were evaluated in SARS-CoV-2 infected Calu-3 cells (human lung epithelial cell line), with
252 remdesivir also tested as a positive control. From these five, remdesivir and haloperidol
253 inhibited viral replication (Figure 4A), and the inhibitory effect was also observed by
254 microscopy (Figure 4B). Toxicity assessments for haloperidol, clofazimine, valproic acid,
255 and fluticasone were evaluated using viability assays (Alamarblue) in Calu-3 cells treated
256 with each compound for 72h (n=1, with 2 technical replicates). No significant differences
257 between controls and the biological replicates were detected using a non-parametric test
258 (Kruskal-Wallis) (Figure S6). Fluticasone and bacampicillin showed some toxicity in a
259 dose-dependent manner at the highest doses tested. Haloperidol, clofazimine and
260 valproic acid did not show significantly reduced cell viability (Figure S6).

261 Additionally, 13 drugs (bacampicillin, ciclopirox, ciclosporin, clofazimine,
262 dicycloverine, fludrocortisone, isoxicam, lansoprazole, metixene, myricetin, pentoxifylline,
263 sirolimus, tretinoin) were independently assessed in a live SARS-CoV-2 antiviral assay.
264 Remdesivir was again used as a positive control. This testing involved six serial dilutions
265 of each drug to inhibit the replication of SARS-CoV-2 in 293T-ACE2 cells using an
266 immunofluorescence-based antiviral assay³⁴. All antiviral assays were paired with

267 cytotoxicity assays using identical drug concentrations in uninfected human 293T-ACE2
268 cells. Positive control remdesivir and 10 of our predicted drugs (bacampicillin, ciclopirox,
269 ciclosporin, clofazimine, dicycloverine, isoxicam, metixene, pentoxifylline, sirolimus, and
270 tretinoin) showed antiviral efficacy against SARS-CoV-2, reducing viral infection by at
271 least 50%, that was distinguishable from their cytotoxicity profile when tested in this cell
272 line (Figure 5). Several inhibitors showed micromolar to sub-micromolar antiviral efficacy,
273 including clofazimine, ciclosporin, ciclopirox, and metixene. These results not only confirm
274 our predictive methods, but have also identified several clinically-approved drugs with
275 potential for repurposing for the treatment of COVID-19.

276

277

278

279

280

281

282

283

284

285

286

287

288 **Discussion**

289 Here, we used a transcriptomics-based drug repositioning pipeline to predict
290 therapeutic drug hits for three different input SARS-CoV-2 signatures, each of which
291 came from distinct human cell or tissue origins. We found significant overlap of the
292 therapeutic predictions for these signatures. From 102 total drug hits, 25 drugs reversed
293 at least two signatures ($p = 0.0334$) and 4 drugs reversed three signatures ($p = 0.0599$).
294 The diversity of such signatures yet overlap of highlighted drugs underscores the utility of
295 the current pipeline for identification of drugs which might be therapeutic for the diverse
296 effects of SARS-CoV-2 infection.

297 Twenty-five of our drug hits reversed at least two of the three input signatures
298 (Figure 3). Notably, 14 of the 15 hits from the EXP signature were also hits for the BALF
299 signature, despite being generated from different types of tissue. The EXP signature was
300 generated from intestinal tissue, whereas the BALF signature was generated from
301 constituents of the respiratory tract. Among the common hits reversing at least two of the
302 signatures were two immunosuppressants (cyclosporin and sirolimus), an anti-
303 inflammatory medication (isoxicam), and two steroids (fludrocortisone and fluticasone).
304 Sirolimus (or rapamycin), an immunosuppressant and an mTOR inhibitor, is currently
305 undergoing investigation in several clinical trials in COVID-19 patients (NCT04371640,
306 NCT04341675, NCT04461340). Other hits currently in clinical trials for COVID-19
307 treatment include cyclosporin (NCT04412785, NCT04392531), niclosamide in
308 combination with diltiazem (NCT04558021), and clofazimine in combination with
309 interferon beta-1b (NCT04465695).

310 Among our four drug hits that reversed all three signatures, three drugs
311 demonstrated in vitro antiviral efficacy - bacampicillin, clofazimine, and haloperidol with

312 no toxicity effects (Figure S6). Our group found haloperidol decreased viral growth in
313 SARS-CoV-2 infected Calu-3 cells (Figure 4B) in a dose-dependent manner (Figure 4A).
314 Haloperidol is a psychiatric medication that is indicated for the treatment of psychotic
315 disorders including schizophrenia and acute psychosis. By blocking dopamine (mainly
316 D2) receptors in the brain, haloperidol eliminates dopamine neurotransmission which
317 leads to improvement of psychotic symptoms³⁵. Haloperidol can also bind to the sigma-1
318 and sigma-2 receptors, which are implicated in lipid remodeling and cell stress
319 response¹². As reported by Gordon et al.¹², the SARS-CoV-2 proteins Nsp6 and ORF9c
320 interact with the sigma-1 receptor and the sigma-2 receptor, respectively. Moreover, they
321 found that haloperidol decreased viral replication in SARS-CoV-2-infected Vero E6 cells
322 with, based on their reported pIC50 and pCC50 values, a calculated Selectivity Index (SI)
323 of 53.7¹². An SI greater than 10 is the generally accepted minimum cut-off for an antiviral
324 worth pursuing^{36,37}. In another more recent study, Gordon et al. found in their analysis of
325 a national electronic medical record database that fewer hospitalized COVID-19 patients
326 who were newly prescribed haloperidol and other sigma-binding typical antipsychotic
327 medications progressed to requiring mechanical ventilation compared to those who were
328 newly prescribed atypical antipsychotic medications that do not bind to sigma receptors¹⁴.

329 Our testing of clofazimine demonstrated submicromolar antiviral effects of this
330 drug in SARS-Co-V-2 infected 293T-ACE2 and Vero E6 cells (Figures 4 and S3).
331 Clofazimine is an orally administered antimycobacterial drug used in the treatment of
332 leprosy. By preferentially binding to mycobacterial DNA, clofazimine disrupts the cell cycle
333 and eventually kills the bacterium³⁸. In addition to being an antimycobacterial agent,
334 clofazimine also possesses anti-inflammatory properties primarily by inhibiting T

335 lymphocyte activation and proliferation³⁹. Yuan et al. found that clofazimine inhibits
336 SARS-CoV-2 replication by interfering with spike-mediated viral entry and viral RNA
337 replication⁴⁰. Their work also demonstrated that clofazimine has antiviral efficacy against
338 SARS-CoV-2 in human embryonic stem cell-derived cardiomyocytes and in an ex vivo
339 human lung culture system, as well as antiviral synergy with remdesivir demonstrating
340 the potential of clofazimine as part of a combination treatment regimen for COVID-19⁴⁰.

341 Our group found bacampicillin to have micromolar antiviral efficacy in SARS-Co-
342 V-2 infected 293T-ACE2 cells. Bacampicillin is an orally administered prodrug of
343 ampicillin typically prescribed for treating bacterial infections⁴¹. As identified by SPOKE⁴²,
344 bacampicillin was found to downregulate the GDF15 gene and upregulate the NFKB2
345 (Nuclear Factor Kappa B Subunit 2) gene in studies by Cmap²⁸ and LINCS⁴³. The GDF15
346 protein acts as a cytokine and is involved in stress response after cellular injury, and the
347 NFKB2 is a central activator of genes involved with inflammation and immune function⁴⁴.
348 Circulating levels of GDF15 have been found to be significantly higher in COVID-19
349 patients who die⁴⁵. Zhou et al.'s work revealed NF-kappa B signaling as one of the main
350 pathways of coronavirus infections in humans. While the rapid conversion of bacampicillin
351 to ampicillin in vivo makes this prodrug a less optimal therapeutic candidate for COVID-
352 19, our findings nevertheless provide insights into the immunologic and inflammatory
353 landscape from SARS-CoV-2 infection.

354 Overall, in testing of our drug hits across two human cell line assays, 11 of 16
355 exhibited inhibition of SARS-CoV-2 infection. In particular, three of our four consensus
356 drug hits demonstrated antiviral efficacy, with haloperidol showing reproducible inhibition
357 in Calu-3 cells, and bacampicillin and clofazimine inhibiting viral activity in 293T-ACE2

358 cells without cytotoxicity. Many of our tested drugs can be administered orally, and
359 several are on the WHO Model List of Essential medications, including ciclosporin,
360 clofazimine, and haloperidol⁴⁶. These results suggest that our drug repositioning pipeline
361 can rapidly identify readily available potential therapeutics in antiviral contexts.

362 There are several limitations of our approach that should be recognized. In
363 general, drug repositioning pipelines are reliant upon data being encoded in a computable
364 format. Considering variability of conditions under which experimental and clinical data
365 are obtained, this also implies a particular set of limitations and biases to the ensuing
366 results. Therefore, significant and concerted community efforts are necessary for efficient
367 usage of the existing biomedical and clinical information and extraction of knowledge from
368 this information, which may allow better repositioning of the current drugs⁴⁷.

369 In our work, we applied the KS-based similarity metric on the CMap database.
370 Rank-based methods, such as the KS statistic, may suffer from high false positive rates,
371 as genes not differentially expressed can be ranked high and contribute to the similarity
372 measurement for drug and disease signatures^{48,49}. Moreover, rank-based approaches
373 can also miss many potential drugs as ranking captured just a small part of information in
374 a gene expression profile. Alternative methods have been proposed, such as EMUDRA⁴⁹
375 and XSum⁵⁰; however, they have not been widely adopted by the community. Future work
376 might include evaluating multiple similarity metrics on larger datasets⁴⁹. The studies that
377 we leveraged here are also limited because of their small sample size, which might
378 explain the small gene overlap across the signatures. While the different sample types
379 are able to capture the heterogeneity of the response to viral infection, because data
380 generated from cell lines and organoids (the ALV and EXP signatures, respectively) might

381 not accurately represent the biological changes and responses in human infection, *in*
382 *vivo* experiments are needed to better understand the biology of the disease and the
383 effect of these drugs on it. Moreover, although the BALF signature was generated from
384 fluid recovered from lavage of infected human tissues, this primary response data was
385 aggregated from a very limited sample size (2 cases and 3 controls). As the sample types
386 across the three datasets were very different, we used different fold change thresholds to
387 identify the signature genes to be used for drug repurposing. Gathering samples from a
388 larger number of patients should generate a more robust gene expression signature and
389 better inform therapeutic predictions. Furthermore, the drug profiles from CMap were
390 generated from cell line data; drug data generated from more relevant tissue cultures (e.g.
391 lung tissue) may generate more appropriate comparisons. Finally, our validation
392 approaches focus on *in vitro* studies, which are limited, and warrant further *in-vivo* testing
393 of the proposed compounds.

