

1 **Full Title:** Colocalization of expression transcripts with COVID-19 outcomes is rare across cell states, cell types
2 and organs.

3 **Short Title:** Factors influencing colocalization of gene expression with COVID-19 outcomes.

4 **Authors:** Julian Daniel Sunday Willett^{1,2,3}, Tianyuan Lu^{1,2,3}, Tomoko Nakanishi^{1,2,4,5,6}, Satoshi Yoshiji^{1,2,4,5,6},
5 Guillaume Butler-Laporte¹, Sirui Zhou^{1,2}, Yossi Farjoun¹, J. Brent Richards^{1,2,8,9,10}

6
7 **Corresponding author:**

8 Brent Richards

9 brent.richards@mcgill.ca

10 Pavillon H-413, Jewish General Hospital

11 3755 Cote Ste Catherine

12 Montreal, QC, Canada, H3T 1E2

13

14 **Affiliations:**

15 1. Centre for Clinical Epidemiology, Department of Medicine, Lady Davis Institute, Jewish General Hospital,
16 McGill University, Montréal, Québec, Canada.

17 2. McGill University, Montreal, Quebec, Canada

18 3. Quantitative Life Sciences Program, McGill University, Montreal, Quebec, Canada

19 4. Department of Human Genetics, McGill University, Montreal, Quebec, Canada

20 5. Kyoto-McGill International Collaborative Program in Genomic Medicine, Graduate School of Medicine,
21 Kyoto University, Kyoto, Japan

22 6. Research Fellow, Japan Society for the Promotion of Science, Tokyo, Japan.

23 7. Genome Centre, McGill University, Montreal, Quebec, Canada

24 8. Departments of Medicine, Human Genetics, Epidemiology and Biostatistics, McGill University, Montréal,
25 Quebec, Canada

26 9. Department of Twin Research, King's College London, London, United Kingdom

27 10. Five Prime Sciences Inc, Montréal, Québec, Canada

28

Colocalization of expression transcripts with COVID-19 outcomes is dependent on cell-type and cell-state.

29 **Abstract:**

30 Identifying causal genes at GWAS loci can help pinpoint targets for therapeutic interventions. Expression
31 studies can disentangle such loci but signals from expression quantitative trait loci (eQTLs) often fail to
32 colocalize—which means that the genetic control of measured expression is not shared with the genetic control of
33 disease risk. This may be because gene expression is measured in the wrong cell type, physiological state, or organ.
34 We tested whether Mendelian randomization (MR) could identify genes at loci influencing COVID-19 outcomes
35 and whether the colocalization of genetic control of expression and COVID-19 outcomes was influenced by cell
36 type, cell stimulation, and organ.

37 We conducted MR of *cis*-eQTLs from single cell (scRNA-seq) and bulk RNA sequencing. We then tested
38 variables that could influence colocalization, including cell type, cell stimulation, RNA sequencing modality, organ,
39 symptoms of COVID-19, and SARS-CoV-2 status among individuals with symptoms of COVID-19. The outcomes
40 used to test colocalization were COVID-19 severity and susceptibility as assessed in the Host Genetics Initiative
41 release 7.

42 Most transcripts identified using MR did not colocalize when tested across cell types, cell state and in
43 different organs. Most that did colocalize likely represented false positives due to linkage disequilibrium. In general,
44 colocalization was highly variable and at times inconsistent for the same transcript across cell type, cell stimulation
45 and organ. While we identified factors that influenced colocalization for select transcripts, identifying 33 that
46 mediate COVID-19 outcomes, our study suggests that colocalization of expression with COVID-19 outcomes is
47 partially due to noisy signals even after following quality control and sensitivity testing. These findings illustrate the
48 present difficulty of linking expression transcripts to disease outcomes and the need for skepticism when observing
49 eQTL MR results, even accounting for cell types, stimulation state and different organs.

50

51 **Author Summary:** The genetic determinants of disease and gene expression often do not colocalize (which means
52 they do not share a single causal signal). While some researchers have identified factors that could explain this
53 disconnect, such as immune stimulation or tissue studied, understanding of this complex phenomenon remains
54 incomplete. A deeper understanding could help identify additional genes that mediate disease, affording promising
55 targets for treatment or prevention of disease. We used RNA sequencing data collected at the single cell and bulk
56 tissue level to identify genes whose expression influenced COVID-19 outcomes. We assessed which variables

Colocalization of expression transcripts with COVID-19 outcomes is dependent on cell-type and cell-state.

57 influencing colocalization, including cell type, cell stimulation, RNA sequencing modality, organ, symptoms of
58 COVID-19, and SARS-CoV-2 status among individuals with symptoms of COVID-19. We observed that
59 colocalization of specific candidate genes identified by MR was highly variable and influenced by multiple factors,
60 including cell state and cell population. These results illustrate that even after assessing multiple variables that may
61 influence colocalization, there existed few examples of genes identified by MR that colocalized with gene
62 expression. Future studies would benefit from larger transcriptomics study cohorts and more advanced statistical
63 methods which better account for differences in linkage disequilibrium panels between data sources.

64

65 **Keywords:** COVID-19, colocalization, Mendelian randomization, eQTL, scRNA

Colocalization of expression transcripts with COVID-19 outcomes is dependent on cell-type and cell-state.

66 **Introduction:**

67 Severe COVID-19 is partially influenced by immune hyperstimulation (Merad et al. 2022; Tan et al. 2021).
68 The immune response is mediated by different immune cell subtypes, acting at varying time points during infection
69 across tissues (Ong et al. 2020; Tan et al. 2021). Understanding the dynamics of this process may pinpoint targets
70 helpful for COVID-19 interventions, as previously shown (De Biasi et al. 2020; Kundu et al. 2022; Mathew et al.
71 2020).

72 One way to investigate mechanisms influencing COVID-19 outcomes is to determine the underlying
73 contributory genetic factors. GWAS has identified 87 loci associated with COVID-19 outcomes, but it is often
74 unclear which gene(s) at such loci drive this association (Covid-19 Host Genetics Initiative 2021). Resolving a
75 GWAS locus to its causal gene(s) is non-trivial (Forgetta et al. 2022). One way to identify causal genes at GWAS
76 loci is to examine whether associated SNPs influence outcomes in an appropriate cell type. However, genetic
77 determinants of gene expression (expression quantitative trait loci; eQTL) have often failed to “colocalize” with
78 disease outcomes (Connally et al. 2022). Colocalization, in this context, means that gene expression and the disease
79 outcome share a single common causal SNP (Connally et al. 2022). This lack of colocalization is concerning and not
80 fully resolved. However, given the central importance of gene expression in disease incidence and progression,
81 efforts are required to explain the paradox that the genetic determinants of gene expression often appear different
82 than the those of disease, even for known causal genes in known causal cell types or tissues (Connally et al. 2022).
83 We sought to determine if this lack of colocalization could be explained when cell type, cell stimulation, method of
84 sequencing and organ were taken into account.

