

- 39 ¹⁵Department of Medical Sciences, Uppsala University, Uppsala, Sweden
40 ¹⁶Berlin Institute of Health, Charité University Medicine Berlin, Berlin, Germany
41 ¹⁷MRC Epidemiology Unit, University of Cambridge, Cambridge, UK
42 ¹⁸Institute of Health Informatics, University College London, London, UK
43 ¹⁹Health Data Research, University College London, London, UK
44 ²⁰British Heart Foundation Centre of Research Excellence, University of Cambridge, Cambridge,
45 UK
46 ²¹Health Data Research UK Cambridge, Wellcome Genome Campus and University of
47 Cambridge, Cambridge, UK
48 ²²Rutherford Fund Fellow, Department of Public Health and Primary Care, University of
49 Cambridge, Cambridge, UK
50 ²³MRC Biostatistics Unit, University of Cambridge, Cambridge, UK
51 ²⁴Centre for Inflammatory Disease, Dept of Immunology and Inflammation, Imperial College,
52 London, UK
53 ²⁵Health Data Research UK, UK
54 ²⁶VA Informatics and Computing Infrastructure, VA Salt Lake City Health Care System, Salt
55 Lake City, UT, USA
56 ²⁷Department of Internal Medicine, Epidemiology, University of Utah, Salt Lake City, UT, USA
57 ²⁸Department of Epidemiology and Biostatistics, University of Arizona, Tucson, AZ, USA
58 ²⁹Phoenix VA Health Care System, Phoenix, AZ, USA
59 ³⁰Center for Population Genomics, Massachusetts Veterans Epidemiology Research and
60 Information Center (MAVERIC), VA Boston Healthcare System, Boston, MA, USA
61 ³¹Cardiology, VA Boston Healthcare System, Boston, MA, USA
62 ³²Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA
63 ³³Epidemiology Research and Information Center (ERIC), VA Palo Alto Health Care System,
64 Palo Alto, CA, USA
65 ³⁴Department of Medicine, Stanford University School of Medicine, Palo Alto, CA, USA
66 ³⁵MIRECC, Durham VA Medical Center, Durham, NC, USA
67 ³⁶Department of Psychiatry and Behavioral Sciences, Duke University School of Medicine,
68 Durham, NC, USA
69 ³⁷Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA
70 ³⁸Office of Research and Development, Department of Veterans Affairs, Washington, DC, USA
71 ³⁹VA Tennessee Valley Healthcare System, Nashville, TN, USA
72 ⁴⁰Nephrology & Hypertension, Vanderbilt University, Nashville, TN, USA
73 ⁴¹The Corporal Michael J. Crescenz VA Medical Center, Philadelphia, PA, USA
74 ⁴²Department of Medicine, Perelman School of Medicine, University of Pennsylvania,
75 Philadelphia, PA, USA
76 ⁴³Atlanta VA Health Care System, Decatur, GA, USA
77 ⁴⁴Department of Epidemiology, Emory University Rollins School of Public Health, Atlanta, GA,
78 USA

79 ⁴⁵Medicine, Cardiovascular, VA Boston Healthcare System and Brigham & Women's Hospital,
80 Boston, MA, USA

81 ⁴⁶Department of Veterans Affairs, Tennessee Valley Healthcare System, Vanderbilt University,
82 Nashville, TN, USA

83 ⁴⁷Medicine, Epidemiology, Vanderbilt Genetics Institute, Vanderbilt University Medical Center,
84 Nashville, TN, USA

85 ⁴⁸Division of Aging, Brigham and Women's Hospital, Harvard Medical School, Boston, MA,
86 USA

87 ⁴⁹National Institute for Health Research Blood and Transplant Research Unit in Donor Health
88 and Genomics, University of Cambridge, Cambridge, UK

89 ⁵⁰National Institute for Health Research Cambridge Biomedical Research Centre, University of
90 Cambridge and Cambridge University Hospitals, Cambridge, UK

91

92

93 **Authors for correspondence:** asb38@medschl.cam.ac.uk (A.S.B.),

94 jpcasasromero@bwh.harvard.edu (J.P.C.)

95 **Key words:** Mendelian randomization, drug repurposing, COVID-19, actionable druggable

96 genome

97

98 **Abstract**

99 Drug repurposing provides a rapid approach to meet the urgent need for therapeutics to address
100 COVID-19. To identify therapeutic targets relevant to COVID-19, we conducted Mendelian
101 randomization (MR) analyses, deriving genetic instruments based on transcriptomic and
102 proteomic data for 1,263 actionable proteins that are targeted by approved drugs or in clinical
103 phase of drug development. Using summary statistics from the Host Genetics Initiative and the
104 Million Veteran Program, we studied 7,554 patients hospitalized with COVID-19 and >1 million
105 controls. We found significant Mendelian randomization results for three proteins (ACE2:
106 $P=1.6\times 10^{-6}$, IFNAR2: $P=9.8\times 10^{-11}$, and IL-10RB: $P=1.9\times 10^{-14}$) using *cis*-eQTL genetic
107 instruments that also had strong evidence for colocalization with COVID-19 hospitalization. To
108 disentangle the shared eQTL signal for *IL10RB* and *IFNAR2*, we conducted phenome-wide
109 association scans and pathway enrichment analysis, which suggested that *IFNAR2* is more likely
110 to play a role in COVID-19 hospitalization. Our findings prioritize trials of drugs targeting
111 IFNAR2 and ACE2 for early management of COVID-19.

112

113 INTRODUCTION

114 COVID-19 has caused a global pandemic resulting in excess mortality, stress on healthcare
115 systems and economic hardship. Even if an efficacious vaccine against SARS-CoV-2 virus
116 emerges in 2021, it could take several years to achieve herd immunity. Hence, there is a need to
117 rapidly identify drugs that can minimize the burden of COVID-19. Although large randomized
118 trials have begun to successfully identify drugs that can be repurposed to address COVID-19,^{1,2}
119 most drugs evaluated so far have failed to show efficacy and have been largely confined to
120 hospitalized or critically-ill patients. Therefore, it is pressing to identify additional drugs that can
121 be repurposed for early management in COVID-19.

122

123 Large-scale human genetic studies are now widely used to inform drug development programs as
124 drug target-disease pairs supported by human genetics have a greater odds of success in drug
125 discovery pipelines.^{3,4} For example, identification of variants in *PCSK9* associated with lower
126 risk of coronary disease led to the successful development of PCSK9 inhibitors, which are now
127 licensed for prevention of cardiovascular events.⁵ The value of human genetics for drug
128 discovery and development has also been realized for infectious diseases. Human genetic studies
129 showed that genetic variation in the *CCR5* gene provides protection against infection by human
130 immunodeficiency virus (HIV) type-1. These findings were key for the development of
131 Maraviroc, an antagonist of CCR5, approved by the FDA for the treatment of patients with HIV-
132 1.⁶

133

134 Genetic variants acting in “*cis*” on druggable protein levels or gene expression that encode
135 druggable proteins can provide powerful tools for informing therapeutic targeting, as they mimic

136 the on-target (beneficial or harmful) effects observed by pharmacological modification.⁷ Such
137 Mendelian randomization (MR) analyses have been used to suggest repurposing opportunities
138 for licensed drugs.^{8,9} MR analysis that focuses on actionable druggable genes, defined as genes
139 that encode the protein targets of drugs that are licensed or in the clinical phase of drug
140 development, could therefore serve as a fast and robust strategy to identify drug-repurposing
141 opportunities to prevent the complications and mortality due to COVID-19.

142

143 To identify further potential repurposing opportunities to inform trials of COVID-19 patients, we
144 conducted large-scale MR and colocalization analyses using gene expression and soluble protein
145 data for 1,263 actionable druggable genes that encode protein targets for approved drugs or drugs
146 in clinical development. By combining trans-ancestry genetic data from 7,554 hospitalized
147 COVID-19 patients and more than 1 million population-based controls from the COVID-19 Host
148 Genetics Initiative¹⁰ (HGI) and the Million Veteran Program¹¹ (MVP), we provide support for
149 two therapeutic strategies.

150

151 **RESULTS**

152 **Overall analysis plan**

153 **Figure 1** describes the overall scheme of the analyses. First, we identified all proteins that are
154 therapeutic targets of approved or clinical-stage drugs. Next, we selected conditionally-
155 independent genetic variants that act locally on plasma levels of these proteins or tissue-specific
156 gene expression that encode these proteins. We proposed that these variants were instrumental
157 variables and conduct two-sample MR analyses¹² using a trans-ancestry meta-analysis of 7,554
158 cases from MVP and publicly available data (HGI outcome B2 from release 4 version 1,
159 downloaded October 4th 2020, **Supplementary Table 1**). Given that all MR analyses relies on
160 several assumptions, some¹³ unverifiable, we conducted a multi-stage strategy to minimize
161 confounding and biases. For MR results that passed our significance threshold after accounting
162 for multiple testing, we performed colocalization to ensure MR results were not due to
163 confounding by linkage disequilibrium (LD). Those with evidence of colocalization were
164 investigated further using an independent proteomics platform (Olink). Finally, we conducted
165 phenome-wide scans and pathway enrichment of relevant variants to reduce risks of horizontal
166 pleiotropy and other biases due to MR violations as well as to understand potential biological
167 mechanisms.

168

169 **Actionable druggable proteins**

170 Using data available in ChEMBL version 26, we identified 1,263 human proteins as ‘actionable’
171 (i.e. therapeutic targets of approved or clinical-stage drugs) (**Supplementary Table 2**). Of these,
172 we noted 700 proteins that are targets for drugs with potential relevance to COVID-19 from cell-
173 based screening, registers of clinical trials against COVID-19 or approved
174 immunomodulatory/anticoagulant drugs (given the clear role of these pathways in COVID-19

175 outcomes), or have biological evidence for the role of the protein in SARS-CoV-2 infection
176 (**Supplementary Table 3**).

177

178 **Genetic proposed instruments for actionable druggable proteins**

179 Using GTEx version 8 (V8)¹⁴, we identified all conditionally-independent expression
180 quantitative trait loci (eQTLs) in 49 tissues that act in *cis* (within 1 Mb on either side of the
181 encoded gene), that covered 1,016 of the 1,263 druggable genes in at least one tissue
182 (**Supplementary Table 2 and 4**). We also selected *cis*-pQTLs for plasma proteins measured
183 using the SomaScan platform in 3,301 participants of the INTERVAL study¹⁵ (**Supplementary**
184 **Table 5**) and 10,708 Fenland cohort participants¹⁶ (**Supplementary Table 6**) that covered a total
185 of 67 proteins. In total 1,021 proteins had genetic proposed instruments using either eQTLs or
186 pQTLs, and 62 had proposed instruments using both.