394 The drug development response for SARS-CoV-2 / COVID-19 is rapidly
395 developing. One drug, remdesivir, recently received FDA approval for the treatment of
396 COVID-19, and numerous other drugs are being actively explored for possible therapeutic
397 value in COVID-19 cases. Utilizing a diverse set of transcriptomic SARS-CoV-2
398 signatures, our drug repositioning pipeline identified 25 therapeutic candidates. Validation
399 experiments revealed antiviral activity for 11 of 16 drug hits. Further clinical investigation
400 into these drug hits, *in vivo* assays as well as potential combination therapies is warranted
401 to further investigate both the anti-viral as well as side effect profile of the drugs.

402 **Materials and Methods**

403 ***Study design***

404 We have previously developed and used a transcriptomics based bioinformatics
405 approach for drug repositioning in various contexts including inflammatory bowel disease,
406 dermatomyositis, and spontaneous preterm birth. For a list of differentially expressed
407 genes, the computational pipeline compares the ranked differential expression of a
408 disease signature with that of a profile^{16,19,28}. A reversal score based on the Kolmogorov-
409 Smirnov statistic is generated for each disease-drug pair, with the idea that if the drug
410 profile significantly reverses the disease signature, then the drug could be potentially
411 therapeutic for the disease.

412

413 ***SARS-CoV-2 gene expression signatures***

414 Blanco-Melo et al. generated a differential gene expression signature using RNA-
415 seq on human adenocarcinomic alveolar basal epithelial cells infected with SARS-CoV-2
416 propagated from Vero E6 cells (GSE147507, 67 samples)²⁴. Due to the fast-moving
417 nature of the research topic, we opted to use this cell line data in lieu of waiting for
418 substantial patient-level data. This work identified 120 differentially expressed genes
419 (DEGs) – 100 upregulated and 20 downregulated. We used these 120 genes as the ALV
420 signature for our computational pipeline (Table S1).

421 Lamers et al. performed RNA-seq on their organoid samples, from which
422 differentially expressed genes were calculated²⁶. These samples were grown in a medium
423 with a Wnt surrogate supplement and infected with SARS-CoV-2 propagated from Vero
424 E6 cells (GSE149312, 22 samples). They detected 434 significant DEGs (FDR < 0.05).
425 We additionally applied a fold-change cutoff ($|\log_2 \text{FC}| > 2$), resulting in 125 genes used
426 as the EXP signature (Table S2).

427 Xiong et al. performed RNA-seq analysis of BALF samples from two COVID-19
428 patients (two samples per patient) and three healthy controls²⁷. We processed their raw
429 read counts in order to construct a differential signature (see below for details). FASTQ
430 files were downloaded from the Genome Sequence Archive^{51,52} under accession number
431 CRA002390. Paired-end reads were mapped to the hg19 human reference genome using
432 Salmon (v.1.2.0) and assigned Ensembl genes. After read quality control, we obtained
433 quantifications for 55,640 genes in all samples. In order to identify genes differentially
434 expressed between cases and controls for the BALF samples, we quantified gene
435 expression as raw counts. Raw counts were used as inputs to DESeq2 (v.1.24.0 R
436 package) to call differentially expressed genes (DEGs). After adjusting for the sequencing
437 platform, the default settings of DESeq2 were used. Principal components were
438 generated using the DESeq2 function (Figure S2), and heat maps were generated using
439 the Bioconductor package pheatmap (v.1.0.12) using the rlog-transformed counts (Figure
440 S3). Values shown are rlog-transformed and row-normalized. Volcano plots were
441 generated using the Bioconductor package EnhancedVolcano (v.1.2.0) (Figure S4).
442 Retaining only protein-coding genes and applying both a significance threshold and a
443 fold-change cutoff ($FDR < 0.05$, $|\log_2 FC| > 4$), we obtained 1,349 genes to be used as
444 the BALF signature (Table S3).

445

446 ***Pathway enrichment analysis***

447 Functional enrichment gene-set analysis for GSEA (Gene Set Enrichment
448 Analysis) was performed using fgsea (v.1.12.0 R package) and the input gene lists were
449 ranked by log2 fold change. The 50 Hallmark Gene Sets used in the GSEA analysis were

450 downloaded from MSigDB Signatures database^{29,53}. For GO (Gene Ontology) terms,
451 identification of enriched biological themes was performed using the DAVID database⁵⁴.

452

453 ***Drug gene expression profiles***

454 Drug gene expression profiles were sourced from Connectivity Map (CMap), a
455 publicly-available database of drugs tested on cancer cell lines²⁸. CMap contains a set of
456 differential gene expression profiles generated from treating cultured human cells with a
457 variety of different drugs and experimental compounds. These profiles were generated
458 using DNA microarrays to assay mRNA expression. These drug profiles are ranked
459 genome-wide profiles (~22,000 genes) of the effects of the drugs on various cell lines.
460 6,100 gene expression profiles are presented in CMap. A total of 1,309 compounds were
461 tested in up to 5 different cell lines. The overlap between the gene lists of CMap and the
462 SARS-CoV-2 signature is 109 genes.

463

464 ***Computational gene expression reversal scoring***

465 To compute reversal scores, we used a non-parametric rank-based method similar
466 to the Kolmogorov-Smirnov test statistic. This analysis was originally suggested by the
467 creators of the CMap database and has since been implemented in a variety of different
468 settings^{16–19,22,28}. As also described by others, the drug signature is compared with the
469 gene expression profiles. By splitting the gene signature into two lists containing only
470 upregulated genes and downregulated genes, a so-called connectivity score is estimated
471 via several auxiliary variables using a nonparametric rank-ordered Kolmogorov–Smirnov
472 (KS) test⁴⁸. Similar to past works, we applied a pre-filtering step to the CMap profiles to

473 maintain only drug profiles which were significantly correlated with another profile of the
474 same drug. Drugs were assigned reversal scores based on their ranked differential gene
475 expression profile relative to the SARS-CoV-2 ranked differential gene expression
476 signature. A negative reversal score indicated that the drug had a profile which reversed
477 the SARS-CoV-2 signature; that is, up-regulated genes in the SARS-CoV-2 signature
478 were down-regulated in the drug profile and vice versa.

479

480 ***Statistical analysis***

481 P-values were adjusted using the false discovery rate (FDR; Benjamini-Hochberg)
482 procedure. P-values for individual drug hits were obtained by comparing reversal scores
483 to a distribution of random scores. Negative reversal scores were considered significant
484 if they met the criterion $FDR < 0.05$. For drugs tested multiple times (e.g. different cell
485 lines), we used the most reversed profile (lowest negative score). For significance values
486 of the number of drugs reversing multiple signatures, we constructed distributions of the
487 common reversal (reversing two of three signatures) and the consensus reversal
488 (reversing three of three signatures) by randomly sampling the same number of drug
489 profiles for each signature from CMap.

490

491 ***Single-cell data analysis***

492 Quantification files were downloaded from GEO GSE145926. An individual Seurat object
493 for each sample was generated using Seurat v.3. While the data has been filtered by
494 10x's algorithm, we still needed to ensure the remaining cells are clean and devoid of
495 artifacts. We calculated three confounders for the dataset: mitochondrial percentage,

496 ribosomal percentage, and cell cycle state information. For each sample, cells were
497 normalized for genes expressed per cell and per total expression, then multiplied by a
498 scale factor of 10,000 and log-transformed. Low quality cells were excluded from our
499 analyses— this was achieved by filtering out cells with greater than 5,000 and fewer than
500 300 genes and cells with high percentage of mitochondrial and ribosomal genes (greater
501 than 10% for mitochondrial genes, and 50% for ribosomal genes). SCTransform is a
502 relatively new technique that uses "Pearson Residuals" (PR) to normalize the data. PRs
503 are independent of sequencing depth⁵⁵. We "regress out" the effects of mitochondrial and
504 ribosomal genes, and the cell cycling state of each cell, so they do not dominate the
505 downstream signal used for clustering and differential expression. We then performed a
506 lineage auto-update disabled r dimensional reduction (RunPCA function). Then, each
507 sample was merged together into one Seurat object. Data were then re-normalized and
508 dimensionality reduction and significant principal components were used for downstream
509 graph-based, semi-unsupervised clustering into distinct populations (FindClusters
510 function) and uniform manifold approximation and projection (UMAP) dimensionality
511 reduction was used. For clustering, the resolution parameter was approximated based on
512 the number of cells according to Seurat guidelines; a vector of resolution parameters was
513 passed to the FindClusters function and the optimal resolution of 0.8 that established
514 discernible clusters with distinct marker gene expression was selected. We obtained a
515 total of 21 clusters representing the major immune and epithelial cell populations. To
516 identify marker genes driving each cluster, the clusters were compared pairwise for
517 differential gene expression (FindAllMarkers function) using the Likelihood ratio test
518 assuming an underlying negative binomial distribution (negbinom). For visualization of

519 gene expression data between different samples a number of Seurat functions were used:
520 FeaturePlot, VlnPlot and DotPlot.

521

522 **Cell Lines**

523 For studies at the Gladstone Institutes, Calu-3 cells, a human lung epithelial cell line
524 (American Type Culture Collection, ATCC HTB-55), were cultured in advanced MEM
525 supplemented with 2.5% fetal bovine serum (FBS) (Gibco, Life Technologies), 1% L-
526 GlutaMax (ThermoFisher), and 1% penicillin/streptomycin (Corning) at 37°C and 5% CO₂.
527 SARS-CoV-2 Isolate USA-WA1/2020 was purchased from BEI Resources and
528 propagated and titered in Vero E6 cells. For studies carried out at Mount Sinai, SARS-
529 CoV-2 was propagated in Vero E6 cells (ATCC CRL-1586) and 293T-ACE2 cells (ATCC
530 CRL-3216).

531

532 **Compounds**

533 Selection of compounds for testing was based on side effect profiles and compound
534 availability. Bacampicillin (B0070000), ciclopirox (SML2011-50MG), ciclosporin
535 (C2163000), clofazimine (1138904-200MG), dicycloverine (D1060000), fludrocortisone
536 (1273003-200MG), fluticasone (1285873-100MG), haloperidol (H1512-5G), isoxicam
537 (I1762-1G), lansoprazole (1356916-150MG), metixene (M1808000), myricetin (M6760-
538 10MG), pentoxifylline (1508901-200MG), sirolimus (S-015-1ML), tretinoin (1674004-
539 5X30MG), and valproic acid (1708707-500MG) were purchased from Sigma-Aldrich.
540 Remdesivir (GS-5734) was purchased from Selleckchem.

541

542 Compounds were resuspended in DMSO according to manufacturer's instructions and
543 serially diluted to the relevant concentrations for treatment of infected cells.