85 One factor that may influence colocalization is the population of cells studied. Gene expression is typically
86 determined in bulk tissue, which provides a mixture of cells from the tissue. Such signal dilution, combined with
87 complex factors such as cell-cell interactions, may explain why bulk tissue eQTLs often fail to colocalize with
88 disease outcomes (Connally et al. 2022). Single-cell sequencing studies (scRNA-seq) assay gene expression in
89 specific cell types and thus single-cell sequencing can provide a less heterogeneous assessment of gene expression.
90 Comparing colocalization between bulk and single-cell sequencing, studying additional variables that influence the
91 association, could resolve the contributory factors. This knowledge could help identify causal genes at GWAS loci
92 and accelerate drug development by targeting pathways causal for disease. Targets supported by MR and
93 colocalization evidence are more likely to anticipate clinical trial results, where the target of the medicine in the trial

Colocalization of expression transcripts with COVID-19 outcomes is dependent on cell-type and cell-state.

94 is a circulating protein and its causal influence upon disease is supported by both MR and colocalization. (Zheng et
95 al. 2020).

96 Several studies have investigated the relationship between gene expression and COVID-19 outcomes using
97 older releases of expression data or COVID-19 outcomes. Pairo-Castineira *et al.* found that increased *TYK2* and
98 decreased *IFNAR2* expression in whole blood were associated with life-threatening COVID-19 (Pairo-Castineira et
99 al. 2021). Schmiedel *et al.* found several genes whose expression in specific immune cell types and tissues,
100 including resting and activated naive CD4+ cells, influenced and colocalized with genetic determinants of COVID-
101 19 outcomes (Schmiedel et al. 2021). D'Antonio et al. found genes that colocalized with COVID-19 loci in whole
102 blood, including *ABO* and *IFNAR2*, and identified the causal variants using fine-mapping (D'Antonio et al. 2021).

103 Recently, Soskic *et al.* profiled the changes in gene expression in CD4+ T-cells following stimulation with
104 anti-CD3/anti-CD28 human T-activators (Soskic et al. 2022). We aimed to determine if cell and cell-state specific
105 gene expression could identify novel determinants of COVID-19 outcomes, suggesting which cells are responsible
106 for COVID-19 mortality risk and when. Further, we aimed to determine if such cellular specificity may clarify
107 colocalization of expression and GWAS data. Repeating this analysis in other cell types, bulk whole-blood in
108 individuals with symptoms of COVID-19, with and without a recent PCR-confirmed infection, and bulk tissues in
109 individuals assessed prior to the pandemic lacking any symptoms would identify differences in colocalization due to
110 sequencing modality and different clinical states. The results would afford insights into whether the genetic control
111 of gene expression and disease risk is clarified when resolving to single cells, specific cellular states and clinical
112 characteristics of patients sampled.

113 To answer these questions, we undertook a four-stage study design (**Fig. 1**). First, we conducted Mendelian
114 randomization (MR) of *cis*-eQTLs obtained from single-cell RNA-sequencing (scRNA-seq) data from CD4+ T-cell
115 subtypes at varying times after stimulation with gene expression as exposures and COVID-19 severity as an
116 outcome. Second, we tested these MR-identified genes for colocalization with COVID-19 outcomes across all time
117 points following stimulation of CD4+ T-cells. Third, we compared these colocalization results to those from bulk
118 whole-blood RNA sequencing obtained from individuals with COVID-19 symptoms, who were either SARS-CoV-2
119 PCR positive or negative. Fourth, we compared the colocalization results to bulk unstimulated whole blood and 47
120 other tissues obtained from individuals assessed in GTEx v.8, whose tissues were apparently undiseased at time of
121 sampling.

Colocalization of expression transcripts with COVID-19 outcomes is dependent on cell-type and cell-state.

122 These findings identified 33 genes whose expression may influence COVID-19 outcomes. Colocalization
123 was highly variable and not consistent across the factors that may influence it. These results underline the
124 complexity of factors that influence the colocalization of genetic expression with disease.

125

126 **Results:**

127 *Cohort demographics.* All datasets used in this study included individuals solely of European ancestry to reduce
128 potential confounding by population stratification. ScRNA-seq data for CD4+ T cells in whole blood was obtained
129 from Soskic *et al.*, who isolated cells from 119 healthy, British-ancestry individuals, with a mean age of 47 years,
130 where 44% were females (Soskic et al. 2022). Bulk whole-blood RNA sequencing of individuals with symptoms of
131 COVID-19, who were SARS-CoV-2 positive or negative, was obtained from BQC19, a Quebec cohort of
132 individuals recruited from hospitals presenting with COVID-19 symptoms. BQC19 RNA-sequencing data
133 comprised 112 individuals with symptoms of COVID-19 and SARS-CoV-2 positive PCR tests and 166 individuals
134 who also had symptoms of COVID-19 but a SARS-CoV-2 negative PCR test. The mean age of BQC19 participants
135 was 54 years and 53% were females. The GTEx Consortium cohort comprised 838 post-mortem donors, including
136 715 individuals of European ancestry, of which 33.5% were female (GTEx Consortium 2020).

137

138 *MR and sensitivity testing.* To identify genes influencing COVID-19 outcomes, we used either a Wald ratio or
139 inverse-variance weighted MR analysis at 81 non-MHC COVID-19 associated loci (**Supplemental Table 1**) across
140 three COVID-19 outcomes (severe disease, hospitalized disease, and susceptibility to disease) with sensitivity
141 analyses for each test, detailed in **Methods**. We identified 33 genes whose expression was shown by MR to
142 influence COVID-19 severity and susceptibility (**Fig. 2, Table 1, Supplemental Table 2**).

143 *Single-cell results.* Of the 29,430 combinations of CD4+ cell:stimulation-state:gene:outcome for the whole-blood
144 single-cell eQTLs, 1,225 transcripts (4.2%) had a multiple-testing, Benjamini-Hochberg adjusted p-value ≤ 0.05 in
145 the MR analyses. We limited results to only those arising from *cis*-eQTLs to reduce potential bias from horizontal
146 pleiotropy. *Cis*-eQTLs were defined as genetic variants associated with transcript level within ± 500 kb of the
147 transcriptional start site. We also only retained *cis*-eQTLs that were within ± 500 kb of the lead SNP associated with
148 a COVID-19 outcomes (**Supplemental Table 3**). Only 13 of these 1,225 *cis*-eQTLs passed colocalization sensitivity
149 testing, defined as having a probability of the gene's expression and the COVID-19 outcome sharing a single causal

Colocalization of expression transcripts with COVID-19 outcomes is dependent on cell-type and cell-state.

150 variant (PP_{H4}) greater than 0.80 (**Fig. 1-2**). These data demonstrate that only a small proportion of *cis*-single-cell
151 eQTLs identified via MR colocalized with COVID-19 outcomes, suggesting that such MR findings do not
152 consistently reflect a common causal genetic signal shared between the transcript and COVID-19 outcomes.