187

188 **Mendelian randomization and colocalization**

189 Using our (eQTL and pQTL) proposed instruments, we performed two-sample MR on trans-
190 ancestry summary statistics for hospitalized COVID-19 cases from MVP and HGI
191 (**Supplementary Table 1**). Using GTEx *cis*-eQTLs as proposed instruments, we found
192 significant ($P < 4.0 \times 10^{-5}$, 0.05 Bonferroni-corrected for 1,263 proteins) MR results for six genes
193 (*IL10RB*, *CCR1*, *IFNAR2*, *PDE4A*, *ACE2* and *CCR5*) in at least one tissue (**Table 1**), and four
194 additional genes (*CA5B*, *CA9*, *NSTN* and *SLC9A3*) with suggestive MR results ($P < 5 \times 10^{-4}$ and
195 $P > 4 \times 10^{-5}$, **Figure 2**). No proposed instruments involving *cis*-pQTLs reached our suggestive
196 threshold in any of the analyses. For three significant genes (*IL10RB*, *IFNAR2*, *ACE2*) there was
197 strong evidence of colocalization (posterior probability of shared causal variant across two traits

198 - hypothesis 4 [PP.H4] >0.8) between at least one proposed instrumental variant and our trans-
199 ancestry meta-analysis of COVID-19 hospitalization (**Table 1**). Beta-coefficients of MR
200 estimates for *ACE2* were positive in all tissues (**Table 1**), meaning higher *ACE2* expression is
201 associated with higher risk of COVID-19 hospitalization. MR beta-coefficients for *IFNAR2* and
202 *IL10RB* were negative and positive, respectively, in all tissues except one for each gene (skeletal
203 muscle for *IFNAR2*; cultured fibroblasts for *IL10RB*; **Table 1**).

204

205 **IL10RB and IFNAR2**

206 Interferon alpha receptor 2 (IFNAR2) and interleukin 10 receptor beta (IL-10RB) both act as
207 receptors for interferons (IFN). IFNAR2 forms a complex with IFNAR1, which together act as a
208 receptor for type I IFN (IFN- α , β , ω , κ , ϵ), while IL-10RB acts as a receptor for type III IFN
209 (IFN- λ) when complexed with interferon lambda receptor-1 (IFNLR1)¹⁷, or IL-10 when
210 complexed with IL-10RA. IL-10RB and IFNAR2 are encoded by adjacent genes and some *cis*-
211 eQTLs for *IL10RB* are also *cis*-eQTLs for *IFNAR2* (**Supplementary Table 7, Figure 3**), making
212 it difficult to determine which gene may be responsible for the association with COVID-19 and
213 requiring further investigation.

214

215 All significant MR results for *IFNAR2/IL10RB* that colocalized with COVID-19 hospitalization
216 contained one of nine strongly correlated ($r^2 > 0.75$ in 1000G European ancestry participants)
217 variants (rs11911133, rs1051393, rs2300370, rs56079299, rs17860115, rs13050728, rs2236758,
218 rs12053666, and rs1131668), which are *cis*-eQTLs for *IL10RB* in eleven tissues and for *IFNAR2*
219 in four tissues (**Supplementary Table 8**). Within this LD block (hereafter rs13050728-LD
220 block), rs13050728 is the eQTL most strongly associated with COVID-19 hospitalization (per T-

221 allele odds ratio = 1.17; 95% CI = 1.12-1.23; $P=1.88\times 10^{-12}$; **Supplementary Table 7**). Variants
222 outside the rs13050728-LD block were not strongly associated with COVID-19 hospitalization
223 (**Figure 3**).

224

225 *pQTLs for IL10RB*

226 Using stepwise conditional analysis on Olink measurements of plasma IL-10RB we identified
227 two *cis*-pQTLs, rs2266590 ($P=1.04\times 10^{-136}$) and rs2239573 ($P= 2.66\times 10^{-19}$), which explained
228 5.4% and 1.2%, respectively, of the variance in plasma IL-10RB. rs2266590 was also an eQTL
229 for *IL10RB* in three tissues and *IFNAR2* in one tissue, while rs2239573 was also an eQTL for
230 *IL10RB* in two tissues (**Supplementary Table 8**). rs2266590 and rs2239573 lie in intron 5 and 1,
231 respectively, of the *IL10RB* gene and are located in separate regions of high epigenetic
232 modification (h3k27ac marking), indicating enhancer regions (**Figure 3**). rs2266590 and
233 rs2239573 were not associated with COVID-19 hospitalization ($P= 0.85$ for rs2266590, $P=0.66$
234 for rs2239573, **Supplementary Figure 1**) and MR using these two *cis*-pQTLs yields a null result
235 ($P=0.74$).

236

237 A third *cis*-pQTL (rs2834167, $P=1.1\times 10^{-8}$) for plasma IL-10RB measured on the SomaScan
238 platform was previously identified in 3,200 Icelanders over the age of 65.¹⁸ rs2834167 is a
239 missense variant (Lys>Glu) and is not correlated with either of the *cis*-pQTLs for plasma IL-
240 10RB measured by Olink ($r^2=0.01$ for rs2266590, $r^2=0.03$ for rs2239573 in 1000G EUR).
241 Although rs2834167 was associated with *IL10RB* expression in 18 tissues, it was not associated
242 with *IFNAR2* expression in any tissue (**Supplementary Table 8**). The A allele at rs2834167,
243 which is associated with lower *IL10RB* gene expression but higher plasma IL-10RB, was

244 inversely associated with COVID-19 (per-A-allele OR= 0.91; 95%CI= 0.87-0.95; $P=5.3\times 10^{-5}$).
245 Because Emilsson et al.¹⁸ did not report full summary statistics we could not perform
246 colocalization between this pQTL and COVID-19 hospitalization. However, rs2834167 as an
247 eQTL does not colocalize (PP.H4<0.8) with COVID-19 in any tissue (**Table 1**). These three *cis*-
248 pQTLs, while possibly functional variants altering plasma IL-10RB levels, suggest that the
249 plasma IL-10RB levels are not likely the mediator of the association between this locus and
250 COVID-19 hospitalization. This suggests the MR assumption are unlikely to hold for IL10RB.
251 IFNAR2 was not measured on the SomaScan or Olink platforms.

252

253 *Phenome-wide scan of rs13050728*

254 To identify other phenotypes associated with rs13050728, we performed a phenome-wide scan
255 of publicly available data on PhenoScanner¹⁹ and GTEx, and unpublished proteomic data in
256 INTERVAL (see methods). rs13050728 was associated with tryptase gamma 1 (TPSG1, $P=$
257 1.5×10^{-5}) and vascular endothelial growth factor 2 (VEGFR2, $P= 2.6\times 10^{-5}$, **Supplementary**
258 **Table 9**), and both showed strong evidence of colocalization with COVID-19 hospitalization
259 (PP.H4=0.96 for VEGFR2, PP.H4=0.96 for TPSG1, **Figure 4**). The C allele at rs13050728
260 associated with higher *IFNAR2* expression in all tissues (except skeletal muscle), lower risk of
261 COVID-19 hospitalization, and lower levels of plasma VEGFR2 and TPSG1 (**Supplementary**
262 **Table 9**). This mimics agonistic effects of IFNAR2 through recombinant type-I IFNs, which are
263 known to have an anti-angiogenic effect, at least in part through reduced VEGF/VEGFR2
264 signaling^{20,21}, and decrease tryptase levels in a phase-2 trial using recombinant type-I IFN in
265 patients with mastocytosis²², a condition that causes proliferation of mast cells. rs13050728 was
266 not associated at $P<4\times 10^{-5}$ (our Bonferroni corrected P value) with any phenotype beyond

267 plasma VEGFR2 and TPSG1 and gene expression of *IFNAR2* and *IL10RB* (**Supplementary**
268 **Table 9**), indicating that this variant is unlikely to exhibit widespread horizontal pleiotropy.
269 Also, the chances of substantial bias due to MR violations is low²³ since the variant is not
270 strongly associated with other risk factors that could alter the likelihood of SARS-CoV-2 testing
271 or hospitalization of COVID-19 patients.

272

273 *Pathway enrichment analysis of rs13050728*

274 Using information from all GTEx V8 tissues we identified 476 genes whose expression levels
275 were associated with rs13050728 at a nominal significance level ($P < 0.05$). Taking into
276 consideration an adjusted P value for multiple testing within the WikiPathway corpus, only two
277 biological pathways were significantly associated among all 624 pathways present in this
278 database: Host-pathogen interaction of human corona viruses - IFN induction (adjusted P value =
279 0.0028) and Type I IFN Induction and Signaling During SARS-CoV-2 Infection (adjusted P
280 value = 0.0098). In addition, among Gene Ontology (GO) and Reactome pathways, several gene
281 sets were also significantly enriched. Notably, among enriched pathways were those related to
282 IFN type I or antiviral response (**Supplementary Figure 2A**).

283

284 **ACE2**

285 Angiotensin converting enzyme 2 (ACE2) converts angiotensin II into angiotensin (1-7) as part
286 of the RAA system, and more importantly, is the viral receptor for SARS-CoV-2. We identified
287 seven *cis*-eQTLs in seven tissues (**Supplementary Table 10**) for *ACE2* which are strongly
288 correlated ($r^2 > 0.75$ in 1000G EUR, **Supplementary Table 11**) with rs4830976 being the eQTL
289 in the region most strongly associated with COVID-19 hospitalization.