544

545 ***Infection Experiments***

546 Work involving live SARS-CoV-2 was performed in the BSL3 facility at the
547 Gladstone Institutes with appropriate approvals. Calu-3 cells were seeded in 96-well
548 plates for 24h, infected with SARS-CoV-2 at a multiplicity of infection (MOI) of 0.05, and
549 treated with compounds. 72 hours post infection, supernatant was collected for RNA
550 extraction and the RNA was analyzed using RT-qPCR to quantify viral genomes present
551 in the supernatant. SARS-CoV-2 specific primers targeting the E gene region: 5'-
552 *ACAGGTACGTTAATAGTTAATAGCGT-3'* (Forward) and 5'-
553 *ATATTGCAGCAGTACGCACACA-3'* (Reverse) were used to quantify cDNA on the 7500
554 Fast Real-Time PCR system (Applied Biosystems). Cells were fixed with
555 paraformaldehyde and used for immunofluorescence analysis with dsRNA antibody
556 (SCICONS) and DAPI stain. Images were acquired and analyzed using ImageXpress
557 Micro Confocal High-Content Imaging System.

558 ***In Vitro Microneutralization Assay for SARS-CoV-2 Serology and Drug Screening***

559 For studies at Mount Sinai, SARS-CoV-2 was propagated in Vero E6 cells (ATCC
560 CRL-1586) and 293T-ACE2 cells (ATCC CRL-3216), as previously described^{12,34}. Two
561 thousand cells were seeded into 96-well plates in DMEM (10% FBS) and incubated for
562 24 h at 37 °C, 5% CO₂. Then, 2 h before infection, the medium was replaced with 100 µl
563 of DMEM (2% FBS) containing the compound of interest at concentrations 50% greater
564 than those indicated, including a DMSO control. The Vero E6 cell line used in this study

565 is a kidney cell line; therefore, we cannot exclude that lung cells yield different results for
566 some inhibitors. Plates were then transferred into the Biosafety Level 3 (BSL3) facility
567 and 100 PFU (MOI = 0.025) was added in 50 μ l of DMEM (2% FBS), bringing the final
568 compound concentration to those indicated. Plates were then incubated for 48 h at 37 °C.
569 After infection, supernatants were removed and cells were fixed with 4% formaldehyde
570 for 24 h before being removed from the BSL3 facility. The cells were then immunostained
571 for the viral NP protein (an in-house mAb 1C7, provided by Dr. Thomas Moran) with a
572 DAPI counterstain. Infected cells (488 nM) and total cells (DAPI) were quantified using
573 the Celigo (Nexcelcom) imaging cytometer. Infectivity is measured by the accumulation
574 of viral NP protein in the nucleus of the Vero E6 cells and 293T-ACE2 cells (fluorescence
575 accumulation). Percentage infection was quantified as ((infected cells/total cells) –
576 background) \times 100 and the DMSO control was then set to 100% infection for analysis.
577 The IC50 and IC90 for each experiment were determined using the Prism (GraphPad)
578 software. Cytotoxicity was also performed using the MTT assay (Roche), according to the
579 manufacturer's instructions. Cytotoxicity was performed in uninfected VeroE6 cells with
580 same compound dilutions and concurrent with viral replication assay. All assays were
581 performed in biologically independent triplicates.

582

583 ***Code Availability Statement***

584 The data used for the repositioning pipeline are all publicly available. The code for the
585 drug repositioning pipeline was adapted from reference 19 and is available at
586 https://github.com/brianlle/sirota_lab_covid_drug_repositioning.

587

588 **Acknowledgements**

589 This work was funded in part by the Program for Breakthrough Biomedical
590 Research (PBBR) Grant. This research was also partly funded by the Defense Advanced
591 Research Projects Agency (HR0011-19-2-0020); by CRIP (Center for Research for
592 Influenza Pathogenesis), a NIAID supported Center of Excellence for Influenza Research
593 and Surveillance (CEIRS, contract # HHSN272201400008C); by supplements to NIAID
594 grant U19AI135972 and DoD grant W81XWH-20-1-0270; by the generous support of the
595 JPB Foundation and the Open Philanthropy Project (research grant 2020-215611 (5384));
596 and by anonymous donors to A. García-Sastre. M. Ott acknowledges support through a
597 gift from the Roddenberry Foundation and by NIH 5DP1DA038043. We would like to
598 thank members of the Sirota Lab for useful discussion, Dr. Boris Oskotsky for technical
599 support, Edna Rodas for administrative support, Randy Albrecht for support with the BSL3
600 facility and procedures at the ISMMS, and Richard Cadagan for technical assistance.

601

602 **Author contributions**

603 B.L., T.O. and M.S. designed and coordinated the study. B.L. led the drug
604 repurposing efforts. G.A., K.Y., I.K., and C.L. helped with data analyses. SARS-CoV-2
605 virus assays were led by A.V.G., G.R.K., K.L., R.R., K.W., A.G.S., and M.O. All the
606 authors contributed to making figures, writing and editing the manuscript.

607

608 **Competing interests**

609 M.S. is on the advisory board of twoXAR. The García-Sastre Laboratory has
610 received research support from Pfizer, Senhwa Biosciences and 7Hills Pharma. A.G.S.

611 has consulting agreements for the following companies involving cash and/or stock:
612 Vivaldi Biosciences, Contrafect, 7Hills Pharma, Avimex, Vaxalto, Accurius and
613 Esperovax. Other authors declare no competing financial interests.

614

615

616 **References**

- 617 1. Ghisolfi S, Almås I, Sandefur JC, Carnap T von, Heitner J, Bold T. Predicted COVID-19
618 fatality rates based on age, sex, comorbidities and health system capacity. *BMJ Glob*
619 *Health*. 2020;5(9):e003094. doi:10.1136/bmjgh-2020-003094
- 620 2. Gupta A, Madhavan MV, Sehgal K, et al. Extrapulmonary manifestations of COVID-19. *Nat*
621 *Med*. 2020;26(7):1017-1032. doi:10.1038/s41591-020-0968-3
- 622 3. Zhang Q, Bastard P, Liu Z, et al. Inborn errors of type I IFN immunity in patients with life-
623 threatening COVID-19. *Science*. Published online September 24, 2020.
624 doi:10.1126/science.abd4570
- 625 4. Mehta P, McAuley DF, Brown M, Sanchez E, Tattersall RS, Manson JJ. COVID-19:
626 consider cytokine storm syndromes and immunosuppression. *The Lancet*.
627 2020;395(10229):1033-1034. doi:10.1016/S0140-6736(20)30628-0
- 628 5. Cava C, Bertoli G, Castiglioni I. A protein interaction map identifies existing drugs targeting
629 SARS-CoV-2. *BMC Pharmacol Toxicol*. 2020;21(1):65. doi:10.1186/s40360-020-00444-z
- 630 6. Cava C, Bertoli G, Castiglioni I. In Silico Discovery of Candidate Drugs against Covid-19.
631 *Viruses*. 2020;12(4):404. doi:10.3390/v12040404
- 632 7. Beigel JH, Tomashek KM, Dodd LE, et al. Remdesivir for the Treatment of Covid-19 —
633 Preliminary Report. *N Engl J Med*. 2020;0(0):null. doi:10.1056/NEJMoa2007764

- 634 8. Commissioner O of the. FDA Approves First Treatment for COVID-19. FDA. Published
635 October 22, 2020. Accessed October 22, 2020. [https://www.fda.gov/news-events/press-](https://www.fda.gov/news-events/press-announcements/fda-approves-first-treatment-covid-19)
636 [announcements/fda-approves-first-treatment-covid-19](https://www.fda.gov/news-events/press-announcements/fda-approves-first-treatment-covid-19)
- 637 9. RECOVERY Collaborative Group, Horby P, Lim WS, et al. Dexamethasone in Hospitalized
638 Patients with Covid-19 - Preliminary Report. *N Engl J Med*. Published online July 17, 2020.
639 doi:10.1056/NEJMoa2021436
- 640 10. DiMasi JA, Hansen RW, Grabowski HG. The price of innovation: new estimates of drug
641 development costs. *J Health Econ*. 2003;22(2):151-185. doi:10.1016/S0167-
642 6296(02)00126-1
- 643 11. Aronson JK. Old drugs – new uses. *Br J Clin Pharmacol*. 2007;64(5):563-565.
644 doi:10.1111/j.1365-2125.2007.03058.x
- 645 12. Gordon DE, Jang GM, Bouhaddou M, et al. A SARS-CoV-2 protein interaction map reveals
646 targets for drug repurposing. *Nature*. Published online April 30, 2020:1-13.
647 doi:10.1038/s41586-020-2286-9
- 648 13. Zhou Y, Hou Y, Shen J, Huang Y, Martin W, Cheng F. Network-based drug repurposing for
649 novel coronavirus 2019-nCoV/SARS-CoV-2. *Cell Discov*. 2020;6:1268-1281.
650 doi:10.1038/s41421-020-0153-3
- 651 14. Gordon DE, Hiatt J, Bouhaddou M, et al. Comparative host-coronavirus protein interaction
652 networks reveal pan-viral disease mechanisms. *Science*. Published online October 15,
653 2020. doi:10.1126/science.abe9403
- 654 15. Shah B, Modi P, Sagar SR. In silico studies on therapeutic agents for COVID-19: Drug
655 repurposing approach. *Life Sci*. 2020;252:117652. doi:10.1016/j.lfs.2020.117652
- 656 16. Sirota M, Dudley JT, Kim J, et al. Discovery and Preclinical Validation of Drug Indications
657 Using Compendia of Public Gene Expression Data. *Sci Transl Med*. 2011;3(96):96ra77.
658 doi:10.1126/scitranslmed.3001318

- 659 17. Dudley JT, Sirota M, Shenoy M, et al. Computational Repositioning of the Anticonvulsant
660 Topiramate for Inflammatory Bowel Disease. *Sci Transl Med.* 2011;3(96):96ra76.
661 doi:10.1126/scitranslmed.3002648
- 662 18. Cho HG, Fiorentino D, Lewis M, Sirota M, Sarin KY. Identification of alpha-adrenergic
663 agonists as potential therapeutic agents for dermatomyositis through drug-repurposing
664 using public expression datasets. *J Invest Dermatol.* 2016;136(7):1517-1520.
665 doi:10.1016/j.jid.2016.03.001
- 666 19. Chen B, Wei W, Ma L, et al. Computational Discovery of Niclosamide Ethanolamine, a
667 Repurposed Drug Candidate That Reduces Growth of Hepatocellular Carcinoma Cells
668 In Vitro and in Mice by Inhibiting Cell Division Cycle 37 Signaling. *Gastroenterology.*
669 2017;152(8):2022-2036. doi:10.1053/j.gastro.2017.02.039
- 670 20. Chen B, Ma L, Paik H, et al. Reversal of cancer gene expression correlates with drug
671 efficacy and reveals therapeutic targets. *Nat Commun.* 2017;8:16022.
672 doi:10.1038/ncomms16022
- 673 21. Jahchan NS, Dudley JT, Mazur PK, et al. A Drug Repositioning Approach Identifies Tricyclic
674 Antidepressants as Inhibitors of Small Cell Lung Cancer and Other Neuroendocrine
675 Tumors. *Cancer Discov.* 2013;3(12):1364-1377. doi:10.1158/2159-8290.CD-13-0183
- 676 22. Le BL, Iwatani S, Wong RJ, Stevenson DK, Sirota M. Computational discovery of
677 therapeutic candidates for preventing preterm birth. *JCI Insight.* 5(3).
678 doi:10.1172/jci.insight.133761
- 679 23. Xing J, Shankar R, Drelich A, et al. Reversal of Infected Host Gene Expression Identifies
680 Repurposed Drug Candidates for COVID-19. *bioRxiv.* Published online April 9,
681 2020:2020.04.07.030734. doi:10.1101/2020.04.07.030734
- 682 24. Blanco-Melo D, Nilsson-Payant BE, Liu W-C, et al. Imbalanced Host Response to SARS-
683 CoV-2 Drives Development of COVID-19. *Cell.* Published online May 13, 2020.
684 doi:10.1016/j.cell.2020.04.026