153

154 *Bulk results.* Next, we assessed *cis*-eQTLs from the bulk whole-blood RNA sequencing in the BQC19 cohort, which
155 comprised individuals with symptoms of COVID-19 who had either a positive, or negative PCR test for SARS-
156 CoV-2. Of the 5,105 combinations of study-group:gene:outcome for the BQC19 *cis*-eQTLs, 80 transcripts (1.6%)
157 were identified by MR to have effects on COVID-19 outcomes (**Supplemental Table 4**). Only 4 of these 80
158 transcripts passed colocalization sensitivity testing (**Fig. 1-2**). We next assessed whether we would observe similar
159 results if we used whole blood and other organ bulk RNA sequencing from GTEx v.8 (GTEx Consortium 2020) as
160 the source of the *cis*-eQTLs, that included individuals without symptoms or seropositivity for COVID-19. In GTEx
161 bulk whole-blood, we found that of the 2,660 combinations of gene:outcome for the GTEx *cis*-eQTLs, 75 (2.8%)
162 were identified by MR testing to have effects upon COVID-19 outcomes, with only two colocalizing (**Fig. 1-2**). In
163 all 48 GTEx tissues, we found that of the 125,088 combinations of tissue:gene:outcome tested, 3190 (2.6%) were
164 estimated to influence COVID-19 outcomes by MR but only 98 of these colocalized (**Fig. 2**).

165

166 Thus, taken together, across the three different sources of gene expression data (scRNA-seq whole blood CD4+ T-
167 cells, bulk whole blood RNA sequencing in patients with symptoms of COVID-19, and bulk RNA sequencing in
168 individuals whose tissues were apparently health across 48 tissues, including whole blood), we observed 33 unique
169 putatively causal transcripts across 115 specific states that colocalized, which represent 2.3% of those transcripts
170 that survived MR testing and multiple testing thresholds. These findings suggest that most transcripts identified to be
171 associated with COVID-19 outcomes via MR fail to colocalize even across single cell and bulk sequencing, as well
172 as different cellular states and patient states.

173

174 *MR implicated several cis-eQTLs that increase or decrease the risk for COVID-19 outcomes with few overlapping*
175 *transcripts between scRNA-seq and bulk RNA sequencing results.* Across outcomes, all cell types from the scRNA-
176 seq whole-blood data had at least one *cis*-eQTL estimated to be causal for a COVID-19 outcome via MR, except T
177 effector memory re-expressing CD45RA and T regulatory cells (**Fig. 2**). Of the four genes estimated to have

Colocalization of expression transcripts with COVID-19 outcomes is dependent on cell-type and cell-state.

178 colocalized causal effects from the whole-blood scRNA-seq MR experiments, *RALGDS* expression was the only *cis*-
179 eQTL that decreased the risk of COVID-19 outcomes. Specifically, *RALGDS* expression reduced the risk of severe
180 disease (OR = 0.78, 95% CI: 0.71-0.87; adjusted $p = 1.1 \times 10^{-4}$) and susceptibility to disease (OR = 0.88, 95% CI:
181 0.86-0.90; adjusted $p = 7.4 \times 10^{-18}$), when expressed in solely T effector memory cells 16 hours after stimulation
182 (**Fig. 2**). *RALGDS*' influence on hospitalized COVID-19 was not clearly different from the null (OR = 0.87, 95% CI:
183 0.79-0.96; adjusted $p = 0.08$) but passed MR sensitivity testing and colocalized. *NAPSA* expression in T naive cells
184 40 hours after stimulation increased risk of hospitalized (OR = 1.17, 95% CI: 1.08-1.28; adjusted $p = 5.0 \times 10^{-3}$) and
185 susceptibility to (OR = 1.08, 95% CI: 1.05-1.11; adjusted $p = 1.7 \times 10^{-5}$) COVID-19 (**Fig. 2**). *NAPSA*'s influence on
186 severe COVID-19 was not different from the null (OR = 1.20, 95% CI: 1.06-1.37; adjusted $p = 0.07$) but passed MR
187 sensitivity testing and colocalized. *NAPSA* also increased risk for every outcome in T effector memory cells 40
188 hours after stimulation (OR = 1.18, 95% CI: 1.10-1.27; adjusted $p = 4.0 \times 10^{-4}$ for severe, OR = 1.15, 95% CI: 1.10-
189 1.20; adjusted $p = 5.9 \times 10^{-9}$ for hospitalized, OR = 1.04, 95% CI: 1.03-1.06; adjusted $p = 1.2 \times 10^{-4}$ for
190 susceptibility) (**Fig. 2**).

191 Interestingly, none of the colocalizing transcripts from scRNA-seq overlapped with colocalizing bulk RNA
192 sequencing *cis*-eQTLs from BQC19 or GTEx in matching tissues. Generally, signals in bulk sequencing were harder
193 to separate from noise, compared to single-cell results. Increased *IFNAR2* expression was protective in individuals
194 with symptoms of COVID-19 without PCR-confirmed SARS-CoV-2 against severe COVID-19, or individuals who
195 were negative for or had perhaps not yet tested positive for COVID-19 (OR = 0.75, 95% CI: 0.66-0.86; adjusted $p =$
196 3.0×10^{-3}) with insufficient evidence to suggest protection for individuals who did test positive (adjusted $p = 0.04$)
197 where it failed weighted mode MR sensitivity testing but colocalized. In contrast, *IFNAR2* was protective for only
198 individuals who tested positive for SARS-CoV-2 against hospitalized COVID-19 (OR = 0.86, 95% CI: 0.80-0.94;
199 adjusted $p = 0.03$) with insufficient evidence to suggest protection for BQC19 SARS-CoV-2 negative individuals
200 (adjusted $p = 8.1 \times 10^{-3}$) where it failed MR Egger intercept sensitivity testing ($p = 0.02$) while colocalizing (**Fig. 2**).
201 This was perhaps a false negative. There were several other genes that colocalized for only a subset of outcomes in
202 other tissues, although some consistently colocalized in the same tissue across outcomes, such as ABO in the testis
203 and DPP9 in sun-exposed skin (**Table 1, Supplemental Table 2**).

204

Colocalization of expression transcripts with COVID-19 outcomes is dependent on cell-type and cell-state.

205 *Colocalization of specific MR-identified cis-eQTLs depended on cell stimulation.* To investigate the variables that
206 influenced colocalization, we conducted colocalization on each gene that was estimated causal in MR. Most MR-
207 identified transcripts did not colocalize (**Fig. 1, Supplemental Tables 6-8**). The proportion of estimated causal
208 transcripts that colocalized and passed single causal variant sensitivity testing was 1.1% for whole blood single-cell
209 eQTLs, 5.0% for BQC19, 2.7% for GTEx whole blood, and 3.1% for all organs in GTEx (**Fig. 3**). Colocalization of
210 2/4 single-cell eQTLs was specific to cell type (**Fig. 2**) with 3/3 from stimulated cells specific to cell state with
211 colocalization for *RALGDS*, for example, specific to T effector memory cells 16 hours post-stimulation for severe
212 and susceptibility to COVID-19 (**Fig. 4**).

213 Colocalization of transcripts occurred at varying times following stimulation (**Fig. 2**). Colocalization did
214 not appear to depend on SARS-CoV-2 PCR result in BQC19 among individuals with symptoms of COVID-19, as
215 seen with *IFNAR2* (**Fig. 2, Fig. 5**). Colocalization of select transcripts appeared organ specific, as with *MUC5B*
216 (**Fig. 6**), although there was more noise in bulk sequencing data, as was observed for *ABO*. Specifically, *ABO*
217 colocalized in some organs for only a single outcome and *IL10RB* tested in GTEx had an opposite direction of effect
218 in cultured fibroblasts than tibial nerve (**Fig. 2**).