290

291 *pQTLs for ACE2*

292 Stepwise conditional analysis for plasma ACE2 measured by Olink revealed one pQTL,
293 rs5935998 ($P=1.4\times 10^{-21}$), which is in high LD with a previously reported *cis*-pQTL
294 (rs12558179) for ACE2 ($r^2=0.89$ in 1000G EUR)²⁴, and a secondary suggestive signal
295 (rs4646156, $P=3.20\times 10^{-7}$). rs5935998 and rs4646156 are concordant in their effect on COVID-
296 19 hospitalization (higher ACE2 levels corresponds to higher risk of COVID-19 hospitalization
297 for both) resulting in a strong, positive MR association (MR beta-coefficient: 0.34; 95% CI:
298 0.17-0.51; $P=8.1\times 10^{-5}$). Although neither rs5935998 or rs4646156 strongly colocalized with
299 COVID-19 hospitalization (PP.H4=0.49 for rs5935998, PP.H4=0.08 for rs4646156,
300 **Supplementary Figure 3**), the two pQTLs, while statistically independent, are mildly correlated
301 ($r^2=0.2$ in 1000G EUR), which can make colocalization difficult to interpret.²⁵ One possible
302 explanation is that these two pQTLs confer an effect on COVID-19 hospitalization that
303 converges on the rs4830976-LD-block, as both are moderately correlated with rs4830976
304 ($r^2=0.32$ for rs5935998, $r^2=0.42$ for rs4646156 in 1000G EUR, **Supplementary Figure 3**)

305

306 *Phenome-wide scan of rs4830976*

307 rs4830976 is associated ($P<4\times 10^{-5}$) with and colocalized (PP.H4>0.8) with expression of nearby
308 genes *CA5B*, *CLTRN* (also known as *TMEM27*), and *VEGFD* (**Supplementary Table 12**) in at
309 least one tissue, indicating that this variant may be instrumenting on gene expression beyond
310 *ACE2*. However, given the biological prior that *ACE2* acts as the receptor of SARS-CoV-2,
311 *ACE2* is probably more likely than *CA5B*, *CLTRN* or *VEGFD* to be responsible for COVID-19
312 hospitalization. There were no other reported phenome-wide scan results at $P < 4\times 10^{-5}$ for

313 rs4830976, which is at least in part due to the lack of reported X-chromosome results from a
314 large proportion of GWAS.

315

316 *Pathway enrichment analysis of rs4830976*

317 Exploring the landscape of genes differentially expressed according to genotype in GTEx V8, we
318 observed 1397 genes differentially expressed at a nominal P value less than 0.05. Over-
319 representation analysis identified 238 significantly enriched biological pathways among
320 differentially expressed genes (**Supplementary figure 2B**). Among these, signaling by
321 interleukins, regulation of cytokine production, and antigen processing and presentation, might
322 prove biologically relevant in COVID-19 infection.

323

324 **DISCUSSION**

325 To identify drug-repurposing opportunities to inform trials against COVID-19, we conducted a
326 large-scale MR analysis of protein and gene expression data. We first updated the “actionable”
327 genome to an enlarged set of 1,263 human proteins and provided evidence for 700 of these as
328 targets for drugs with some potential relevance to COVID-19. By testing more than a thousand
329 of these using several of the largest currently available human genetic datasets, we provide
330 evidence for drug targets of type-I IFNs (IFNAR2) and ACE2 modulators (ACE2) as priority
331 candidates for evaluation in randomized trials of early management in COVID-19.

332

333 Our finding that ACE2 may play an important role in COVID-19 is unsurprising given its well-
334 known relevance to SARS-CoV-2. Since ACE2 acts as the primary receptor for SARS-CoV-2,
335 increased expression of ACE2 has been hypothesized to lead to increased susceptibility to
336 infection. ACE2 plays a vital role in the RAAS signaling pathway, providing negative regulation
337 through the conversion of Angiotensin II to Angiotensin 1-7. This action has anti-inflammatory
338 and cardioprotective effects²⁶ and plays a protective role in acute respiratory distress
339 syndrome.^{27,28} ACE2 is a single-pass membrane protein but can be cleaved from the membrane
340 to a soluble form which retains the enzymatic function to cleave Angiotensin II. It has therefore
341 been hypothesized that administration of human recombinant soluble ACE2 (hrsACE2) could be
342 an effective treatment for COVID-19, through distinct mechanisms in two phases of COVID-19.
343 First, hrsACE2 can bind the viral spike glycoprotein of SARS-CoV-2, which could prevent
344 cellular uptake of SARS-CoV-2 by reducing binding to the membrane-bound form of ACE2
345 (early phase). This suggestion is supported by the finding that APN01, a hrsACE2 therapeutic,
346 showed a 1000-5000 fold reduction in SARS-CoV-2 viral load in primate kidney epithelial
347 (Vero) cells, as well as inhibiting infection of human blood vessel and kidney organoids.²⁹ In the

348 later phase, hrsACE2 could reduce sequelae of SARS-CoV-2 infection by reducing inflammation
349 in the lungs and other infected tissues. A case report of a hospitalized COVID-19 patient
350 supported this hypothesis by showing that 7-day administration of APN01 was associated with a
351 reduction in SARS-CoV-2 viral load and inflammatory markers.³⁰ APN01 is currently being
352 tested in a phase II trial to reduce mortality and invasive mechanical ventilation in 200
353 hospitalized COVID-19 patients³¹.

354

355 One of the main challenges of our analysis was to determine whether IFNAR2 or IL10RB (or
356 both) was driving the association with COVID-19 hospitalization, given that they share *cis*-
357 eQTLs used as proposed instruments for our MR analysis. Multiple lines of evidence indicate
358 that IFNAR2 appears to be primarily responsible for the signal observed. First, our phenome-
359 wide scan using the lead *IFNAR2/IL10RB cis*-eQTL reproduced known effects of type-I IFNs
360 (the therapeutic target of IFNAR2) on VEGFR2 and TPSG1.²⁰⁻²² Second, our pathway
361 enrichment analysis using the same eQTL revealed pathways associated with type-I IFN receptor
362 (IFNAR2) signaling. Last, three independent *cis*-pQTLs that are also *cis*-eQTLs for *IL10RB* did
363 not show evidence of association with COVID-19, suggesting that plasma IL-10RB
364 concentrations are unlikely to be etiologically relevant to COVID-19.

365

366 Evidence of a role for type-I IFN in COVID-19 is rapidly emerging. Studies using *in vitro* (A549
367 pulmonary cell lines), animal (ferrets) and *ex vivo* (human lung tissue) models have all shown
368 lower expression of genes encoding type-I IFNs after exposure to SARS-CoV-2 compared to
369 other respiratory viruses.^{32,33} This has been confirmed *in vivo* by studies showing significantly
370 impaired type-I IFN response – including almost no IFN-beta activity - in the peripheral blood of
371 severe COVID-19 patients compared to mild to moderate COVID-19 patients.³⁴ More

372 importantly, lower levels of IFN alpha-2 among recently hospitalized COVID-19 patients were
373 associated with a substantial increase in the risk of progression to critical care, supporting our
374 observation that lower genetically-predicted *IFNAR2* expression was associated with higher risk
375 of COVID-19 hospitalization.³⁴ Additionally, auto-antibodies for type 1 IFNs were found in a
376 much higher proportion of individuals with severe COVID-19 than those with asymptomatic or
377 mild SARS-CoV-2 infection.³⁵

378

379 Whole exome and genome sequencing studies on severe COVID-19 patients have identified rare
380 mutations that implicate type I IFN signaling. Zheng et al.³⁶ found severe COVID-19 patients
381 were enriched for rare variants predicted to cause loss of protein function at 13 genes involved in
382 type-I IFN response. A cases-series of four severe COVID-19 patients under the age of 35 found
383 a rare LOF mutation in *TLR7* and decreased type 1 IFN signaling.³⁷

384

385 Several *in vitro* studies have shown that when diverse type of cells (including animal and human)
386 and human organoids were pre-treated with type-I or -III IFNs, a reduction in SARS-CoV-2
387 replication was observed when compared with controls³⁸⁻⁴² (**Supplementary Table 13**). Though
388 these *in vitro* studies are encouraging, the evidence from randomized trials for type I IFNs in
389 early COVID-19 stages is limited. Hung et al.⁴³ showed that randomization to a combination of
390 IFN beta-1b, ribavirin and lopinavir-ritonavir was superior to lopinavir-ritonavir alone in
391 shortening the duration of viral shedding, alleviating symptoms and reducing the length of the
392 hospital stay. Importantly, these benefits were confined to a subgroup who were hospitalized
393 within 7 days of onset of symptoms where IFN beta-1b was administered to the intervention arm.
394 These results, together with our genetic findings on COVID-19 hospitalization and the
395 established role of type-I IFNs as first line of response against viral agents suggest recombinant

396 type-I IFN as potential intervention during early stages of COVID-19. To date, there is no large
397 randomized trial on IFN beta for early treatment of COVID-19 patients who are at high risk of
398 hospitalization.

399

400 Trial evidence on the use of IFN-beta in late stages of COVID-19 has emerged in the last month.
401 The SOLIDARITY trial, which randomized 2,050 hospitalized COVID-19 patients to IFN beta-
402 1a, found no effect on mortality overall (relative risk (RR)=1.16, [95%CI: 0.96-1.39]), but
403 possibly a trend across subgroups of COVID-19 severity at randomization (RR=1.40 [95%CI:
404 0.82-2.40] for those on ventilator, RR=1.13 [95%CI: 0.86-1.50] for those not ventilated but on
405 oxygen, and RR=0.80 [95%CI: 0.27-2.35] in those with neither).⁴⁴ The Adaptive COVID-19
406 Treatment Trial 3 (ACTT-3) trial stopped enrollment of severely ill COVID-19 patients for a
407 trial on IFN beta-1a and remdesivir due to adverse events but continued enrolling patients with
408 less severe disease.⁴⁵ These findings indicate no role for the use of IFN beta during late stages of
409 COVID-19, where the cytokine storm has already established.

410

411 Our study has several strengths. We provide an updated catalog of all actionable protein targets
412 and drugs that are amenable to causal inference investigation through human genetics. By
413 combining several of the most comprehensive datasets currently available and linking genetic
414 variation with gene expression, plasma protein levels and COVID-19, we were able to robustly
415 evaluate over one thousand actionable drug targets. We took multiple steps to minimize potential
416 biases and confounding that could invalidate our actionable druggable-genome-wide MR
417 analysis, such as using colocalization methods to minimize the chances of false positive results
418 due to confounding by LD. We were careful to reduce the impact of horizontal pleiotropy by
419 restricting our proposed instruments to variants acting in *cis* and performing phenome-wide scan

420 to ensure they were only associated with gene expression of the tested gene or downstream
421 phenotypes (vertical pleiotropy), but unmeasured sources can remain.