- 685 25. Hoffmann M, Kleine-Weber H, Schroeder S, et al. SARS-CoV-2 Cell Entry Depends on
686 ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell*.
687 2020;181(2):271-280.e8. doi:10.1016/j.cell.2020.02.052
- 688 26. Lamers MM, Beumer J, Vaart J van der, et al. SARS-CoV-2 productively infects human gut
689 enterocytes. *Science*. 2020;369(6499):50-54. doi:10.1126/science.abc1669
- 690 27. Xiong Y, Liu Y, Cao L, et al. Transcriptomic characteristics of bronchoalveolar lavage fluid
691 and peripheral blood mononuclear cells in COVID-19 patients. *Emerg Microbes Infect*.
692 2020;9(1):761-770. doi:10.1080/22221751.2020.1747363
- 693 28. Lamb J, Crawford ED, Peck D, et al. The Connectivity Map: using gene-expression
694 signatures to connect small molecules, genes, and disease. *Science*.
695 2006;313(5795):1929-1935. doi:10.1126/science.1132939
- 696 29. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: A knowledge-
697 based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S*
698 *A*. 2005;102(43):15545-15550. doi:10.1073/pnas.0506580102
- 699 30. Mootha VK, Lindgren CM, Eriksson K-F, et al. PGC-1 α -responsive genes involved in
700 oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat Genet*.
701 2003;34(3):267-273. doi:10.1038/ng1180
- 702 31. Liao M, Liu Y, Yuan J, et al. Single-cell landscape of bronchoalveolar immune cells in
703 patients with COVID-19. *Nat Med*. 2020;26(6):842-844. doi:10.1038/s41591-020-0901-9
- 704 32. Wishart DS, Feunang YD, Guo AC, et al. DrugBank 5.0: a major update to the DrugBank
705 database for 2018. *Nucleic Acids Res*. 2018;46(D1):D1074-D1082.
706 doi:10.1093/nar/gkx1037
- 707 33. Keiser MJ, Roth BL, Armbruster BN, Ernsberger P, Irwin JJ, Shoichet BK. Relating protein
708 pharmacology by ligand chemistry. *Nat Biotechnol*. 2007;25(2):197-206.
709 doi:10.1038/nbt1284

- 710 34. Amanat F, White KM, Miorin L, et al. An In Vitro Microneutralization Assay for SARS-CoV-2
711 Serology and Drug Screening. *Curr Protoc Microbiol.* 2020;58(1):e108.
712 doi:10.1002/cpmc.108
- 713 35. Seeman P, Kapur S. Schizophrenia: more dopamine, more D2 receptors. *Proc Natl Acad*
714 *Sci U S A.* 2000;97(14):7673-7675. doi:10.1073/pnas.97.14.7673
- 715 36. Hafidh RR, Abdulmir AS, Abu Bakar F, Sekawi Z, Jahansheri F, Jalilian FA. Novel antiviral
716 activity of mung bean sprouts against respiratory syncytial virus and herpes simplex virus
717 -1: an in vitro study on virally infected Vero and MRC-5 cell lines. *BMC Complement Altern*
718 *Med.* 2015;15. doi:10.1186/s12906-015-0688-2
- 719 37. Berezin V, Abdulkhakimova D, Trenochnikova L, et al. Antiviral activities of extremophilic
720 actinomycetes extracts from Kazakhstan's unique ecosystems against influenza viruses
721 and paramyxoviruses. *Virologia J.* 2019;16. doi:10.1186/s12985-019-1254-1
- 722 38. PubChem. Clofazimine. Accessed May 21, 2020.
723 <https://pubchem.ncbi.nlm.nih.gov/compound/2794>
- 724 39. Cholo MC, Steel HC, Fourie PB, Germishuizen WA, Anderson R. Clofazimine: current
725 status and future prospects. *J Antimicrob Chemother.* 2012;67(2):290-298.
726 doi:10.1093/jac/dkr444
- 727 40. Yuan S, Yin X, Meng X, et al. Clofazimine is a broad-spectrum coronavirus inhibitor that
728 antagonizes SARS-CoV-2 replication in primary human cell culture and hamsters. *Res Sq.*
729 Published online October 7, 2020. doi:10.21203/rs.3.rs-86169/v1
- 730 41. PubChem. Bacampicillin. Accessed May 21, 2020.
731 <https://pubchem.ncbi.nlm.nih.gov/compound/441397>
- 732 42. Himmelstein DS, Lizee A, Hessler C, et al. Systematic integration of biomedical knowledge
733 prioritizes drugs for repurposing. Valencia A, ed. *eLife.* 2017;6:e26726.
734 doi:10.7554/eLife.26726

- 735 43. Subramanian A, Narayan R, Corsello SM, et al. A Next Generation Connectivity Map: L1000
736 platform and the first 1,000,000 profiles. *Cell*. 2017;171(6):1437-1452.e17.
737 doi:10.1016/j.cell.2017.10.049
- 738 44. Maglott D, Ostell J, Pruitt KD, Tatusova T. Entrez Gene: gene-centered information at NCBI.
739 *Nucleic Acids Res*. 2005;33(Database Issue):D54-D58. doi:10.1093/nar/gki031
- 740 45. Luis García de Guadiana Romualdo, Mulero MDR, Olivo MH, et al. Circulating levels of
741 GDF-15 and calprotectin for prediction of in-hospital mortality in COVID-19 patients: A
742 case series. *J Infect*. Published online August 12, 2020. doi:10.1016/j.jinf.2020.08.010
- 743 46. WHO | WHO Model Lists of Essential Medicines. WHO. Accessed October 19, 2020.
744 <http://www.who.int/medicines/publications/essentialmedicines/en/>
- 745 47. Poroikov V, Druzhilovskiy D. Chapter 1 - Drug Repositioning: New Opportunities for Older
746 Drugs. In: Roy K, ed. *In Silico Drug Design*. Academic Press; 2019:3-17.
747 doi:10.1016/B978-0-12-816125-8.00001-8
- 748 48. Musa A, Ghoraie LS, Zhang S-D, et al. A review of connectivity map and computational
749 approaches in pharmacogenomics. *Brief Bioinform*. 2018;19(3):506-523.
750 doi:10.1093/bib/bbw112
- 751 49. Zhou X, Wang M, Katsyv I, Irie H, Zhang B. EMUDRA: Ensemble of Multiple Drug
752 Repositioning Approaches to improve prediction accuracy. *Bioinformatics*.
753 2018;34(18):3151-3159. doi:10.1093/bioinformatics/bty325
- 754 50. Cheng J, Yang L, Kumar V, Agarwal P. Systematic evaluation of connectivity map for
755 disease indications. *Genome Med*. 2014;6:95. doi:10.1186/s13073-014-0095-1
- 756 51. National Genomics Data Center Members and Partners. Database Resources of the
757 National Genomics Data Center in 2020. *Nucleic Acids Res*. 2020;48(D1):D24-D33.
758 doi:10.1093/nar/gkz913
- 759 52. Wang Y, Song F, Zhu J, et al. GSA: Genome Sequence Archive<sup/>. *Genomics*
760 *Proteomics Bioinformatics*. 2017;15(1):14-18. doi:10.1016/j.gpb.2017.01.001

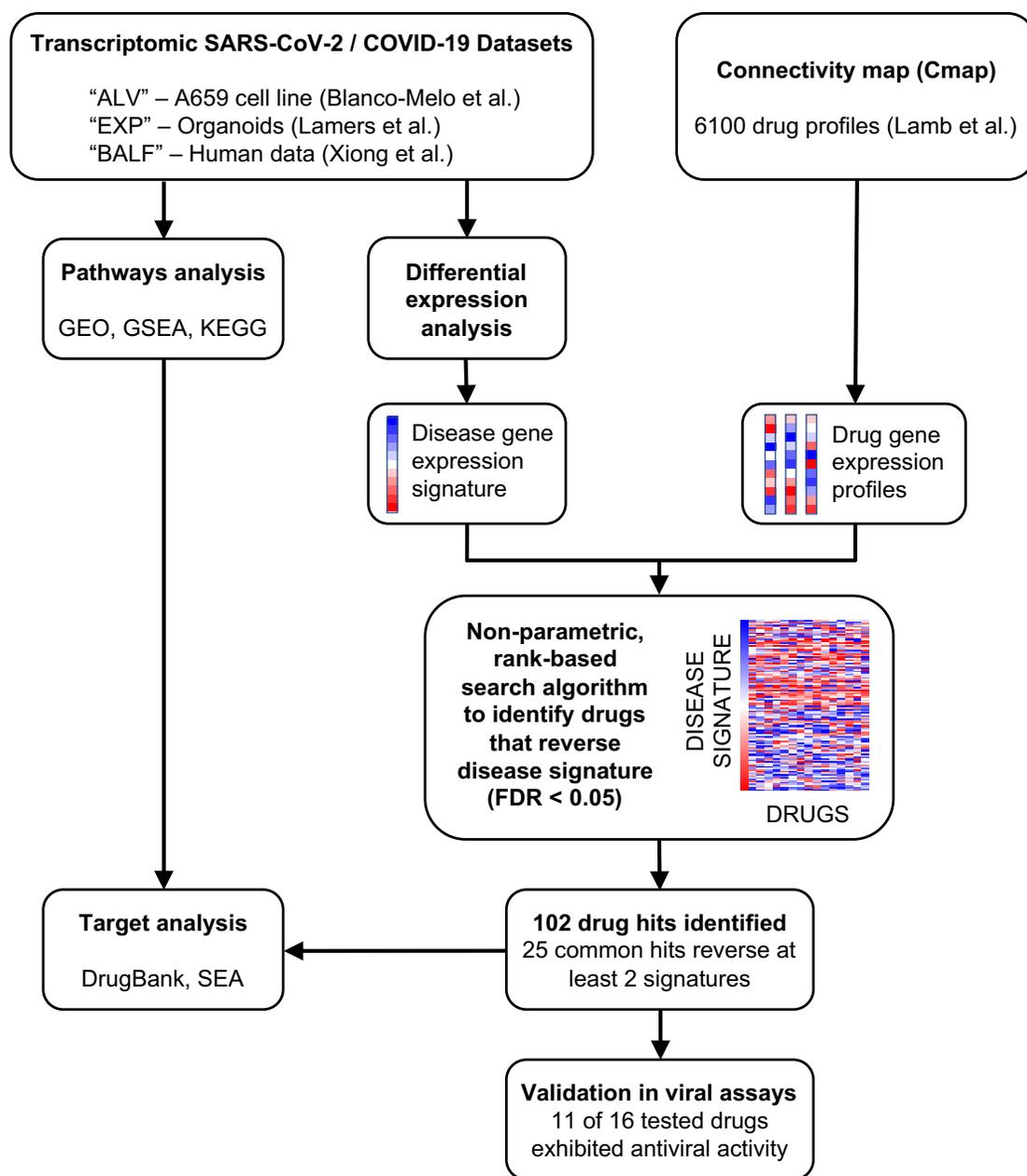
- 761 53. Liberzon A, Subramanian A, Pinchback R, Thorvaldsdóttir H, Tamayo P, Mesirov JP.
762 Molecular signatures database (MSigDB) 3.0. *Bioinformatics*. 2011;27(12):1739-1740.
763 doi:10.1093/bioinformatics/btr260
- 764 54. Dennis G, Sherman BT, Hosack DA, et al. DAVID: Database for Annotation, Visualization,
765 and Integrated Discovery. *Genome Biol*. 2003;4(9):R60.
- 766 55. Hafemeister C, Satija R. Normalization and variance stabilization of single-cell RNA-seq
767 data using regularized negative binomial regression. *Genome Biol*. 2019;20.
768 doi:10.1186/s13059-019-1874-1