219 While many transcripts did not colocalize, many transcripts that did have evidence of the single causal
220 variant assumption being violated, with multiple peaks present within or close to the cognate 1 Mb window,
221 suggesting bias due to linkage disequilibrium (**Fig. 7**). Of 19 colocalizing transcripts in single-cell CD4+ T cells, 6
222 had evidence of violating this assumption. Of 9 colocalizing transcripts in bulk whole blood from patients with
223 symptoms of COVID-19 in BQC19, we observed 6 that violated this assumption. Of 4 from bulk whole blood in
224 GTEx, 2 violators. Of 378 from all tissues in GTEx, 265 transcripts violated this assumption. These results underpin
225 the limitations of present study sample sizes and existing methods designed to clarify causal colocalizing expression
226 signals.

227

228 **Discussion:**

229 In this study, we attempted to identify factors that influence colocalization of *cis*-eQTL MR findings for
230 COVID-19 outcomes. Most MR-identified single-cell eQTLs and bulk eQTLs did not colocalize, suggesting that
231 linkage disequilibrium and limited study sizes yielding fewer converging signals may strongly impact the validity of
232 eQTL MR studies that assess colocalization, and necessitate more stringent quality control. Previous research has

Colocalization of expression transcripts with COVID-19 outcomes is dependent on cell-type and cell-state.

233 demonstrated a lack of colocalization of expression findings and suggested that this may be resolved by single-cell
234 eQTL analyses, assessing stimulation state, or clearly defining the state of the individual when blood samples were
235 drawn. Here we show that colocalization of these signals is highly variable, and it is not fully explained by changing
236 cell type, cell stimulation, symptoms of COVID-19, or organ. Taken together, these findings suggest that
237 colocalization of eQTLs with disease outcomes is difficult using current technologies, reference panels and
238 statistical methods, and with present study sample sizes that in single-cell studies typically is limited to 100
239 individuals.

240 While we overall found colocalization to be limited, our results were reassuring in that we observed some
241 trends and results consistent with past studies and hypotheses (**Fig. 2**) (Connally et al. 2022; Soskic et al. 2022). For
242 putatively causal transcripts from CD4+ T cells in whole blood, colocalization was influenced by cell stimulation, or
243 cell state (**Fig. 2a**), like other groups' observations when conducting colocalization without MR (Soskic et al. 2022).
244 On comparing putatively causal transcripts between single-cell and bulk results, with the latter evaluating transcripts
245 in several different cells in a matching tissue, we found no overlap between the modalities in the same tissue, that
246 could suggest that sequencing modality plays a role (**Fig. 2**). Genes that colocalized in CD4+ T cells did not
247 colocalize in spleen that is enriched with T cells (**Fig. 2**). This could underscore the impact of sequencing multiple
248 cell types in bulk sequencing, suggesting that results must be contextualized against sequencing modality and
249 sample cellular heterogeneity. We found organ to play a moderate role in colocalization when comparing putatively
250 causal genes in GTEx (**Fig. 2b**), supporting others' findings with non-COVID-19 outcomes (Rocheleau et al. 2022).
251 The data on organs' roles were perhaps noisy. While we found outcome to have a limited role in influencing
252 colocalization for select transcripts, different from others (Rocheleau et al. 2022; Soskic et al. 2022), this could be
253 due to the greater similarity in our outcomes.

254 This paper has limitations. The scRNA-seq data came from cells stimulated by a standard T cell stimulator
255 rather than SARS-CoV-2. Such stimulation has been employed by several existing works studying immunity in
256 COVID-19 (De Biasi et al. 2020; Kundu et al. 2022; Mathew et al. 2020) and may have some relevance to the
257 stimulation endured when T-cells encounter SARS-CoV-2. We only investigated CD4+ T cells in whole blood (**Fig.**
258 **2**). There are other cell types present in greater proportions, which could explain the minimal overlap in
259 colocalization between our single-cell and bulk sequencing results. We could only locate whole blood *cis*-sceQTLs
260 from an admixed study that did not release ancestry-stratified results (Randolph et al. 2021). Ancestry-stratified

Colocalization of expression transcripts with COVID-19 outcomes is dependent on cell-type and cell-state.

261 results are important for Mendelian randomization and colocalization to limit bias from indirect pleiotropic
262 mechanisms, including linkage disequilibrium that varies by ancestral group. Our GTEx results show that many
263 tissues could contribute to COVID-19 outcomes and pathogenesis (**Fig. 2**). Single-cell analysis of all tissues,
264 particularly lung, could help understand the multi-system basis of post-COVID-19 syndrome (Mehandru and Merad
265 2022). We were unable to access single-cell lung data for this work (Lamontagne et al. 2018). Our outcomes were
266 limited to COVID-19 severity and susceptibility, when expression could also influence complications of COVID-19,
267 such as post-COVID-19 syndrome and schizophrenia (Baranova et al. 2022b; Mehandru and Merad 2022). We have
268 developed two-step MR methods that link gene expression to COVID-19 outcomes, and then to these complications
269 (Yoshiji et al. 2023), which has already been used to validate BMI's role on COVID-19 outcomes as discovered by
270 others (Baranova et al. 2023).

271 Our study was limited to individuals of European ancestry. While the HGI has found loci with genome-
272 significant variants in other ancestries, none of the loci from Admixed American, African, East Asian, or South
273 Asian ancestry for any outcome overlapped with loci found to influence COVID-19 outcomes in this study. Given
274 fewer loci in non-European datasets, this could be due to sample size and underscores the importance of multi-
275 ancestry analyses. However, even within European ancestry individuals, subtle differences in LD patterns can
276 influence colocalization, which we observed (**Fig. 4-6**) (Kanai et al. 2022; Kanai et al. 2021). Such differences may
277 have impacted the lack of colocalization observed here. Sample size was limited for the eQTL datasets that we
278 employed, which increases the risk of biased results, which we possibly observed where some genes colocalized for
279 only a single outcome or did not colocalize in one outcome (**Fig. 2**). We adjusted for this by using methods well
280 established in the literature, including several stringent and conservative sensitivity tests and means of quality
281 control for MR and colocalization. While using a colocalization window around a lead variant of ± 500 kb is more
282 likely to limit bias from pleiotropy, it increases the risk of missed signals. Data must still be carefully appraised for
283 possible false positives, as may be the case with *IL10RB* being estimated protective in cultured fibroblasts for all
284 outcomes but increasing the risk of COVID-19 outcomes in the tibial nerve for all outcomes (**Fig. 2**).

285 While we found symptoms of COVID-19 to influence genomic colocalization for select transcripts and
286 states, as shown by *IFNAR2* between severe and hospitalized COVID-19, this trend could have been influenced by a
287 batch effect. SARS-CoV-2 status' effect on colocalization among individuals in BQC19 could be biased as negative
288 results could represent false negatives and individuals with COVID-19 symptoms could have had a different viral

Colocalization of expression transcripts with COVID-19 outcomes is dependent on cell-type and cell-state.

289 illness. We mitigated this by deriving eQTLs using the same pipeline used by GTEx and limiting the conclusions we
290 made given this context (GTEx Consortium 2020). Generally, we investigated variables in multiple datasets, given
291 that a variable's demonstrated role in multiple cohorts supports making a generalization.