422

423 Our analysis also has limitations. Though we make use of instrumental variants from multiple
424 data sources, none cover the entire actionable druggable genome, were ancestry-specific or were
425 derived from COVID-19 patients. However, we managed to recover credible biological targets
426 from our analysis that were consistent across ancestral groups. Identifying the most relevant
427 tissue or cell-type can be challenging for interpreting MR analyses of gene expression. In our
428 case, a relevant tissue could be: one invaded by SARS-CoV-2, an organ associated with clinical
429 complications of COVID-19, a tissue where the COVID-19-relevant protein is produced, or a
430 tissue that would be the likely site of action for the target drug. We opted to use a data-driven
431 strategy that incorporates all tissues available in GTEx V8. For *IFNAR2*, we recovered
432 fibroblasts (the main cell type responsible for IFN-beta production), esophageal mucosa^{46,47} (a
433 tissue invaded by SARS-CoV-2), and skeletal muscle⁴⁸ (associated with the neurological
434 manifestations of COVID-19). For *ACE2*, we recovered brain tissue, an organ known to be
435 invaded by SARS-CoV-2 and associated with clinical manifestations.^{49,50}

436

437 In conclusion, our trans-ancestry MR analysis covering all actionable druggable genes identified
438 two drug repurposing opportunities (type-I IFNs and hsrACE2) as interventions that need to be
439 evaluated in adequately powered randomized trials to investigate their efficacy and safety for
440 early management of COVID-19.

441

442 **METHODS**

443 **Identification of actionable druggable genes suitable for repurposing against COVID-19**

444 Information about drugs and clinical candidates, and their therapeutic targets, was obtained from
445 the ChEMBL database (release 26⁵¹, **Supplementary Methods**). For the purposes of our
446 COVID-19 drug repurposing efforts, actionable proteins were defined as those that are
447 therapeutic targets of approved drugs and clinical candidates or are potential targets of approved
448 drugs. Therapeutic targets were identified from the drug mechanism of action information in
449 ChEMBL and linked to their component proteins. Each protein was assigned a confidence level
450 based on the type and size of target annotated, and the resulting list was filtered to remove non-
451 human proteins and those with lower confidence assignments (cases where the therapeutic target
452 consists of more than 10 proteins or the protein is known to be a non-drug-binding subunit of a
453 protein complex). For approved drugs, additional potential human target proteins were identified
454 from pharmacological assay data in ChEMBL with recorded affinity/efficacy measurements \leq
455 100nM (represented by a pChEMBL value \geq 7).

456
457 A total of 1,263 unique human proteins were identified as ‘actionable’ from data available in
458 ChEMBL. These consisted of 531 proteins that are therapeutic targets of approved drugs, 381
459 additional proteins that are therapeutic targets of clinical candidates and 351 additional proteins
460 that are bound by approved drugs, but not annotated as the therapeutic targets. While the
461 biological relevance of the latter group of targets in the context of the approved drug indications
462 may be unclear, the high affinity/efficacy measurements suggest the drug should be capable of
463 modulating these proteins, should they be found to be relevant to COVID-19 (although likely not
464 in a selective manner). Proteins were further annotated with biological and drug information
465 relating to their potential role in SARS-CoV-2 infection (**Supplementary methods**) such as

466 change in abundance during infection, interaction with viral proteins or the activity of drugs in
467 antiviral cell-based assays. Of the 1,263 actionable proteins identified previously, 300 were
468 annotated as biologically relevant in SARS-CoV-2 infection and 547 were targets of drugs with
469 some evidence of COVID-19 relevance from cell-based assays, clinical trials or the ATC
470 classification (**Supplementary Table 2**).

471

472 **Selection of proposed instruments**

473 *eQTL proposed instruments*

474 We proposed eQTL instruments using raw data from GTEx Version 8 by performing conditional
475 analysis on normalized gene expression in European ancestry individuals in 49 tissues that had at
476 least 70 samples. We used Matrix eQTL⁵² and followed the same procedure as outlined by the
477 GTEx consortium (<https://gtexportal.org/home/>). Briefly, after filtering the genotypes (genotype
478 missingness <0.05, MAF<0.01, HWE<0.000001, removing ambiguous SNPs), within each
479 tissue, we performed GWAS between variants and gene expression adjusting for sex, the first 5
480 principal components of European genetic ancestry, PEER factors, sequencing platform and
481 protocol. To identify independent eQTLs, we performed conditional analysis in regions around
482 associations that fell below genome-wide significance, additionally adjusting for the peak variant
483 if there exists an association reaching a P -value of 5×10^{-8} . *Cis*-eQTLs were defined as significant
484 ($P < 5 \times 10^{-8}$) associations within 1Mb on either side of the encoded gene. To convert from build 38
485 to build 37, we used the table available from the GTEx consortium for all variants genotyped in
486 GTEx v8 and hg19 liftover,
487 ([https://storage.googleapis.com/gtex_analysis_v8/reference/GTEX_Analysis_2017-06-](https://storage.googleapis.com/gtex_analysis_v8/reference/GTEX_Analysis_2017-06-05_v8_WholeGenomeSeq_838Indiv_Analysis_Freeze.lookup_table.txt.gz)
488 [05_v8_WholeGenomeSeq_838Indiv_Analysis_Freeze.lookup_table.txt.gz](https://storage.googleapis.com/gtex_analysis_v8/reference/GTEX_Analysis_2017-06-05_v8_WholeGenomeSeq_838Indiv_Analysis_Freeze.lookup_table.txt.gz)). In each tissue,

489 multiple GW-significant ($P < 5 \times 10^{-8}$) eQTLs for the same gene were combined into a single
490 instrument. For example, for *IL10RB* expression in skeletal muscle tissue, there were two
491 conditionally-independent eQTLs (rs2300370 and rs2834167, **Table 1**) that made up that
492 instrument.

493

494 *pQTL proposed instruments*

495 We proposed pQTL instruments from two sources of publicly available data that reported
496 conditionally independent pQTLs for proteins measured by the SomaLogic Inc. (Boulder,
497 Colorado, US) SomaScan^{53,54} platform: (1) Sun *et al.*¹⁵, which reported results for 2,994 proteins
498 in 3,301 INTERVAL participants and (2) Pietzner *et al.*¹⁶, which reported results for 179
499 proteins in 10,708 participants of the Fenland. In both, we restricted proposed instrumental
500 variants to *cis*-pQTLs for actionable proteins, used a *P* value threshold of 5×10^{-8} and removed
501 variants with $MAF < 0.01$. MR was run independently for each data source (i.e. proposed
502 instruments for the same protein in different platforms were tested against COVID-19
503 hospitalization independently).

504

505 **Estimates for COVID-19 hospitalization**

506 To generate outcome summary-statistics, we meta-analyzed results from the Million Veteran
507 Program (MVP), an ongoing, prospective cohort recruiting from 63 Veterans Health
508 Administration (VA) medical facilities (**Supplementary Methods**), and the Host Genetics
509 Initiative,¹⁰ a global collaboration to accumulate GWAS on COVID-19 infection and clinical
510 manifestations.

511

512 In MVP, 1,062 COVID-19 cases (**Supplementary Table 1**) were identified between March 1st
513 and September 17, 2020 using an algorithm developed by the VA COVID National Surveillance
514 Tool (NST). The NST classified COVID-19 cases as positive or negative based on reverse
515 transcription polymerase chain reaction (rRT-PCR) laboratory test results conducted at VA
516 clinics, supplemented with Natural Language Processing (NLP) on clinical documents. The
517 algorithm to identify COVID-19 patients is continually updated to ensure new annotations of
518 COVID-19 are captured from the clinical notes, with chart reviews performed periodically to
519 validate the algorithm.⁵⁵ COVID-19-related hospitalizations were defined as admissions from 7
520 days before up to 30 days after a patient's first positive test for SARS-CoV-2 test. We tested
521 association between all our proposed genetic instruments and COVID-19 hospitalization (versus
522 population controls) in MVP adjusting for age, sex and the first 10 principal components in
523 ancestry-specific strata using PLINK v2 (analysis completed on October 10, 2020). The MVP
524 received ethical and study protocol approval by the Veterans Affairs Central Institutional Review
525 Board and informed consent was obtained for all participants.

526

527 We downloaded publicly available summary statistics for the B2 outcome from Host Genetic
528 Initiative on October 4, 2020 (release 4 version 1). In total, HGI accumulated 6,492 cases of
529 COVID-19 hospitalization through collaboration from 16 contributing studies (**Supplementary**
530 **Table 1**), which were asked to define cases as “hospitalized laboratory confirmed SARS-CoV-2
531 infection (RNA and/or serology based), hospitalization due to corona-related symptoms” versus
532 population controls

533 (https://docs.google.com/document/d/1okamrqYmJfa35CILvCt_vEe4PkvTwwgHq7T3jbeyCI/vi

534 [ew](#)) and use a model that adjusts for age, age², sex, age*sex, PCs, and study specific covariates

535 (https://docs.google.com/document/d/16ethjgi4MzlQeO0KAW_yDYyUHdB9kKbtfuGW4XYV

536 [KQg/view](#)) Results for each ancestry-stratum were meta-analyzed along with the HGI (summary
537 statistics already meta-analyzed from contributing studies) for COVID hospitalization using
538 METAL software⁵⁶ with inverse-variance weighting and fixed effects.

539

540 **Mendelian randomization and colocalization**

541 We conducted MR analyses using the R package TwoSampleMR
542 (<https://mrcieu.github.io/TwoSampleMR/>). We used fixed-effects, inverse-variance weighted
543 MR for proposed instruments that contain more than one variant, and Wald-ratio for proposed
544 instruments with one variant. For proposed instruments with multiple variants, we also tested the
545 heterogeneity across variant-level MR estimates, using the Cochrane Q method
546 (`mr_heterogeneity` option in TwoSampleMR package). We defined significant MR results using
547 a P value threshold of $P < 4.0 \times 10^{-5}$ (0.05 Bonferroni-corrected for 1,263 actionable druggable
548 genes) and identified a list of “suggestive” actionable druggable targets that passed a threshold of
549 $P < 5 \times 10^{-4}$. For statistically significant MR results, we also performed colocalization⁵⁷ between
550 each eQTL and the trans-ancestry meta-analysis on COVID-19 hospitalization using the `moloc` R
551 package (<https://github.com/clagiamba/moloc>) with default priors (probability of shared causal
552 variant for trait 1 and trait 2 is $p1=p2=1 \times 10^{-4}$, probability of shared causal variant across two
553 traits is $p12=1 \times 10^{-5}$). For example, if a proposed instrument contained two variants, we
554 performed colocalization for the primary eQTL GWAS with COVID-19 hospitalization, as well
555 as the secondary eQTL GWAS (i.e. eQTL GWAS after adjusting for peak variant from primary
556 GWAS) with COVID-19 hospitalization. Statistically significant MR hits with posterior
557 probability for hypothesis-4 (PP.H4) > 0.8 (i.e. the probability of a shared causal variant) for a
558 least one instrumental variant were then investigated further using the following analyses.