769

770

771

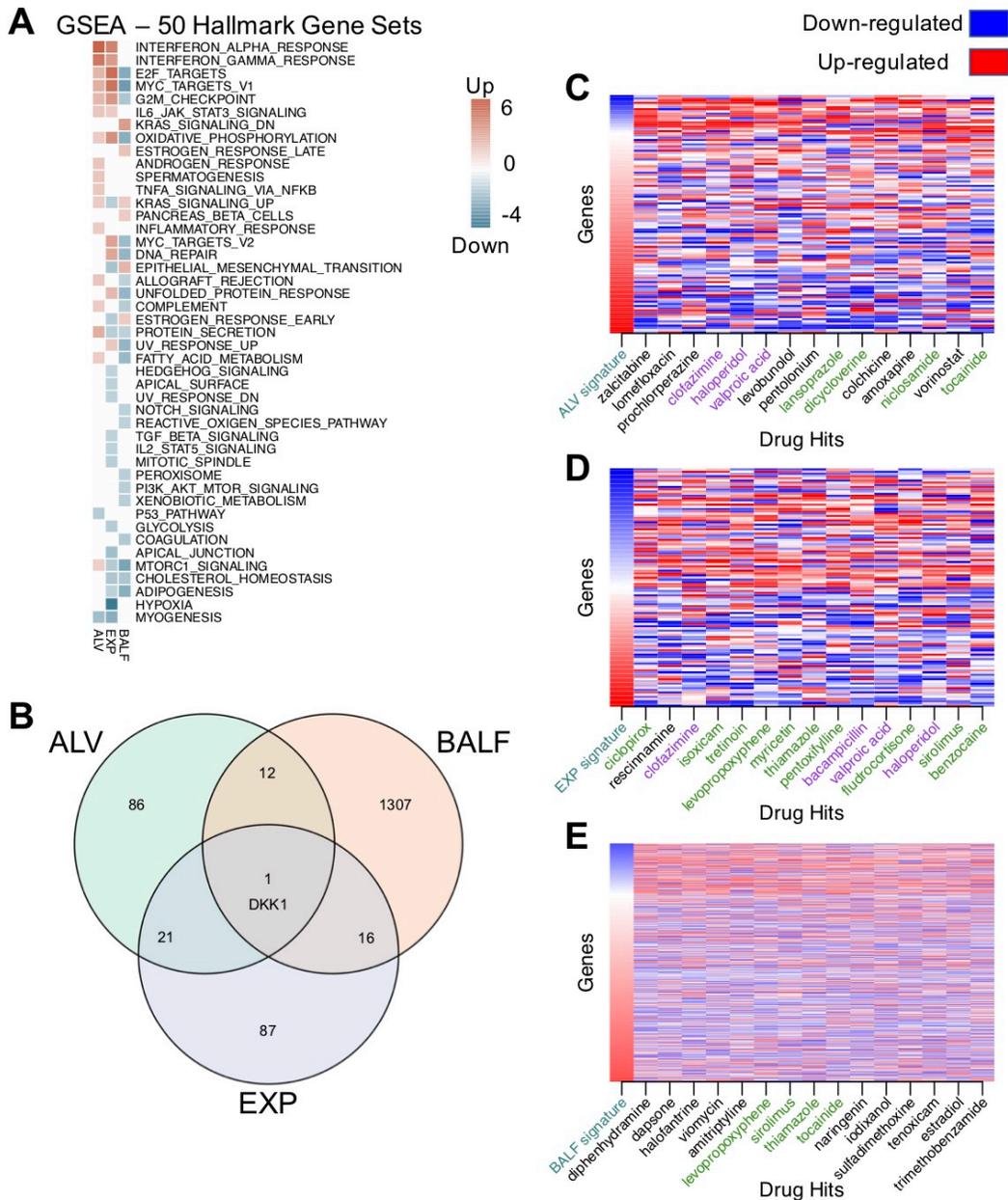
772



773

774 **Figure 1. COVID-19 transcriptomics-based bioinformatics approach for drug repositioning.**

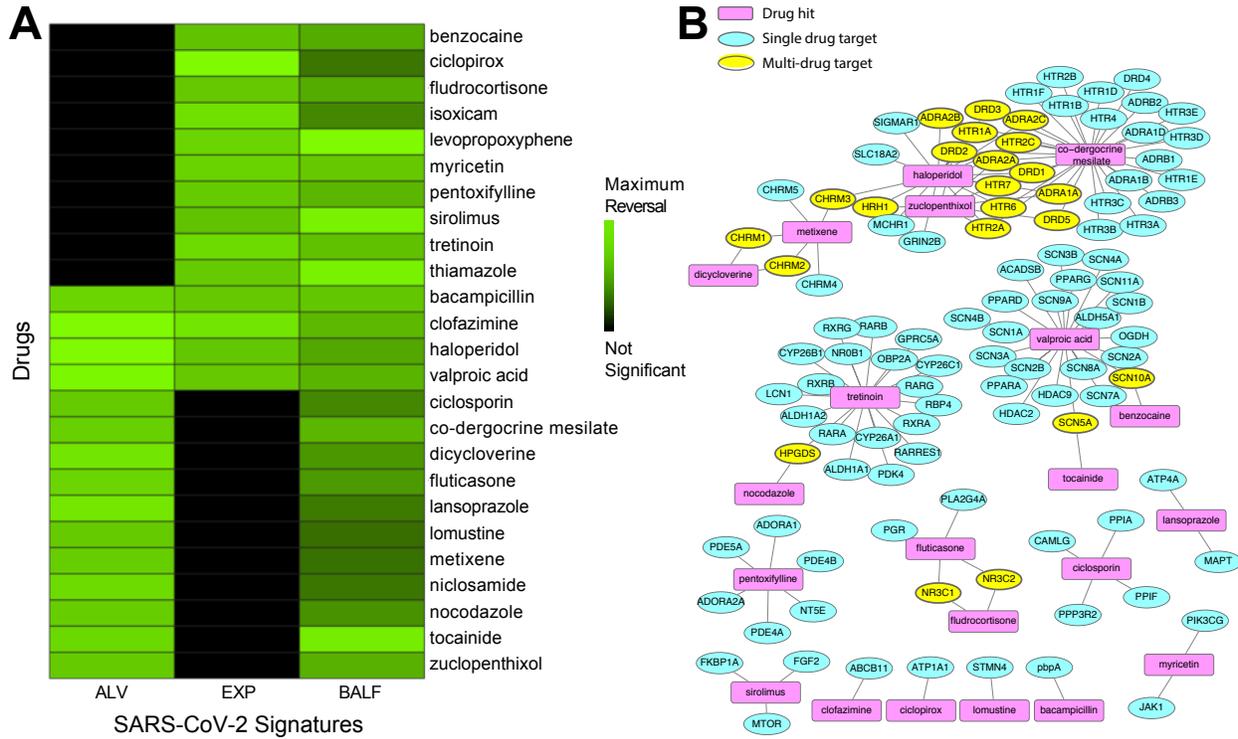
775 We generated lists of statistically significant differentially expressed genes from the analysis of
 776 three published studies of SARS-CoV-2 and COVID-19. The drug repositioning computational
 777 pipeline compares the ranked differential expression of the COVID-19 disease signature with that
 778 of drug profiles from CMap. A reversal score based on the Kolmogorov-Smirnov statistic is
 779 generated for each disease-drug pair. If a drug profile significantly (FDR < 0.05) reverses the
 780 disease signature, then the drug could be therapeutic for the disease. Across all datasets, a total
 781 of 102 drugs have been identified as potentially therapeutic for COVID-19. Twenty-five drugs were
 782 identified in analyses of at least two of the three datasets. We further conducted pathways
 783 analyses and targeted analyses on the results, focusing on the 25 shared hits. Finally, we
 784 validated sixteen of our top predicted hits in live SARS-CoV-2 antiviral assays.



785

786 Figure 2. **SARS-CoV-2 differential gene expression signatures reversed by drug profiles**
 787 **from CMap. (A)** Enrichment analysis using GSEA reveals common pathways among input
 788 signatures. **(B)** DEG overlap from input signatures. Only 1 gene, DKK1, was shared by all 3
 789 signatures. **(C)** Top 15 drug profiles reversing the ALV signature (109 genes). For each column,
 790 the gene expression values were ranked, with rank 1 being the most up-regulated gene (in red)
 791 and the maximum rank (109 for ALV) being the most down-regulated gene (in blue). Drug
 792 names highlighted in green were hits for a second signature, and drug hits highlighted in purple
 793 reversed all three signatures. **(D)** Top 15 drug profiles reversing the EXP signature (108 genes).
 794 **(E)** Top 15 drug profiles reversing the BALF signature (941 genes).