292

293 **Conclusions:** While existing hypotheses suggest that colocalization of transcripts depends on multiple conditions,
294 we found that *cis*-eQTLs identified by MR for COVID-19 outcomes rarely colocalized, even when assessing
295 different cell types, cell states, symptoms of COVID-19 and organs. Taken together, these findings suggest that
296 even after accounting for variables, there was little evidence of colocalization for most genes whose influence on
297 COVID-19 outcomes was identified through MR.

298

299 **Figure Captions**

300 **Fig. 1.** Study overview

301 **Fig. 2.** A total of 33 genes that colocalized across body tissue, evaluated at the (a) single-cell and (b) bulk tissue
302 level, were estimated to have their expression increase (red) and decrease (blue) risk of COVID-19 outcomes. In
303 bulk tissue, COVID+ referred to individuals who had tested positive for COVID-19 with COVID- for those who
304 tested negative and Symptoms+ refers to individuals who presented with symptoms of COVID-19 with Symptoms-
305 referring to those who did not have symptoms of COVID-19. Note that some transcripts show increased expression
306 in some tissues to be associated with COVID-19 outcomes, whereas other tissues show decreased expression to be
307 associated with the same transcript for the same outcome. TCM: T central memory cell. TEM: T effector memory
308 cell. TN: T naïve cell.

309 **Fig 3.** The proportion of estimated causal variants that colocalized and passed sensitivity testing. Orange bars refer
310 to single cell data, gray bulk sequencing.

311 **Fig 4.** Locuszoom plots for *RALGDS* in the setting of variable cell stimulation and cell type, with selected variant of
312 rs8176719. *RALGDS* was estimated causal with colocalization for only T effector memory (TEM) cells 16 hours
313 post-stimulation, highlighting the role of cell stimulation on colocalization. A * indicates a dataset that colocalized.
314 rs8176719 was not detected in T naïve cells 16 hours post-stimulation, so rs7036642 was highlighted, which is in
315 LD with rs8176719.

Colocalization of expression transcripts with COVID-19 outcomes is dependent on cell-type and cell-state.

316 **Fig 5.** Locuszoom plots for *IFNAR2* in the setting of variable presence of COVID-19 symptoms and SARS-CoV-2
317 positivity, with selected variant of rs9636867. *IFNAR2* was estimated causal and colocalized only in patients with
318 COVID-19 symptoms (Symptoms +/-) with SARS-CoV-2 status (COVID +/-) having a less apparent role. A *
319 indicates a dataset that colocalized.

320 **Fig 6.** Locuszoom plots for *MUC5B* in the setting of variable organ in GTEx v8, with selected variant of
321 rs35705950. *MUC5B* was estimated causal and colocalized only in lung for severe COVID-19. There were no *cis*-
322 eQTLs for whole blood.

323 **Fig 7.** Percentage of colocalizing results ($PP_{H4} \geq 0.80$) that had evidence of violating the single causal variant
324 assumption. Orange bars refer to single cell data, gray bulk sequencing.

325

326

327

328 **Table Captions**

329 **Table 1.** Putatively causal associations of gene expression with COVID-19 outcomes. While some genes had
330 opposite effects on risk for a minority of tissues (bolded), the direction of effect of most was consistent when
331 observed in multiple tissues.

332

333 **Supplemental Table Captions**

334 **Supplemental Table 1.** All lead variants in HGI European-ancestry data, excluding MHC loci.

335 **Supplemental Table 2.** All genes that were estimated causal and colocalized in our study, with existing evidence
336 linking genes to outcomes.

337 **Supplemental Table 3.** All MR results using CD4+ T cell expression data.

338 **Supplemental Table 4.** All MR results using expression data from individuals with symptoms of COVID-19 in
339 BQC19.

340 **Supplemental Table 5.** All MR results using expression data from all tissues in GTEx.

341 **Supplemental Table 6.** All colocalization results of putatively causal variants from CD4+ T cell expression data.

342 **Supplemental Table 7.** All colocalization results of putatively causal variants from individuals with symptoms of
343 COVID-19 in BQC19.

Colocalization of expression transcripts with COVID-19 outcomes is dependent on cell-type and cell-state.

344 **Supplemental Table 8.** All colocalization results of putatively causal variants from all tissues in GTEx.

345

346 **Methods:**

347 **Datasets.** We examined *cis*-eQTLs in three datasets to implicate their influence on COVID-19 outcomes and
348 investigate how experimental and physiological conditions impact colocalization (**Fig. 1**). scRNA-seq *cis*-eQTLs
349 from Soskic *et al.* were used to analyze how cell type, cell stimulation, and cell microenvironment affected
350 colocalization (Soskic et al. 2022). Bulk RNA-seq *cis*-eQTLs from Biobanque Quebecoise de la COVID-19
351 (BQC19) were used to investigate how disease state impacted identified *cis*-eQTLs and colocalization in individuals
352 with symptoms of COVID-19, with and without PCR-confirmed SARS-CoV-2 infection. Bulk RNA-seq *cis*-eQTLs
353 from GTEx whole blood were used to compare BQC19 data with data from individuals without symptoms of
354 COVID-19 with all other organs used to determine the role of organ (GTEx Consortium 2020).

355

356 *Whole-blood single-cell eQTLs.* The summary statistics for immune cell expression *cis*-eQTLs before and after
357 stimulation with anti-CD3/anti-CD28 human T-Activators were obtained from Soskic *et al.* (Soskic et al. 2022). The
358 study consisted of 119 individuals of British ancestry with peripheral blood mononuclear cells (655,349 CD4+ T
359 cells) sequenced using scRNA-seq (Soskic et al. 2022). We used the summary statistics for cells at all available time
360 points (unstimulated, 16 hours post-stimulation corresponding to before cell division, 40 hours post-stimulation
361 corresponding to after cell division, five days post-stimulation corresponding to gaining effector function) for cell
362 types present before stimulation and present for at least one time point after stimulation (Soskic et al. 2022). The
363 unstimulated time point acted as a control for stimulation, sequenced 16 hours after culturing without any anti-
364 CD3/anti-CD28 human T-Activators (Soskic et al. 2022). We investigated CD4+ antigen-I and CD4+ memory cell
365 classifications before Leiden-algorithm clustering, implemented by Soskic et al., and T I, T central memory, T
366 effector memory, CD45RA re-expressing T effector memory, and thymus-derived regulatory T cells after clustering
367 (Soskic et al. 2022). Full details describing RNA sequencing, separation of cell types and stimulation are available
368 in Soskic *et al.* (Soskic et al. 2022).

369 *Bulk whole blood eQTLs from subjects with symptoms of COVID-19, with and without SARS-CoV-2 positive PCR*

370 *results.* BQC19 (<https://en.quebecovidbiobank.ca>) is a prospective cohort enrolling participants with PCR-proven

371 SARS-CoV-2 infection and PCR-proven SARS-CoV-2 negative individuals who presented to the hospital with signs

Colocalization of expression transcripts with COVID-19 outcomes is dependent on cell-type and cell-state.