559

560 **Identifying pQTLs using Olink assay**

561 We performed stepwise conditional analysis to identify *cis*-pQTL proposed instruments for
562 proteins that passed our significance and colocalization thresholds and were one of 354 unique
563 proteins measured on four Olink⁵⁸ panels (CVD1, CVD2, Inflammation, and Neuro⁵⁹) in 4,998
564 INTERVAL participants.¹⁵ INTERVAL is a prospective cohort study of ~50,000 blood donors
565 recruited from 25 National Health Service Blood and Transplant centers in England. Participants
566 were genotyped using the UK Biobank Affymetrix Axiom array, followed by phasing using
567 SHAPEIT3 and imputation on the Sanger Imputation Server using a 1000 Genomes Phase 3-
568 UK10K imputation panel. Alleles were tested against Olink proteins using SNPTEST v2.5.2 and
569 adjusted for age, sex, plate, time from blood draw to processing, season and the first 5 principal
570 components. Conditional analysis was performed by adjusting for peak variants until no
571 association fell below 5×10^{-6} .

572

573 **Phenome-wide scan**

574 We conducted a phenome-wide scan for variants with the following goals. First, we want to
575 evaluate that our proposed instruments could reproduce the known phenotype associations (e.g.
576 disease, biomarkers) ascribed to the drug that are due to on-target effects. Secondly, we want to
577 identify if our proposed instruments are associated with comorbidities associated with greater
578 likelihood of SARS-CoV-2 testing or predictors of hospitalization in COVID-19 patients, as this
579 could potentially highlight the presence of certain biases.²³ Also, for genes that were the target of
580 licensed drugs, we checked whether the disease indication was also a risk factor for COVID-19
581 outcomes, as this might introduce a bias analogous to confounding by indication in MR.

582

583 To accomplish these goals, we investigated proposed instruments for associations of a phenome-
584 wide range of outcomes. We searched the GTEx¹⁴ Portal (<https://gtexportal.org/home/>) for gene
585 expression, and Phenoscanner¹⁹ (<http://www.phenoscaner.medschl.cam.ac.uk/>) for proteins,
586 traits and diseases. We additionally queried variants in 354 Olink proteins data that have not
587 been made publicly available (described earlier), and proteins measured by the SomaScan
588 platform not previously published in Sun *et al.*¹⁵

589

590 **Characterizing downstream transcriptional consequences of associated loci**

591 In order to confirm the specificity of the identified loci and to better explore their most important
592 downstream transcriptional consequences, we have studied the transcriptional landscape
593 modulation associated with the selected variants using GTEx V8 data with representation of 49
594 different tissues. For this we have used rs13050728 as the proxy of the *IFNAR2/IL10RB* locus
595 and rs4830976 as the proxy of the *ACE2* locus and conducted a differential gene-expression
596 analysis for all transcripts available in GTEx V8. After fitting models for all genes, enrichment
597 pathway analysis was conducted to retrieve the most enriched pathways using both the
598 differentially expressed (DE) gene list (through an over-representation analysis) and a Gene Set
599 Enrichment Analysis framework (using the R package *clusterProfiler*⁶⁰). For enrichment
600 analysis we have used the corpus from WikiPathways, Gene Ontology and Reactome.

601

602 **Author contribution**

603 J.P.C., A.S.B. and J.M.G. conceived the study design. A.G., A.P.B. and A.R.L. defined the
604 actionable genome, and identified and curated drug information relating to SARS-CoV-2; P.B.
605 and I.B.-H. provided biological annotation relating to SARS-CoV-2; C.G. performed stepwise
606 conditional analysis on GTEx raw data; D.P. tested associations for COVID-19 in MVP; L.G.
607 and J.H.Z. performed meta-analysis of HGI and MVP; L.G. performed Mendelian randomization
608 analysis; L.G. and C.G. performed colocalization analyses; L.G. and B.P.P. performed
609 conditional analysis on Olink proteins; L.G. and E.A. performed phenome-wide scans; A.C.P.
610 performed pathway enrichment analysis; J.N.D., A.S.B., and J.E.P. provided INTERVAL data;
611 Several authors were involved in the curation of the MVP data; L.G., C.G., A.C.P., A.G., D.P.,
612 A.S.B. and J.P.C. wrote the manuscript. J.P.C. oversaw all analyses. All authors critically
613 reviewed the manuscript.

614 **Data availability**

615 Full MR results can be found on dbGAP in the VA Million Veteran Program space (accession:
616 phs001672).

617 **Competing interests**

618 The authors declare no competing interests.

619 **Acknowledgements**

620 We are grateful to the Host Genetic Initiative for making their data publicly available (full
621 acknowledgements can be found here: <https://www.covid19hg.org/acknowledgements/>). This
622 research is based on data from the Million Veteran Program, Office of Research and
623 Development, Veterans Health Administration, and was supported by award #MVP035. This
624 research was also supported by additional Department of Veterans Affairs awards grant
625 #MVP001. This publication does not represent the views of the Department of Veteran Affairs or

626 the United States Government. Full acknowledgements for the VA Million Veteran Program
627 COVID-19 Science Initiative can be found in the supplementary methods. C.G. has received
628 funding from the European Union’s Horizon 2020 research and innovation program under the
629 Marie Skłodowska-Curie grant agreement No 754490 – MINDED project. A.G., P.B. and A.R.L.
630 are funded by the Member States of the European Molecular Biology Laboratory (EMBL). I.B.-
631 H. received funding from Open Targets (grant agreement OTAR-044). The Fenland Study
632 (10.22025/2017.10.101.00001) is funded by the Medical Research Council (MC_UU_12015/1);
633 we are grateful to all the volunteers and to the General Practitioners and practice staff for
634 assistance with recruitment; we thank the Fenland Study Investigators, Fenland Study Co-
635 ordination team and the Epidemiology Field, Data and Laboratory teams; we further
636 acknowledge support for genomics from the Medical Research Council (MC_PC_13046);
637 proteomic measurements were supported and governed by a collaboration agreement between
638 the University of Cambridge and Somalogic. J.E.P. is supported by UKRI Innovation Fellowship
639 at Health Data Research UK (MR/S004068/2). L.R., N.H. and C.L. are supported by the Swedish
640 Research Council.

641 **Figure Legends**

642 **Figure 1. Outline of the analyses performed.** Using multiple data sources, this study tested *cis*-
643 pQTL and *cis*-eQTL proposed instruments for actionable druggable proteins against COVID-19
644 hospitalization summary statistics meta-analyzed from the Host Genetics Initiative and the
645 Million Veteran program. Significant MR associations that also showed evidence for
646 colocalization were investigated further with an independent platform (Olink), phenome-wide
647 scans of relevant variants, and pathway enrichment.

648
649 **Figure 2. Manhattan plot of results from actionable druggable genome-wide Mendelian**
650 **randomization analysis.** Blue solid line indicates the *P* value threshold for significance (*P*
651 $<4.0 \times 10^{-5}$, 0.05 Bonferroni-corrected for 1,263 actionable druggable genes) and red dashed line
652 indicates the suggestive ($P < 5 \times 10^{-4}$) threshold. Genes are labeled by their most significant MR
653 association. For example, the results for *IL10RB* is most significant with *cis*-eQTL proposed
654 instruments derived in skeletal muscle tissue, which is the point labeled. Results are plotted by
655 the gene start position. All MR results with *P* value less than 5×10^{-4} used the GTEx *cis*-eQTLs as
656 proposed instruments.

657
658 **Figure 3. Genomic context, local association plot and LD structure of the *IFNAR2/IL10RB***
659 **region. A,** Local association plot of the interval defined by all unique eQTLs for *IL10RB* or
660 *IFNAR2*. Color code represents the degree of linkage disequilibrium with the most associated
661 marker in 1000G Europeans. **B,** Genomic context of the region. Coding genes are represented by
662 the refseq transcript. Bars represent epigenome Roadmap layered H3K27 acetylation markers.
663 Connecting lines represent significant Hi-C interactions. **C,** Set of rsIDs used as proposed

664 instruments for Mendelian Randomization analysis. Color code represents instruments for
665 *IL10RB* (blue circles), *IFNAR2* (green circles). Red half-circles represent pQTLs for IL-10RB. **D**
666 Linkage disequilibrium structure and blocks defined using European populations from 1000G
667 project.

668

669 **Figure 4. Regional association plots of the *IFNAR2-IL10RB* locus.** Regional association plots
670 for **A**, *IL10RB* gene expression in tibial nerve tissue from GTEx, **B** COVID-19 hospitalization
671 from HGI and MVP trans-ancestry meta-analysis, **C**, **D**, Vascular endothelial growth factor
672 receptor 2 (VEGFR2) and tryptase Gamma (TPSG1), respectively, measured by SomaLogic in
673 3,301 INTERVAL participants. All show the correlation (1000G European ancestry) for
674 rs13050728, the *cis*-eQTL most associated with COVID-19 hospitalization in the *IL10RB*-
675 *IFNAR2* region. All colocalize with each other (PP.H4>0.96 for all).

676

Figure 1.