795



797

798 Figure 3. **Common therapeutic hits from drug repurposing pipeline applied to SARS-CoV-**

799 **2 signatures.** (A) Drug profiles from CMap significantly reversed signatures from the ALV,

800 BALF, and EXP signatures. 25 of the drugs were significant in at least 2 of the signatures. (B)

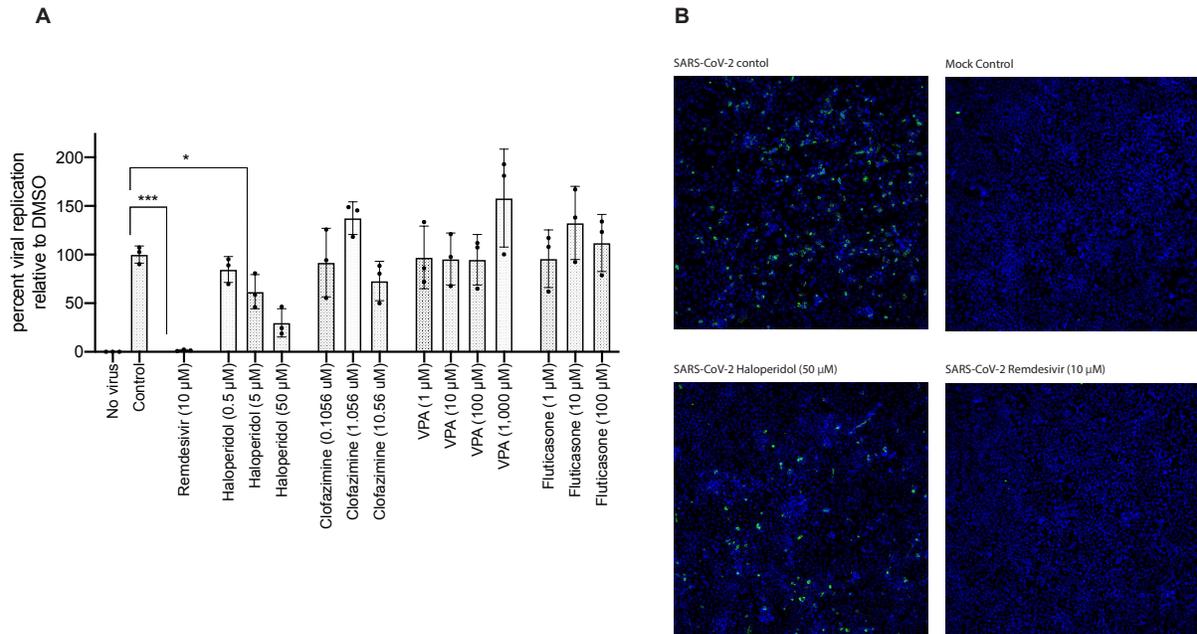
801 Drug-protein target network. For the 25 drugs that reversed at least 2 of the signatures, target

802 information was gathered from DrugBank to identify clusters of drugs from shared targets.

803

804

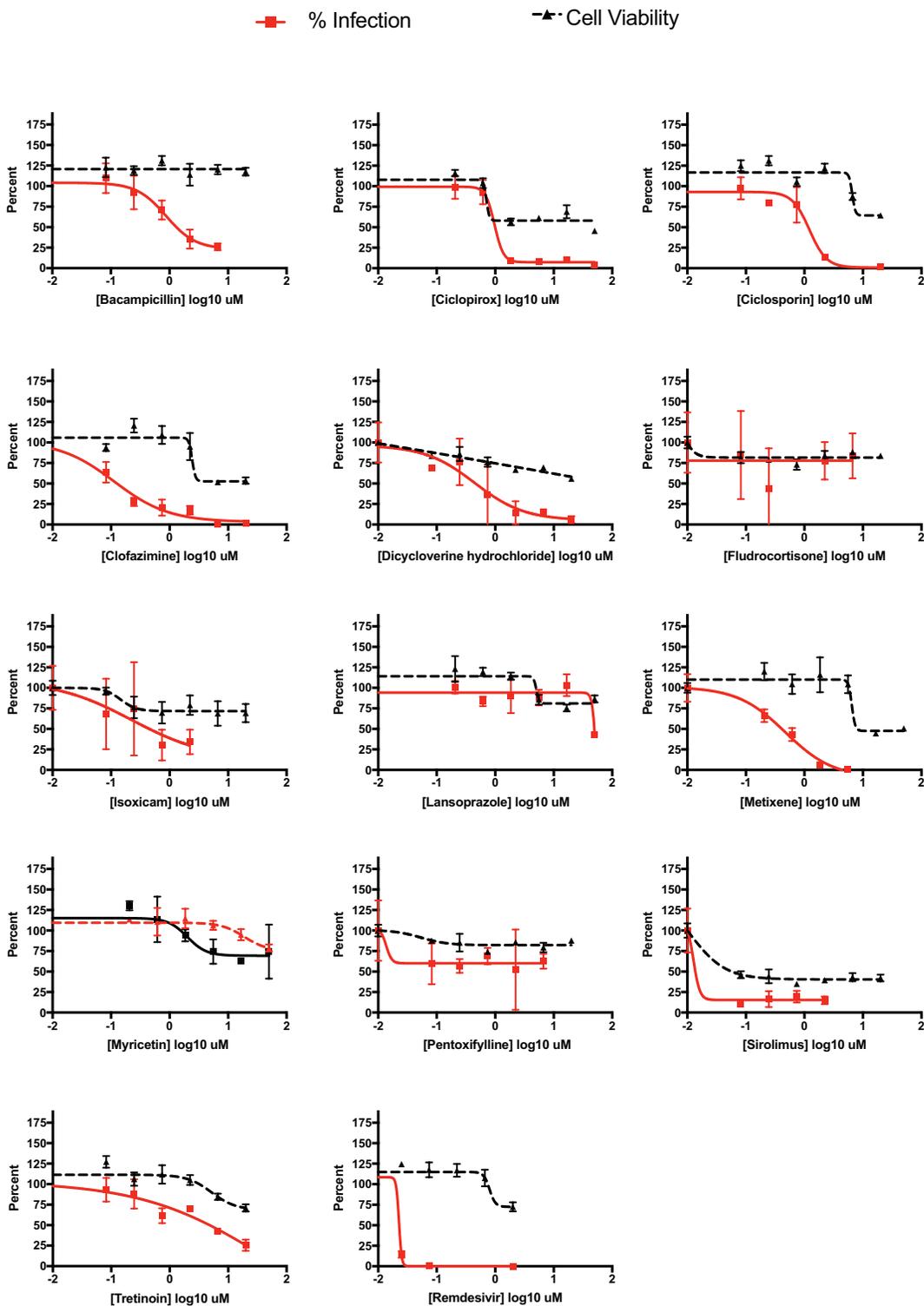
805



806

807 **Figure 4. Haloperidol inhibits viral replication of SARS-CoV-2 in the Calu-3 lung cell line.**
 808 **(A)** Calu-3 cells were infected with SARS-CoV-2 at an MOI of 0.05 for 72h. Viral replication
 809 levels were determined by RT-qPCR from supernatant RNA using specific primers for the E
 810 gene. Viral RNA levels relative to DMSO are graphed. Error bars represent 3 or 4 independent
 811 experiments. One-way ANOVA analysis was used to determine significance. **(B)** Microscopy:
 812 Calu-3 cells were infected with SARS-CoV-2 at an MOI of 0.05 for 72h. Cells were fixed with
 813 paraformaldehyde and used for immunofluorescence analysis with dsRNA antibody (SCICONS)
 814 and DAPI stain. Images were acquired and analyzed using ImageXpress Micro Confocal High-
 815 Content Imaging System.

816



817

818 **Figure 5. Viral inhibition and cell viability tests of 13 compounds in 293T-ACE2 cell**
 819 **assays.** Several drugs inhibit viral infectivity. Red, viral infectivity (anti-NP); black, cell viability.
 820 The lack of a dose response in cell viability probably reflects cytostatic and not cytotoxic effects.
 821 Data are mean \pm s.d.; $n = 3$ biologically independent samples for cell viability data.

Drug hit	Description (current uses)	ALV Reversal Score	EXP Reversal Score	BALF Reversal Score
Bacampicillin	Antibiotic	0.789	0.790	0.596
Benzocaine	Anesthetic	n.s.	0.766	0.546
Ciclopirox	Antifungal	n.s.	1	0.361
Ciclosporin	Immunosuppressant (RA, psoriasis, Crohn's)	0.756	n.s.	0.409
Clofazimine	Antimycobacterial (leprosy)	0.946	0.893	0.558
Co-dergocrine mesilate	Ergoid mesylate (dementia, Alzheimer's, stroke)	0.775	n.s.	0.553
Dicycloverine	Antispasmodic (IBS)	0.847	n.s.	0.461
Fludrocortisone	Corticosteroid	n.s.	0.782	0.519
Fluticasone	Steroid (asthma, COPD)	0.790	n.s.	0.463
Haloperidol	Antipsychotic (schizophrenia)	0.937	0.773	0.507
Isoxicam	NSAID	n.s.	0.873	0.410
Lansoprazole	Proton-pump inhibitor (acid reflux)	0.856	n.s.	0.370
Levopropoxyphene	Antitussive	n.s.	0.835	0.770
Lomustine	Antineoplastic (Hodgkin's disease, brain tumors)	0.748	n.s.	0.338
Metixene	Anticholinergic (Parkinson's)	0.759	n.s.	0.344
Myricetin	Flavonoid	n.s.	0.823	0.603
Niclosamide	Anthelmintic (tapeworms)	0.812	n.s.	0.360
Nocodazole	Antineoplastic	0.766	n.s.	0.439
Pentoxifylline	Vasodilatory and anti-inflammatory (claudication)	n.s.	0.791	0.552
Sirolimus	Immunosuppressive	n.s.	0.768	0.729

Thiamazole	Antithyroid agent (Graves disease)	n.s.	0.796	0.724
Tocainide	Antiarrhythmic	0.798	n.s.	0.714
Tretinoin	Vitamin A derivative (acne, acute promyelocytic leukemia)	n.s.	0.854	0.579
Valproic acid	Anticonvulsant (seizures, bipolar disorder)	0.917	0.786	0.546
Zuclopenthixol	Antipsychotic (schizophrenia)	0.754	n.s.	0.535

822 Table 1. Therapeutic hits reversing at least 2 of input SARS-CoV-2 signatures. A wide
823 variety of drugs were identified by the analysis of multiple signatures. Drug reversal
824 scores are normalized for each signature; drug entries marked "n.s." were not significant
825 for reversing that signature.
826