372 or symptoms consistent with COVID-19. Participants were recruited from eight academic hospitals in the province
373 of Quebec, Canada. A total of 4,704 participants underwent PCR testing with confirmed positive or negative results
374 between January 25th, 2020 to March 20th, 2022. A total of 379 participants had RNA extracted and sequenced with
375 eQTLs called using the GTEx pipeline, available from the Broad Institute ([https://github.com/broadinstitute/gtex-](https://github.com/broadinstitute/gtex-pipeline/tree/master/qlt)
376 [pipeline/tree/master/qlt](https://github.com/broadinstitute/gtex-pipeline/tree/master/qlt)). We used data solely for those of non-Finnish European ancestry, which included 112
377 SARS-CoV-2 positive and 166 SARS-CoV-2 negative samples.

378
379 *GTEx release 8 data.* We obtained *cis*-eQTL summary statistics for GTEx release 8 for whole-blood and all other
380 organs, restricted to individuals of European ancestry, from: <https://www.gtexportal.org/home/> (GTEx Consortium
381 2020).

382
383 *COVID-19 Outcome Data.* European-ancestry-specific summary statistics for COVID-19 outcomes were obtained
384 from the COVID-19 Host Genetics Initiative (HGI) release 7 (<https://www.covid19hg.org/>), which did not include
385 individuals from 23AndMe (Covid-19 Host Genetics Initiative 2021). The COVID-19 outcomes included severe
386 COVID-19 (13,769 cases and 1,072,442 controls), COVID-19 hospitalization (32,519 cases and 2,062,805 controls),
387 and susceptibility to COVID-19 (122,616 cases and 2,475,240 controls). Severe COVID-19 was defined as COVID-
388 19 requiring respiratory support or resulting in death. COVID-19 with hospitalization was defined as an infection
389 requiring hospitalization or death. COVID-19 susceptibility was defined as infection determined by self-report on a
390 questionnaire, electronic medical record diagnosis, or laboratory testing (serology tests or nucleic acid
391 amplification). Controls were all individuals who did not meet an outcome's definition.

392
393 **MR of *cis*-eQTLs with COVID-19 outcomes.** We used MR to estimate the causal relationship between exposures
394 (which here are RNA transcript levels) and COVID-19 outcomes to determine how cell type, cell stimulation, time
395 after stimulation, symptoms of COVID-19, PCR result for SARS-CoV-2 among individuals with symptoms of
396 COVID-19, and organ influenced the relationship between gene expression and COVID-19 outcomes. MR studies
397 use SNPs strongly associated with an exposure as instrumental variables to estimate the effect of an exposure on an
398 outcome. Such studies reduce potential confounding effects because genetic variants are essentially randomized at
399 conception. They also prevent bias due to reverse causation (wherein the outcome influences the exposure) since the

Colocalization of expression transcripts with COVID-19 outcomes is dependent on cell-type and cell-state.

400 assignment of genetic variants always precedes disease onset (Smith and Ebrahim 2003). The main assumptions of
401 MR are that the variants under study are strongly associated with the risk factor of interest, confounders of the
402 exposure-outcome relationship are not associated with the variants, and the variants only affect the outcome through
403 the risk factor (Davies et al. 2018; Skrivanekova et al. 2021). The most problematic of these assumptions is the last,
404 as it is difficult to confidently understand if the SNPs affect the outcome independent of the exposure (i.e. a lack of
405 horizontal pleiotropy). To partially mitigate against such potential bias, we have used only *cis*-eQTLs, which are
406 more likely to act directly through the transcription or translation of the proximal gene, rather than through
407 horizontally pleiotropic pathways. There was overlap of individuals from the *cis*-eQTL studies with individuals in
408 the HGI data.

409 To undertake MR analyses, we used the TwoSampleMR (v0.5.6) package
410 (<https://mrcieu.github.io/TwoSampleMR/>). First, we identified all genome significant ($p \leq 5 \times 10^{-8}$) loci from HGI
411 COVID-19 summary statistics. We isolated all variants 500 kb upstream and downstream of the lead HGI variant for
412 each locus for each exposure and outcome. Second, we excluded the MHC locus (chr6: 28,510,120 – 33,480,577;
413 GRCh38) to reduce potential confounding by linkage disequilibrium structure. Third, the exposure *cis*-eQTLs, ± 500
414 kb from their transcriptional start sites, for each cell type, cell stimulation state, cell stimulation time, SARS-CoV-2
415 status, presence of COVID-19 symptoms, and gene were filtered for variants in LD using the package's clumping
416 function with default settings. Fourth, the exposure data were harmonized with outcome data using default settings.
417 Fifth, we conducted MR with sensitivity tests. For loci with only one remaining *cis*-eQTL following harmonization,
418 we applied Wald Ratio-based MR. For loci with more than one remaining independent *cis* variant, we used inverse-
419 variance weighted MR (MR-IVW). For loci with three or more remaining *cis*-eQTLs, we did sensitivity testing
420 using exposure MR weighted median, MR weighted penalized median, MR weighted mode, MR Egger regression,
421 and MR Egger intercept. MR Steiger was done for every test to assess for reverse causation, wherein the MR results
422 would be better explained by the effects of COVID-19 on the *cis*-eQTL. Results were retained for downstream
423 analysis if they passed these sensitivity tests with a Wald Ratio or IVW Benjamini-Hochberg adjusted p-value,
424 which controls the false-discovery rate by adjusting the p-value by the number of tests, less than or equal to 0.05.

425

426 **Colocalization.** To investigate whether *cis*-eQTLs shared the same single causal signal with COVID-19 outcomes,
427 we used coloc v5.1.1 (<https://chr1swallace.github.io/coloc/>). Colocalization helps to guard against bias due to

Colocalization of expression transcripts with COVID-19 outcomes is dependent on cell-type and cell-state.

428 confounding from linkage disequilibrium. Such confounding can occur when the SNPs that influence an exposure
429 (here *cis*-eQTLs) do not causally influence an outcome (here COVID-19 outcomes) but are associated with each
430 other due to linkage disequilibrium. Consistent with Soskic *et al.*, we required at least 50 variants for each
431 colocalization analysis (Soskic et al. 2022). We employed default priors of $p_1=p_2=10^{-4}$ and $p_{12}=10^{-5}$ where p_1 is the
432 prior probability that only eQTLs had a genetic association in the region, p_2 is a prior probability that only the HGI
433 summary statistics had a genetic association in the region, and p_{12} is the prior probability that the eQTL data and the
434 HGI summary statistics shared the same genetic associations in the region. A $p_{12} \geq 0.8$, or $PP_{H4} \geq 0.8$ was
435 considered evidence of colocalization. Sensitivity tests were conducted using coloc's sensitivity test function for
436 each instance of colocalization to validate the single causal variant assumption and evaluate the robustness of results
437 to different prior settings. Locuszoom plots used to highlight trends were produced using Locuszoom (Pruim et al.
438 2010).

439

440 **Ethics statement.** Soskic *et al.* biological samples were ethically sourced and used in research under an institutional
441 review board-approved protocol (Soskic et al. 2022). BQC19 received ethical approval from the Jewish General
442 Hospital research ethics board (2020-2137) and the Centre Hospitalier de l'Université de Montréal institutional
443 ethics board (MP-02-2020-8929, 19.389). Specimens collected by GTEx were sourced per protocol for an
444 institutional review board (GTEx Consortium 2020).