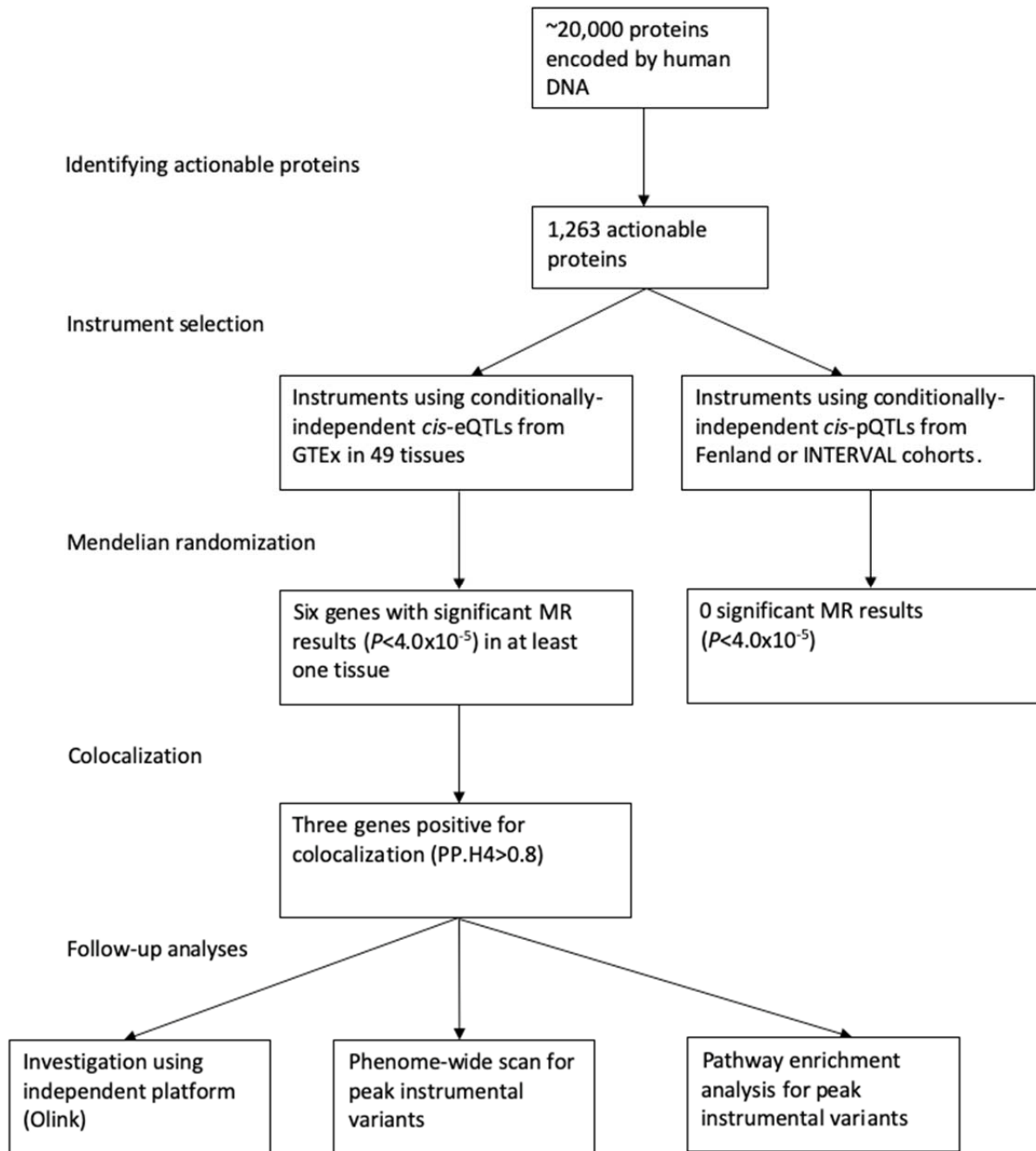


Figure 2.

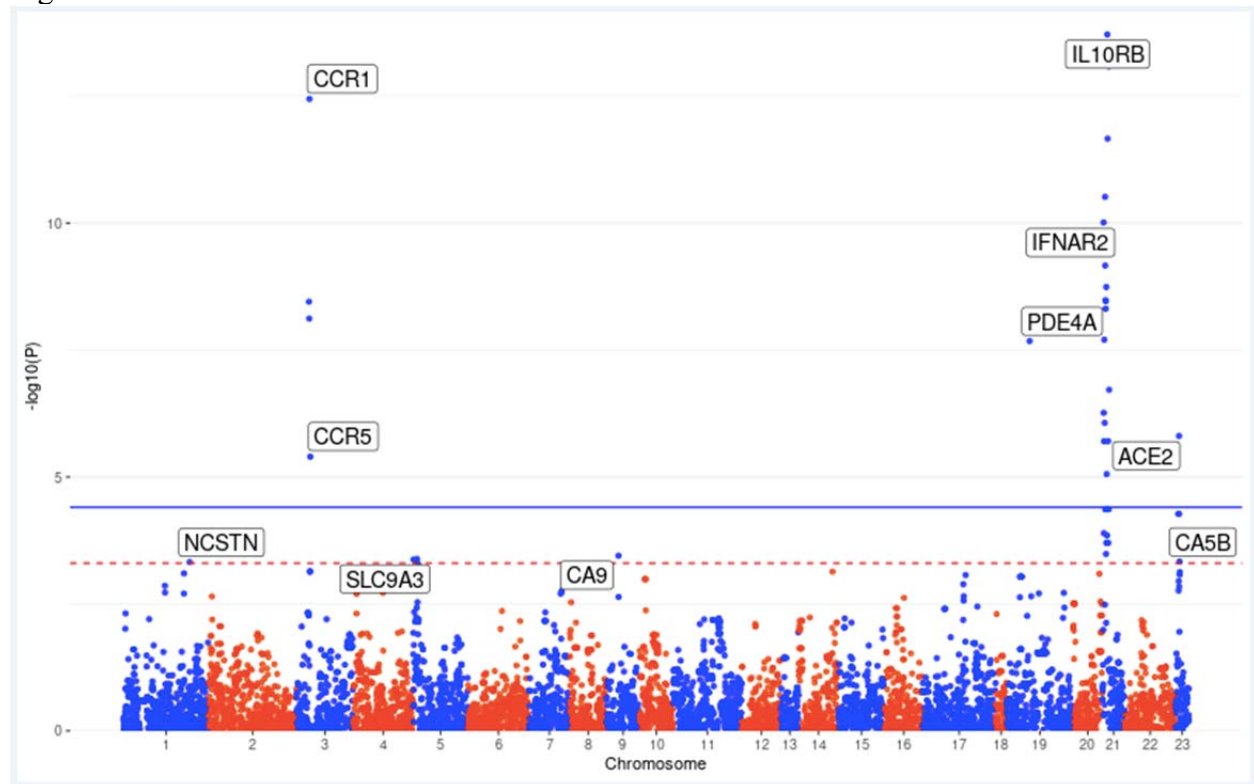


Table 1. Significant ($P < 4.0 \times 10^{-5}$) MR results.

| Gene | Tissue | beta | S.E. | P value | P_{het} | Variants in instrument | Colocalization |
|--------|---------------------------------------|---------|--------|----------|-----------|----------------------------------|--------------------|
| IL10RB | Muscle Skeletal | 0.5081 | 0.0664 | 1.93E-14 | 0.9732 | rs2300370, rs2834167 | 0.98, <0.01 |
| IL10RB | Nerve Tibial | 0.2861 | 0.0383 | 8.15E-14 | 0.0052 | rs13050728, rs2834167, rs2266590 | 0.98, <0.01, <0.01 |
| CCR1 | Cells Cultured fibroblasts | 0.4449 | 0.0612 | 3.60E-13 | NA | rs13095940 | <0.01 |
| IL10RB | Brain Nucleus accumbens basal ganglia | 0.2535 | 0.0361 | 2.18E-12 | 0.0019 | rs2834167, rs17860115 | 0.75, 0.98 |
| IL10RB | Brain Caudate basal ganglia | 0.2628 | 0.0395 | 3.05E-11 | 0.0003 | rs2834167, rs1051393 | 0.01, 0.97 |
| IFNAR2 | Muscle Skeletal | 0.5881 | 0.0909 | 9.75E-11 | NA | rs2300370 | 0.98 |
| IL10RB | Brain Cerebellar Hemisphere | 0.1403 | 0.0227 | 6.87E-10 | 0.0389 | rs2834167, rs2236758 | 0.01, 0.95 |
| IL10RB | Breast Mammary Tissue | 0.6490 | 0.1079 | 1.82E-09 | NA | rs12053666 | 0.95 |
| IL10RB | Brain Frontal Cortex BA9 | 0.1926 | 0.0326 | 3.33E-09 | 0.0366 | rs2834167, rs1131668 | 0.14, 0.97 |
| IL10RB | Brain Cortex | 0.1773 | 0.0300 | 3.43E-09 | 0.0354 | rs2834167, rs1131668 | 0.02, 0.96 |
| CCR1 | Esophagus Gastroesophageal Junction | 0.4667 | 0.0790 | 3.55E-09 | NA | rs13059906 | 0.05 |
| IL10RB | Brain Cerebellum | 0.1146 | 0.0196 | 4.86E-09 | 0.0239 | rs2834167, rs1131668 | <0.01, 0.96 |
| CCR1 | Esophagus Mucosa | 0.4338 | 0.0751 | 7.60E-09 | NA | rs34059564 | <0.01 |
| IFNAR2 | Esophagus Mucosa | -0.5420 | 0.0965 | 1.98E-08 | NA | rs11911133 | 0.92 |
| PDE4A | Artery Aorta | -0.4826 | 0.0861 | 2.11E-08 | 0.0202 | rs370630099, rs45524632 | 0.41, 0.61 |
| IL10RB | Testis | 0.7104 | 0.1364 | 1.92E-07 | NA | rs2284550 | 0.11 |
| IFNAR2 | Skin not Sun Exposed Suprapubic | -0.3360 | 0.0671 | 5.46E-07 | NA | rs8127500 | <0.01 |
| IFNAR2 | Pancreas | -0.4708 | 0.0957 | 8.63E-07 | NA | rs1476415 | 0.06 |
| ACE2 | Brain Frontal Cortex BA9 | 0.1121 | 0.0233 | 1.56E-06 | NA | rs4830976 | 0.95 |
| IFNAR2 | Cells Cultured fibroblasts | -0.3893 | 0.0819 | 1.98E-06 | NA | rs1131668 | 0.92 |
| IL10RB | Cells Cultured fibroblasts | -0.5197 | 0.1093 | 1.98E-06 | NA | rs1131668 | 0.96 |
| CCR5 | Lung | -0.5868 | 0.1272 | 3.99E-06 | NA | rs12639314 | 0.02 |
| IL10RB | Esophagus Gastroesophageal Junction | 0.4678 | 0.1052 | 8.80E-06 | NA | rs56079299 | 0.96 |