Figures

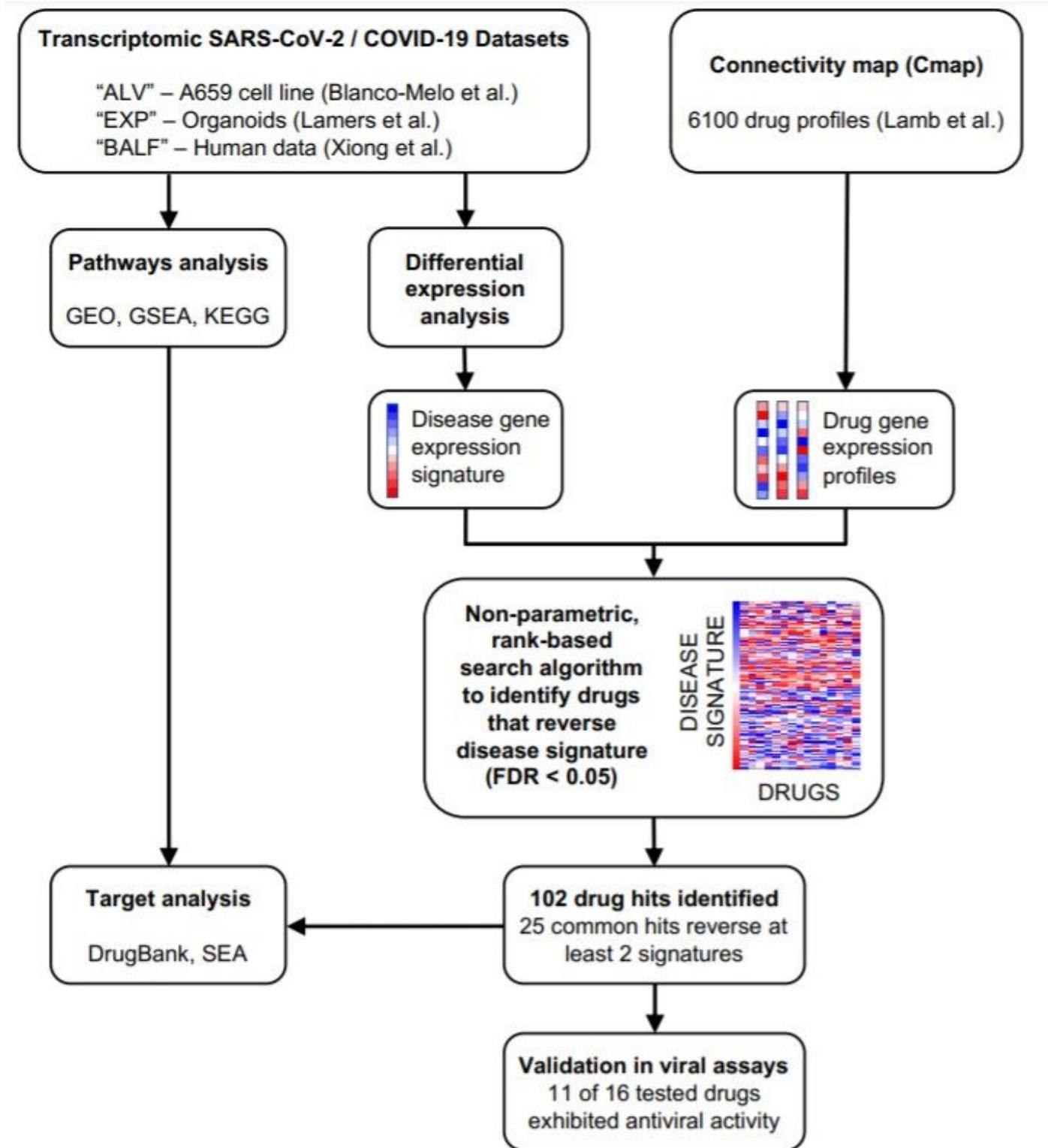
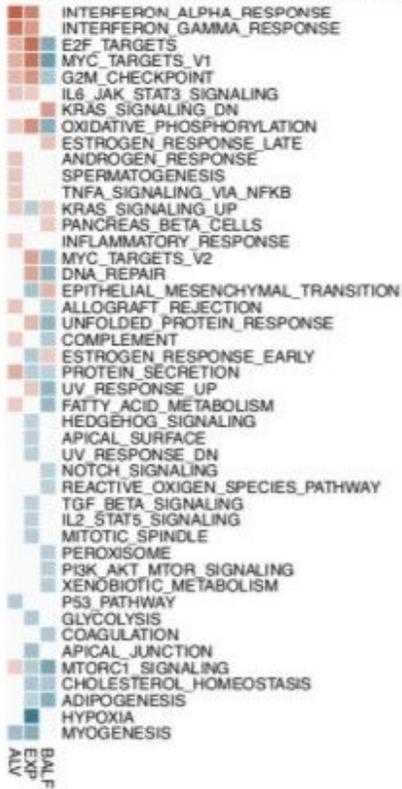


Figure 1

COVID-19 transcriptomics-based bioinformatics approach for drug repositioning. We generated lists of statistically significant differentially expressed genes from the analysis of three published studies of SARS-CoV-2 and COVID-19. The drug repositioning computational pipeline compares the ranked

differential expression of the COVID-19 disease signature with that of drug profiles from CMap. A reversal score based on the Kolmogorov-Smirnov statistic is generated for each disease-drug pair. If a drug profile significantly (FDR < 0.05) reverses the disease signature, then the drug could be therapeutic for the disease. Across all datasets, a total of 102 drugs have been identified as potentially therapeutic for COVID-19. Twenty-five drugs were identified in analyses of at least two of the three datasets. We further conducted pathways analyses and targeted analyses on the results, focusing on the 25 shared hits. Finally, we validated sixteen of our top predicted hits in live SARS-CoV-2 antiviral assays.

A GSEA – 50 Hallmark Gene Sets



Down-regulated ■
Up-regulated ■

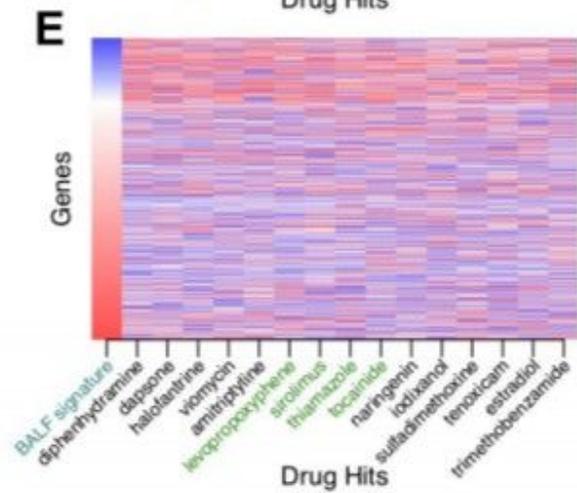
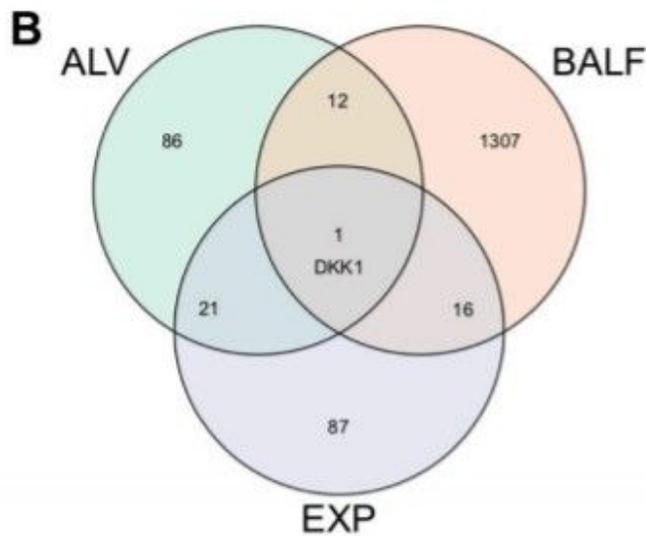
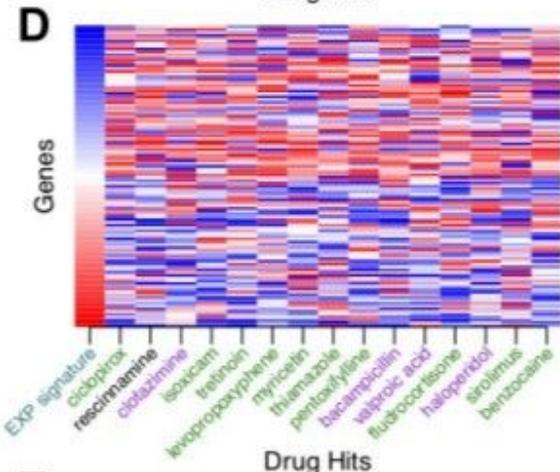
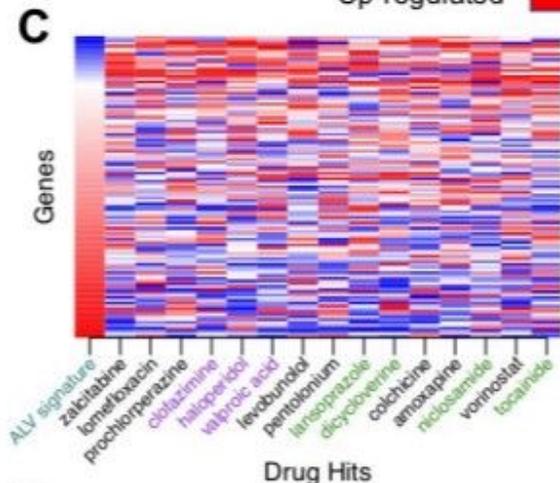


Figure 2

SARS-CoV-2 differential gene expression signatures reversed by drug profiles from CMap. (A) Enrichment analysis using GSEA reveals common pathways among input signatures. (B) DEG overlap from input signatures. Only 1 gene, DKK1, was shared by all 3 signatures. (C) Top 15 drug profiles reversing the ALV signature (109 genes). For each column, the gene expression values were ranked, with rank 1 being the most up-regulated gene (in red) and the maximum rank (109 for ALV) being the most down-regulated gene (in blue). Drug names highlighted in green were hits for a second signature, and drug hits highlighted in purple reversed all three signatures. (D) Top 15 drug profiles reversing the EXP signature (108 genes). (E) Top 15 drug profiles reversing the BALF signature (941 genes).