445

446 **Data availability.** Data from Soskic *et al.* are available through Zenodo (doi:10.5281/zenodo.6006796). BQC-19
447 data, including expression data, is available by application: <https://www.bqc19.ca/>. GTEx release 8 data is available
448 from: <https://www.gtexportal.org/home/>. COVID-19 HGI summary statistics are available from:
449 <https://www.covid19hg.org/>.

450

451 **Code availability.** All analyses were conducted in R v4.0.5 (R Core Team 2021). Code is available at:
452 https://github.com/richardslab/dynamic_QTL_COVID19.

453

454 **Acknowledgements.** JDSW is a graduate student in McGill's Quantitative Life Sciences Ph.D. program. This work
455 was made possible through open sharing of data and samples from BQC19. We thank all participants of BQC19 for

Colocalization of expression transcripts with COVID-19 outcomes is dependent on cell-type and cell-state.

456 their contribution. We thank Vincent Mooser, director of BQC19, for his contributions organizing BQC19. Some
457 transcriptomics sequencing was performed at the McGill University Genome Centre (Montreal, Quebec, Canada) by
458 Ariane Boisclair, Janick St-Cyr, Pouria Jandaghi, Daniel Auld. The remainder of the transcriptomics sequencing was
459 performed at the Centre d'expertise et de services Génome Québec (Montreal, Quebec, Canada), facilitated by the
460 diligent work of Alexandre Montpetit, Isabelle Guillet, Joelle Fontaine, Martine Lavoie, and Genevieve DonPierre.
461 Kevin Liang contributed to paper revisions.

462
463 **Funding sources.** This work is supported by the CIHR (JDSW CIHR 476575). BQC19 was funded by the Fonds de
464 recherche du Québec – Santé, Génome Québec and the Public Health Agency of Canada. The Richards research
465 group is supported by the Canadian Institutes of Health Research (CIHR: 365825; 409511, 100558, 169303), the
466 McGill Interdisciplinary Initiative in Infection and Immunity (MI4), the Lady Davis Institute of the Jewish General
467 Hospital, the Jewish General Hospital Foundation, the Canadian Foundation for Innovation, the NIH Foundation,
468 Cancer Research UK, Genome Québec, the Public Health Agency of Canada, McGill University, Cancer Research
469 UK [grant number C18281/A29019] and the Fonds de Recherche Québec Santé (FRQS). JBR is supported by a
470 FRQS Mérite Clinical Research Scholarship. Support from Calcul Québec and Compute Canada is acknowledged.
471 These funding agencies had no role in the design, implementation or interpretation of this study.

472 **Disclosures:** JBR's institution has received investigator-initiated grant funding from Eli Lilly, GlaxoSmithKline and
473 Biogen for projects unrelated to this research. He is the CEO of 5 Prime Sciences (www.5primesciences.com),
474 which provides research services for biotech, pharma and venture capital companies for projects unrelated to this
475 research.

476
477

Colocalization of expression transcripts with COVID-19 outcomes is dependent on cell-type and cell-state.

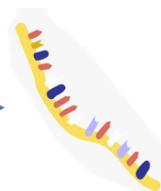
478 **Figures.**

479

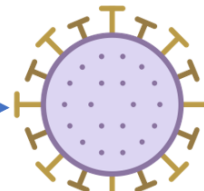
Study Aim: To determine the role of cell type, cell stimulation, COVID-19 symptoms and PCR result among symptomatic patients, and organ on colocalization.



(Un)stimulated CD4+ cells
Whole-blood \pm COVID-19
All GTEx Tissues



Cis-eQTL
Transcripts



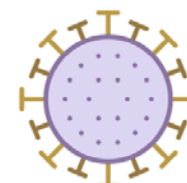
Severe COVID-19
COVID-19 Hospitalization
COVID-19 Susceptibility

480

Mendelian randomization and colocalization of *cis*-eQTLs with COVID-19 Outcomes

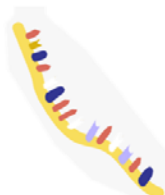


Two-sample MR with sensitivity testing

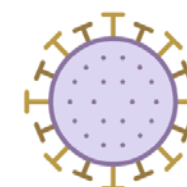


Cell and Cell State (Blood): scRNA-seq: 29,430 tests
Disease State (Blood): bulk RNA-seq: 5,105 tests
Non-Diseased State (Blood): bulk RNA-seq: 2,660 tests

Severe COVID-19
COVID-19 Hospitalization
COVID-19 Susceptibility



Colocalize putative causal genetic expression



Cells and Cell State (Blood): 1,225 transcripts
Disease State (Blood): 80 transcripts
Non-Diseased State (Blood): 75 transcripts