678

679

680

681

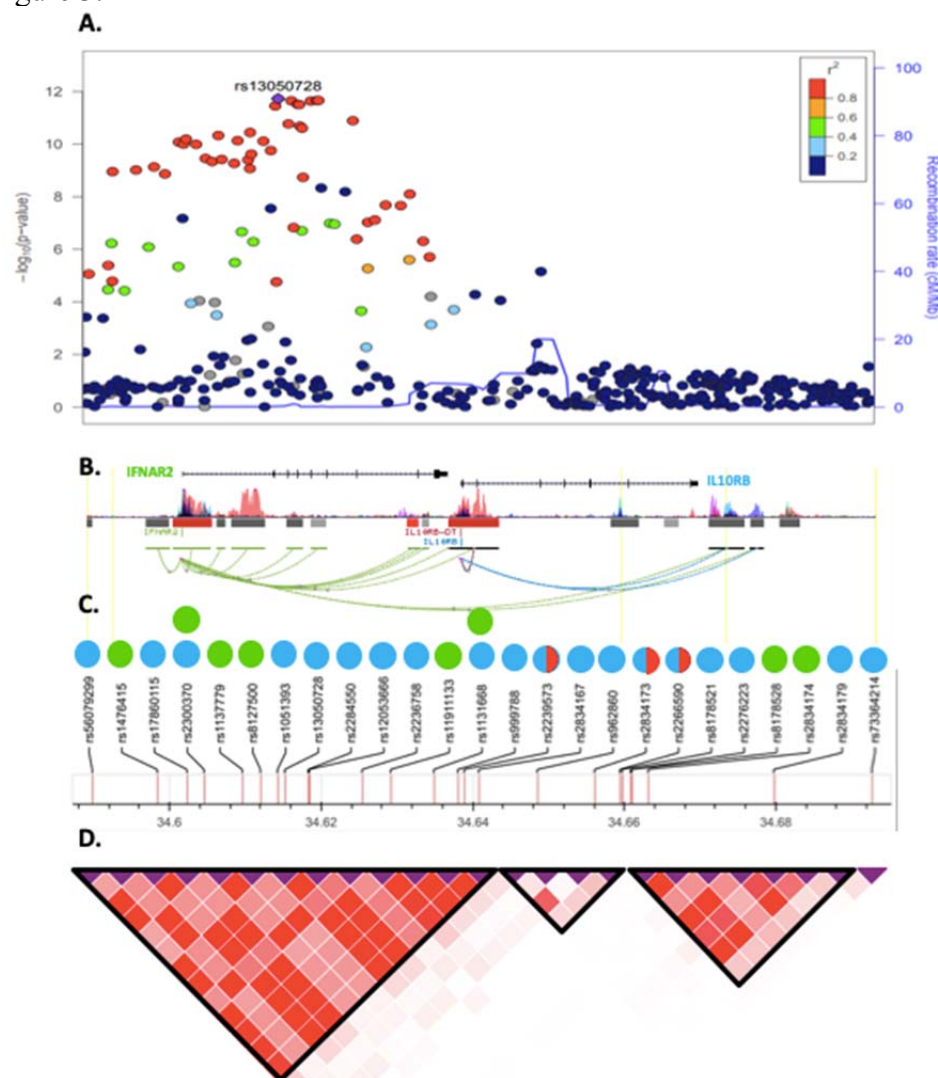
682

683

684

Significant Mendelian randomization results $P < 4.0 \times 10^{-5}$. All results used *cis*-eQTL instruments. No results using *cis*-pQTL instruments yielded results $P < 5 \times 10^{-4}$. P_{het} refers to the heterogeneity P value across individual-variant MR estimates within a genetic instrument calculated using the Cochrane Q method, therefore instruments containing one variant were not tested for heterogeneity. A positive beta estimate indicates that more gene expression is associated with higher risk of COVID-19 hospitalization. “Colocalization” indicates PP.H4 between eQTLs and COVID-19 hospitalization. For example, for IL10RB in skeletal muscle, the primary GWAS with rs2300370 as the peak *cis*-eQTL colocalizes with COVID-19 hospitalization at PP.H4=0.98, and the secondary GWAS (i.e. after adjusting for rs2300370) with rs2834167 as the peak *cis*-eQTL does not colocalize with COVID-19 hospitalization (PP.H4<0.01).

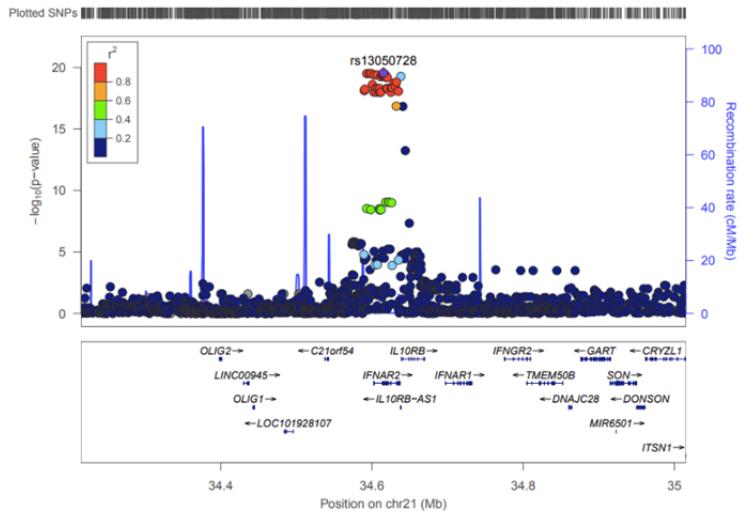
685 Figure 3.