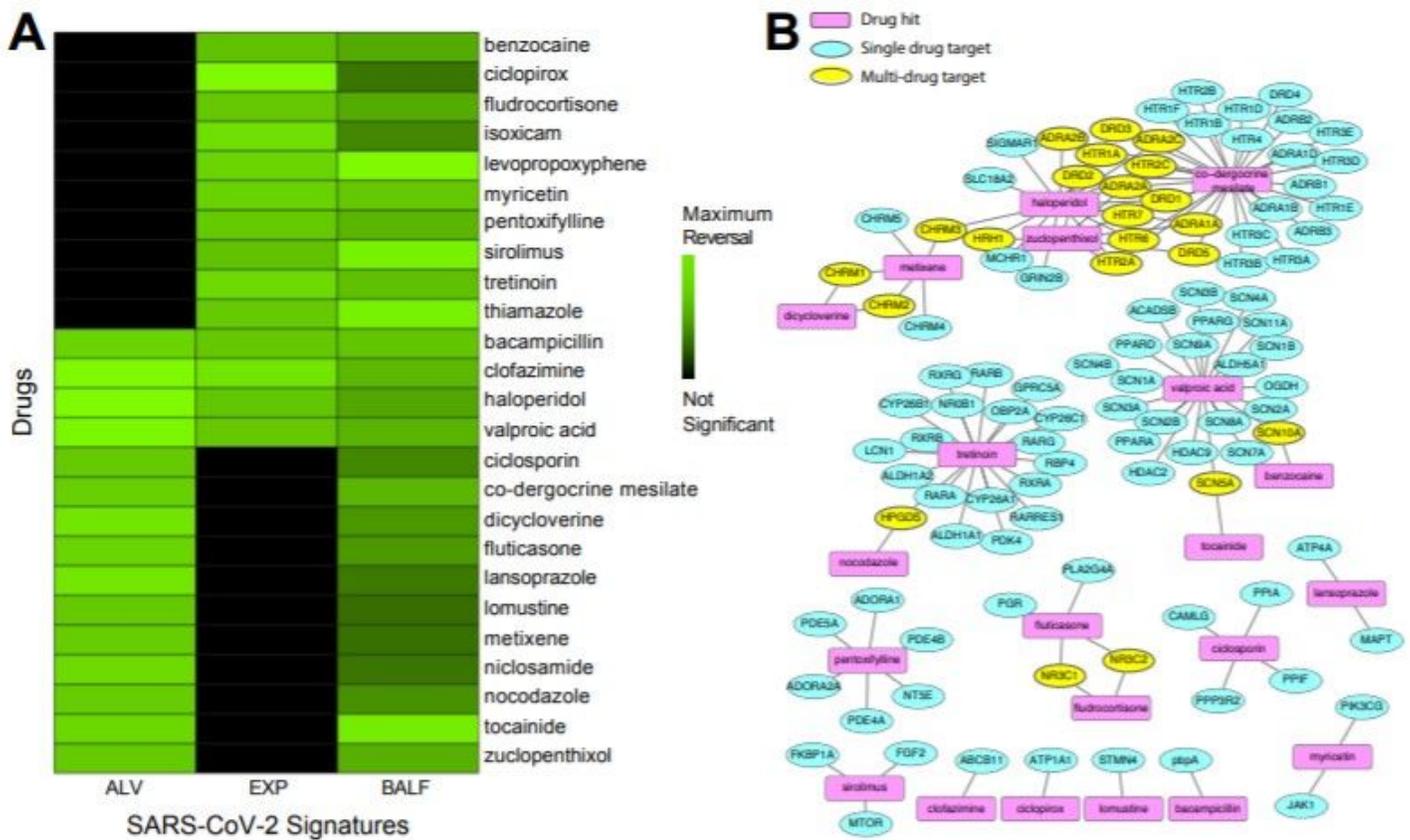


Figure 3

Common therapeutic hits from drug repurposing pipeline applied to SARS-CoV-2 signatures. (A) Drug profiles from CMap significantly reversed signatures from the ALV, BALF, and EXP signatures. 25 of the drugs were significant in at least 2 of the signatures. (B) Drug-protein target network. For the 25 drugs that reversed at least 2 of the signatures, target information was gathered from DrugBank to identify clusters of drugs from shared targets.

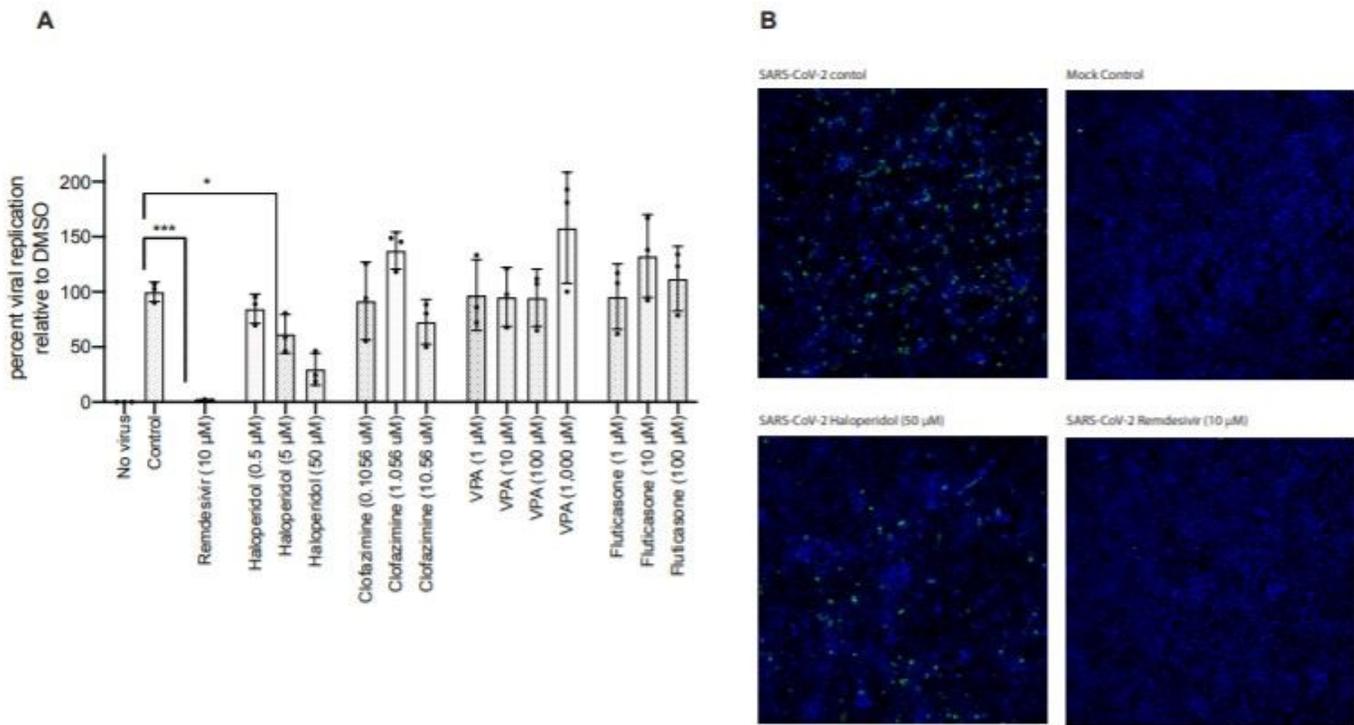


Figure 4

Haloperidol inhibits viral replication of SARS-CoV-2 in the Calu-3 lung cell line. (A) Calu-3 cells were infected with SARS-CoV-2 at an MOI of 0.05 for 72h. Viral replication levels were determined by RT-qPCR from supernatant RNA using specific primers for the E gene. Viral RNA levels relative to DMSO are graphed. Error bars represent 3 or 4 independent experiments. One-way ANOVA analysis was used to determine significance. (B) Microscopy: Calu-3 cells were infected with SARS-CoV-2 at an MOI of 0.05 for 72h. Cells were fixed with paraformaldehyde and used for immunofluorescence analysis with dsRNA antibody (SCICONS) and DAPI stain. Images were acquired and analyzed using ImageXpress Micro Confocal High Content Imaging System.

■ % Infection ▲ Cell Viability

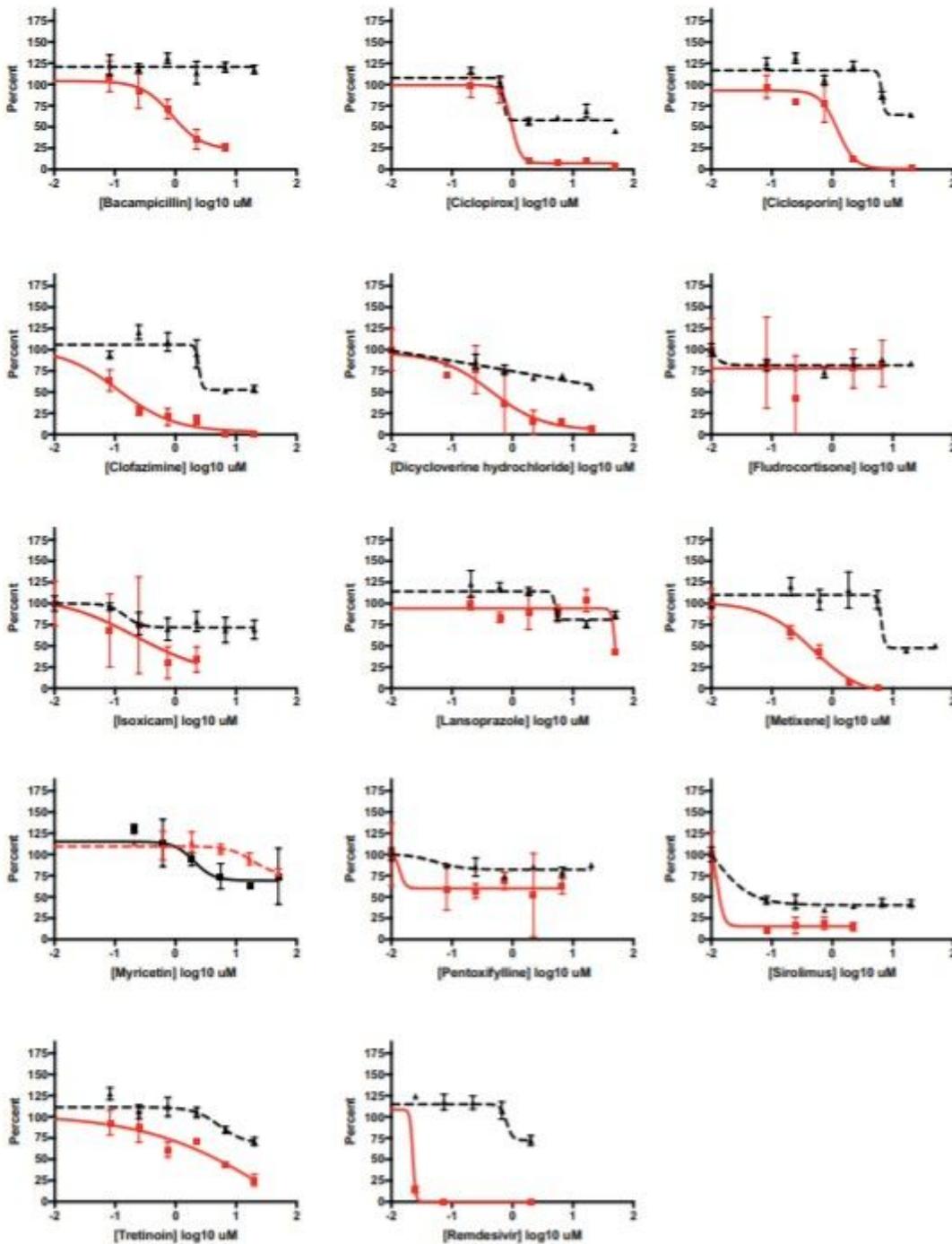


Figure 5

Viral inhibition and cell viability tests of 13 compounds in 293T-ACE2 cell assays. Several drugs inhibit viral infectivity. Red, viral infectivity (anti-NP); black, cell viability. The lack of a dose response in cell viability probably reflects cytostatic and not cytotoxic effects. Data are mean \pm s.d.; n = 3 biologically independent samples for cell viability data.

Supplementary Files

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

- [SRsupplement20210323.pdf](#)
- [SupplementalTables.xlsx](#)