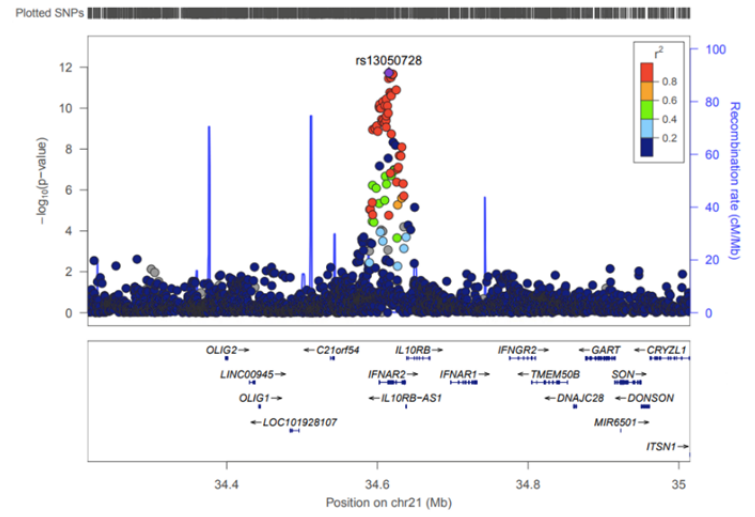
686

687 Figure 4.

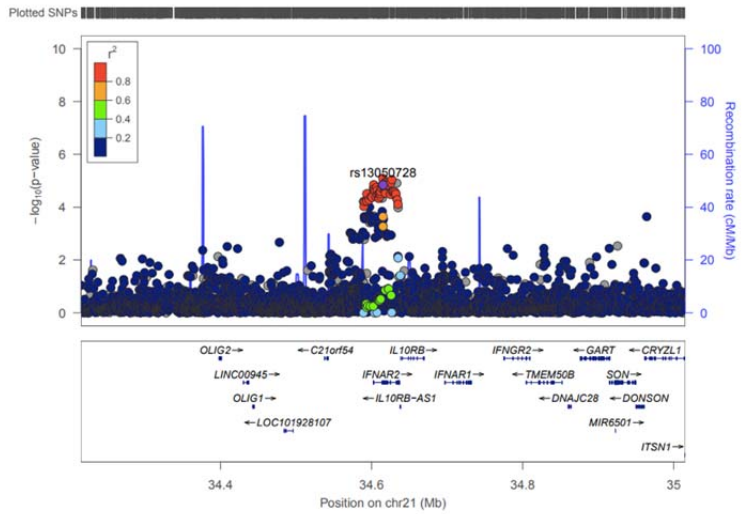
A



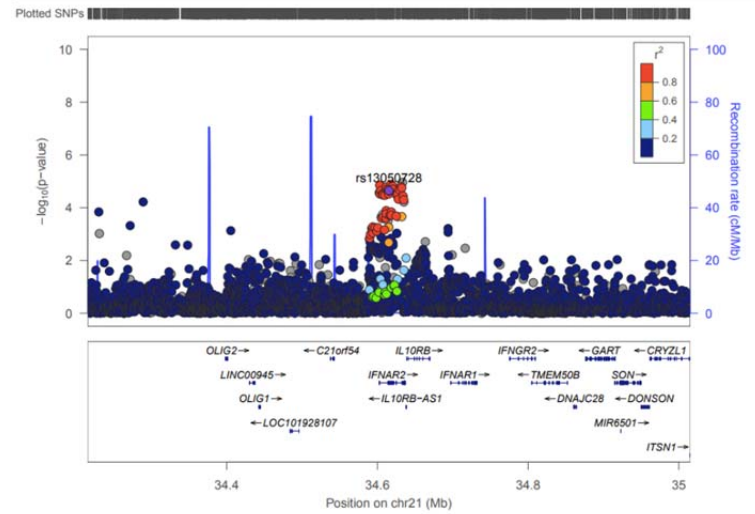
B



C



D



688 References

- 689 1. Horby, P., *et al.* Dexamethasone in Hospitalized Patients with Covid-19 - Preliminary
690 Report. *N Engl J Med* (2020).
- 691 2. Beigel, J.H., *et al.* Remdesivir for the Treatment of Covid-19 - Final Report. *N Engl J*
692 *Med* (2020).
- 693 3. Nelson, M.R., *et al.* The support of human genetic evidence for approved drug
694 indications. *Nat Genet* **47**, 856-860 (2015).
- 695 4. Hingorani, A.D., *et al.* Improving the odds of drug development success through human
696 genomics: modelling study. *Sci Rep* **9**, 18911 (2019).
- 697 5. Cohen, J.C., Boerwinkle, E., Mosley, T.H., Jr. & Hobbs, H.H. Sequence variations in
698 PCSK9, low LDL, and protection against coronary heart disease. *N Engl J Med* **354**,
699 1264-1272 (2006).
- 700 6. Lopalco, L. CCR5: From Natural Resistance to a New Anti-HIV Strategy. *Viruses* **2**,
701 574-600 (2010).
- 702 7. Finan, C., *et al.* The druggable genome and support for target identification and
703 validation in drug development. *Sci Transl Med* **9**(2017).
- 704 8. Swerdlow, D.I., *et al.* The interleukin-6 receptor as a target for prevention of coronary
705 heart disease: a mendelian randomisation analysis. *Lancet* **379**, 1214-1224 (2012).
- 706 9. Kleveland, O., *et al.* Effect of a single dose of the interleukin-6 receptor antagonist
707 tocilizumab on inflammation and troponin T release in patients with non-ST-elevation
708 myocardial infarction: a double-blind, randomized, placebo-controlled phase 2 trial. *Eur*
709 *Heart J* **37**, 2406-2413 (2016).
- 710 10. The COVID-19 Host Genetics Initiative, a global initiative to elucidate the role of host
711 genetic factors in susceptibility and severity of the SARS-CoV-2 virus pandemic. *Eur J*
712 *Hum Genet* **28**, 715-718 (2020).
- 713 11. Gaziano, J.M., *et al.* Million Veteran Program: A mega-biobank to study genetic
714 influences on health and disease. *J Clin Epidemiol* **70**, 214-223 (2016).
- 715 12. Swanson, S.A. & Hernán, M.A. Commentary: how to report instrumental variable
716 analyses (suggestions welcome). *Epidemiology* **24**, 370-374 (2013).
- 717 13. Labrecque, J. & Swanson, S.A. Understanding the Assumptions Underlying Instrumental
718 Variable Analyses: a Brief Review of Falsification Strategies and Related Tools. *Curr*
719 *Epidemiol Rep* **5**, 214-220 (2018).
- 720 14. The Genotype-Tissue Expression (GTEx) project. *Nat Genet* **45**, 580-585 (2013).
- 721 15. Sun, B.B., *et al.* Genomic atlas of the human plasma proteome. *Nature* **558**, 73-79
722 (2018).
- 723 16. Pietzner, M., *et al.* Genetic architecture of host proteins interacting with SARS-CoV-2.
724 *bioRxiv*, 2020.2007.2001.182709 (2020).
- 725 17. Borden, E.C., *et al.* Interferons at age 50: past, current and future impact on biomedicine.
726 *Nat Rev Drug Discov* **6**, 975-990 (2007).
- 727 18. Emilsson, V., *et al.* Co-regulatory networks of human serum proteins link genetics to
728 disease. *Science* **361**, 769-773 (2018).
- 729 19. Staley, J.R., *et al.* PhenoScanner: a database of human genotype-phenotype associations.
730 *Bioinformatics* **32**, 3207-3209 (2016).
- 731 20. von Marschall, Z., *et al.* Effects of interferon alpha on vascular endothelial growth factor
732 gene transcription and tumor angiogenesis. *J Natl Cancer Inst* **95**, 437-448 (2003).

- 733 21. Jia, H., *et al.* Endothelial cell functions impaired by interferon in vitro: Insights into the
734 molecular mechanism of thrombotic microangiopathy associated with interferon therapy.
735 *Thromb Res* **163**, 105-116 (2018).
- 736 22. Casassus, P., *et al.* Treatment of adult systemic mastocytosis with interferon-alpha:
737 results of a multicentre phase II trial on 20 patients. *Br J Haematol* **119**, 1090-1097
738 (2002).
- 739 23. Swanson, S.A., Tiemeier, H., Ikram, M.A. & Hernán, M.A. Nature as a Trialist?:
740 Deconstructing the Analogy Between Mendelian Randomization and Randomized Trials.
741 *Epidemiology* **28**, 653-659 (2017).
- 742 24. Nelson, C.P., *et al.* Genetic Associations With Plasma Angiotensin Converting Enzyme 2
743 Concentration: Potential Relevance to COVID-19 Risk. *Circulation* **142**, 1117-1119
744 (2020).
- 745 25. Wallace, C. Eliciting priors and relaxing the single causal variant assumption in
746 colocalisation analyses. *PLoS Genet* **16**, e1008720 (2020).
- 747 26. Hemnes, A.R., *et al.* A potential therapeutic role for angiotensin-converting enzyme 2 in
748 human pulmonary arterial hypertension. *Eur Respir J* **51**(2018).
- 749 27. Kuba, K., Imai, Y., Rao, S., Jiang, C. & Penninger, J.M. Lessons from SARS: control of
750 acute lung failure by the SARS receptor ACE2. *J Mol Med (Berl)* **84**, 814-820 (2006).
- 751 28. Yang, P., *et al.* Angiotensin-converting enzyme 2 (ACE2) mediates influenza H7N9
752 virus-induced acute lung injury. *Sci Rep* **4**, 7027 (2014).
- 753 29. Monteil, V., *et al.* Inhibition of SARS-CoV-2 Infections in Engineered Human Tissues
754 Using Clinical-Grade Soluble Human ACE2. *Cell* **181**, 905-913.e907 (2020).
- 755 30. Zoufaly, A., *et al.* Human recombinant soluble ACE2 in severe COVID-19. *Lancet*
756 *Respir Med* (2020).
- 757 31. Recombinant Human Angiotensin-converting Enzyme 2 (rhACE2) as a Treatment for
758 Patients With COVID-19. (<https://ClinicalTrials.gov/show/NCT04335136>).
- 759 32. Blanco-Melo, D., *et al.* Imbalanced Host Response to SARS-CoV-2 Drives Development
760 of COVID-19. *Cell* **181**, 1036-1045.e1039 (2020).
- 761 33. Chu, H., *et al.* Comparative Replication and Immune Activation Profiles of SARS-CoV-2
762 and SARS-CoV in Human Lungs: An Ex Vivo Study With Implications for the
763 Pathogenesis of COVID-19. *Clin Infect Dis* **71**, 1400-1409 (2020).
- 764 34. Hadjadj, J., *et al.* Impaired type I interferon activity and inflammatory responses in
765 severe COVID-19 patients. *Science* **369**, 718-724 (2020).
- 766 35. Bastard, P., *et al.* Auto-antibodies against type I IFNs in patients with life-threatening
767 COVID-19. *Science* (2020).
- 768 36. Zhang, Q., *et al.* Inborn errors of type I IFN immunity in patients with life-threatening
769 COVID-19. *Science* (2020).
- 770 37. van der Made, C.I., *et al.* Presence of Genetic Variants Among Young Men With Severe
771 COVID-19. *Jama* **324**, 1-11 (2020).
- 772 38. Lokugamage, K.G., *et al.* SARS-CoV-2 is sensitive to type I interferon pretreatment.
773 *bioRxiv* (2020).
- 774 39. Mantlo, E., Bukreyeva, N., Maruyama, J., Paessler, S. & Huang, C. Antiviral activities of
775 type I interferons to SARS-CoV-2 infection. *Antiviral Res* **179**, 104811 (2020).
- 776 40. Stanifer, M.L., *et al.* Critical Role of Type III Interferon in Controlling SARS-CoV-2
777 Infection in Human Intestinal Epithelial Cells. *Cell Rep* **32**, 107863 (2020).

- 778 41. Clementi, N., *et al.* Interferon- β -1a Inhibition of Severe Acute Respiratory Syndrome-
779 Coronavirus 2 In Vitro When Administered After Virus Infection. *J Infect Dis* **222**, 722-
780 725 (2020).
- 781 42. Dinnon, K.H., *et al.* A mouse-adapted SARS-CoV-2 model for the evaluation of COVID-
782 19 medical countermeasures. *bioRxiv*, 2020.2005.2006.081497 (2020).
- 783 43. Hung, I.F., *et al.* Triple combination of interferon beta-1b, lopinavir-ritonavir, and
784 ribavirin in the treatment of patients admitted to hospital with COVID-19: an open-label,
785 randomised, phase 2 trial. *Lancet* **395**, 1695-1704 (2020).
- 786 44. Pan, H., *et al.* Repurposed antiviral drugs for COVID-19; interim WHO SOLIDARITY
787 trial results. *medRxiv*, 2020.2010.2015.20209817 (2020).
- 788 45. NIAID Stops COVID-19 Trial Enrollment Over Adverse Events. (2020).
- 789 46. Xiao, F., *et al.* Evidence for Gastrointestinal Infection of SARS-CoV-2.
790 *Gastroenterology* **158**, 1831-1833.e1833 (2020).
- 791 47. Bradley, B.T., *et al.* Histopathology and ultrastructural findings of fatal COVID-19
792 infections in Washington State: a case series. *Lancet* **396**, 320-332 (2020).
- 793 48. Mao, L., *et al.* Neurologic Manifestations of Hospitalized Patients With Coronavirus
794 Disease 2019 in Wuhan, China. *JAMA Neurol* **77**, 683-690 (2020).
- 795 49. Puelles, V.G., *et al.* Multiorgan and Renal Tropism of SARS-CoV-2. *N Engl J Med* **383**,
796 590-592 (2020).
- 797 50. Song, E., *et al.* Neuroinvasion of SARS-CoV-2 in human and mouse brain. *bioRxiv*,
798 2020.2006.2025.169946 (2020).
- 799 51. Mendez, D., *et al.* ChEMBL: towards direct deposition of bioassay data. *Nucleic Acids*
800 *Res* **47**, D930-d940 (2019).
- 801 52. Shabalín, A.A. Matrix eQTL: ultra fast eQTL analysis via large matrix operations.
802 *Bioinformatics* **28**, 1353-1358 (2012).
- 803 53. Rohloff, J.C., *et al.* Nucleic Acid Ligands With Protein-like Side Chains: Modified
804 Aptamers and Their Use as Diagnostic and Therapeutic Agents. *Mol Ther Nucleic Acids*
805 **3**, e201 (2014).
- 806 54. Gold, L., *et al.* Aptamer-based multiplexed proteomic technology for biomarker
807 discovery. *PLoS One* **5**, e15004 (2010).
- 808 55. Chapman, A.B., *et al.* A Natural Language Processing System for National COVID-19
809 Surveillance in the US Department of Veterans Affairs. (2020).
- 810 56. Willer, C.J., Li, Y. & Abecasis, G.R. METAL: fast and efficient meta-analysis of
811 genomewide association scans. *Bioinformatics* **26**, 2190-2191 (2010).
- 812 57. Giambartolomei, C., *et al.* A Bayesian framework for multiple trait colocalization from
813 summary association statistics. *Bioinformatics* **34**, 2538-2545 (2018).
- 814 58. Lundberg, M., Eriksson, A., Tran, B., Assarsson, E. & Fredriksson, S. Homogeneous
815 antibody-based proximity extension assays provide sensitive and specific detection of
816 low-abundant proteins in human blood. *Nucleic Acids Res* **39**, e102 (2011).
- 817 59. Olink Target 96 & Target 48 panels for protein biomarker discovery.
- 818 60. Yu, G., Wang, L.G., Han, Y. & He, Q.Y. clusterProfiler: an R package for comparing
819 biological themes among gene clusters. *Omics* **16**, 284-287 (2012).
- 820