Severe COVID-19
COVID-19 Hospitalization
COVID-19 Susceptibility

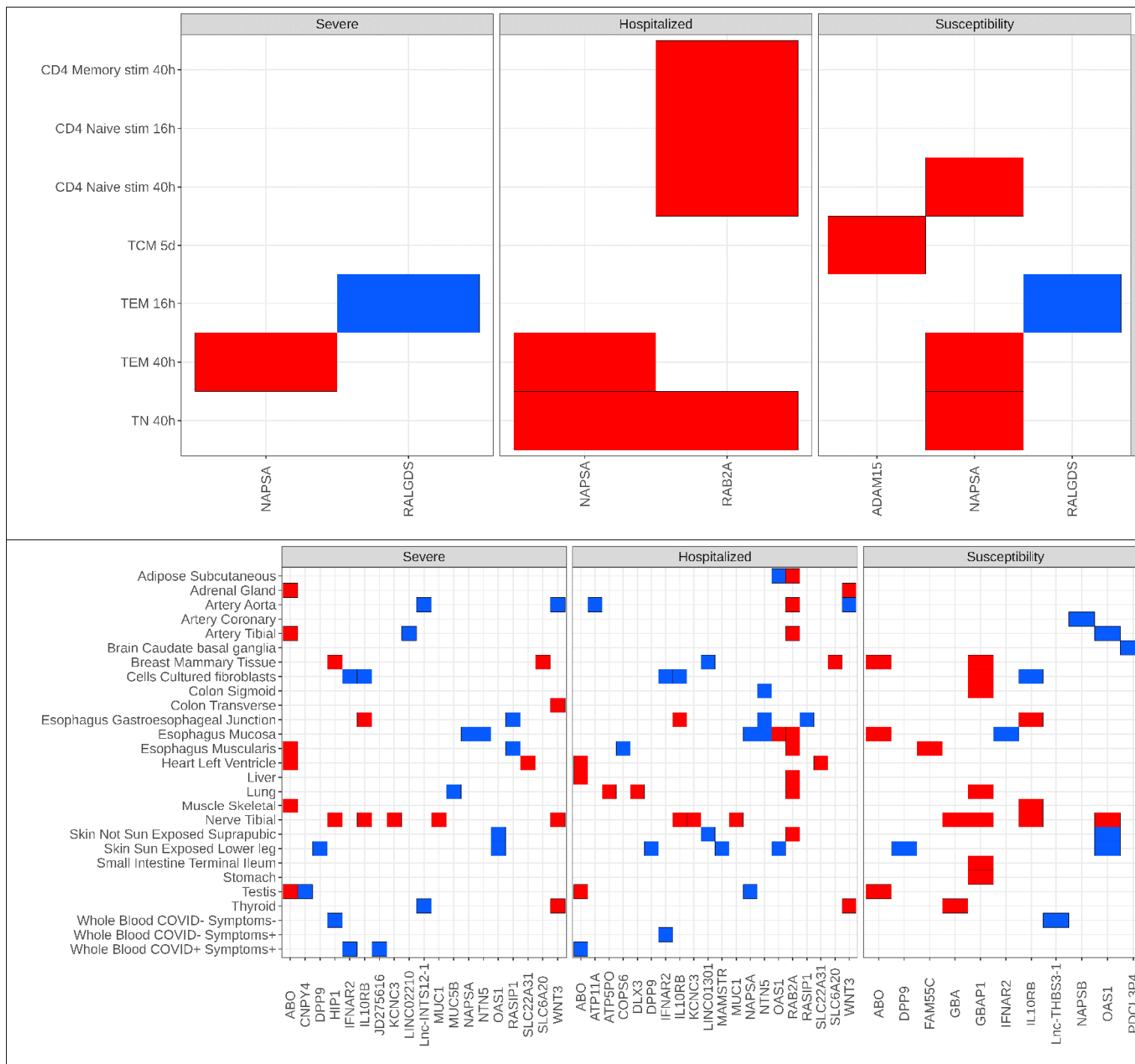
481

482

483

Fig. 1.

Colocalization of expression transcripts with COVID-19 outcomes is dependent on cell-type and cell-state.



484 **Fig. 2.**

485

486

Colocalization of expression transcripts with COVID-19 outcomes is dependent on cell-type and cell-state.

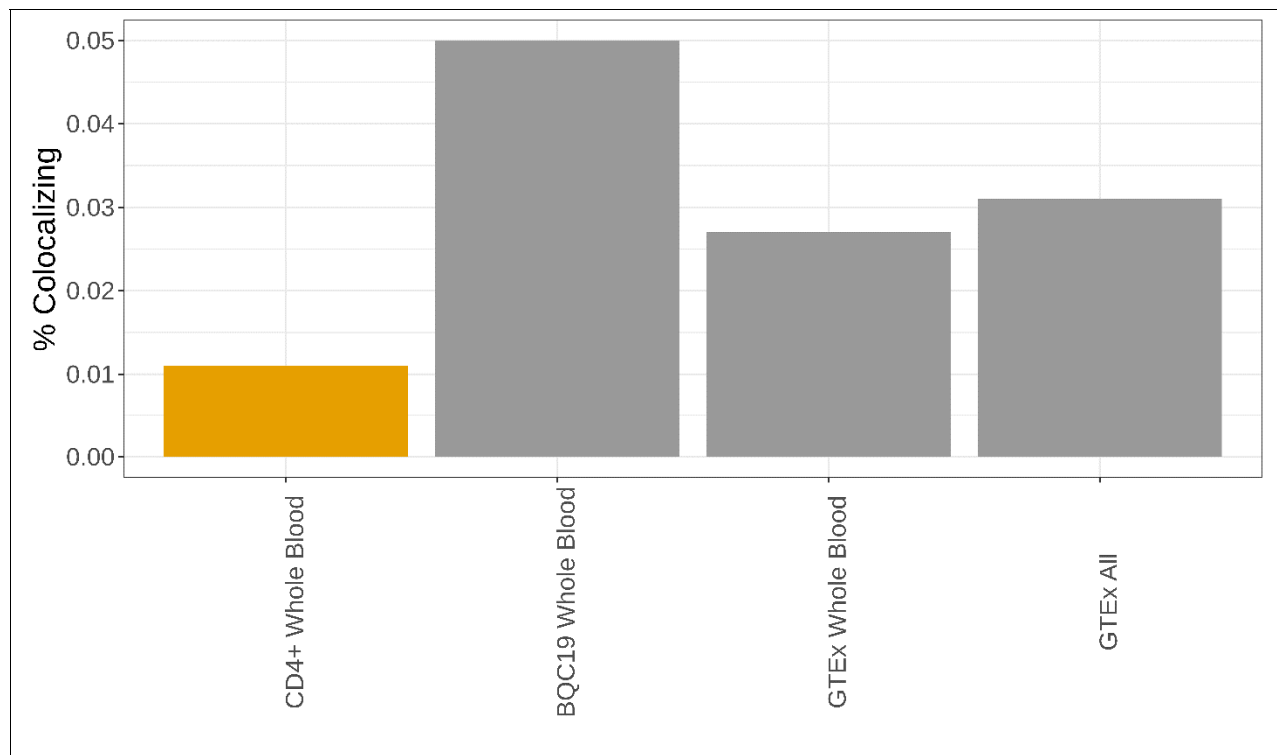
487 **Table 1.**

Gene	Putatively causal cell/tissue	Existing evidence
<i>ABO</i>	Adrenal gland (Severe) – Hazardous Artery tibial (Severe) – Hazardous Breast Mammary Tissue (Susceptible) – Hazardous Esophagus mucosa (Susceptible) – Hazardous Esophagus muscularis (Severe) – Hazardous Heart Left Ventricle (Severe and Hospitalized) – Hazardous Liver (Hospitalized) – Hazardous Muscle skeletal (Severe) – Hazardous Testis (All) – Hazardous Whole blood COVID+ Symptoms+ (Hospitalized) – Protective	MR in non-GTEX and GTEX eQTL and pQTL datasets for lung and whole blood (Baranova et al. 2022a; Hernandez Cordero et al. 2021)
<i>ATP5PO</i>	Lung (Hospitalized) – Hazardous	Associational (Li et al. 2022)
<i>GBAP1</i>	Breast Mammary Tissue (Susceptible) – Hazardous Cultured Fibroblasts (Susceptible) – Hazardous Colon Sigmoid (Susceptible) – Hazardous Lung (Susceptible) – Hazardous Nerve Tibial (Susceptible) – Hazardous Small Intestine Terminal Ileum (Susceptible) – Hazardous Stomach (Susceptible) – Hazardous	MR with colocalization for splice isoform (Nakanishi et al. 2022)
<i>HIP1</i>	Breast mammary tissue (Severe) – Hazardous Nerve tibial (Severe) – Hazardous Whole blood COVID- Symptoms- (Severe) – Protective	None
<i>IFNAR2</i>	Cells cultured fibroblasts (Severe and hospitalized) – Protective Esophagus mucosa (Susceptible) – Protective Whole blood COVID- Symptoms+ (Hospitalized) – Protective Whole blood COVID+ Symptoms+ (Severe) – Protective	MR with colocalization for GTEX v.7 eQTLs and pQTLs for whole-blood and lung (Baranova et al. 2021; Fricke-Galindo et al. 2022; Krishnamoorthy et al. 2023; Liu et al. 2021; Pairo-Castineira et al. 2021)
<i>IL10RB</i>	Cells Cultured fibroblasts (All) – Protective Esophagus Gastroesophageal junction (All) – Hazardous Skeletal muscle (Susceptible) – Hazardous Nerve tibial (All) – Hazardous	MR with colocalization for eQTLs and pQTLs, did not study excitatory neurons (Gaziano et al. 2021)
<i>RAB2A</i>	Adipose subcutaneous (Hospitalized) – Hazardous Artery aorta (Hospitalized) – Hazardous Artery tibial (Hospitalized) – Hazardous CD4 Memory Stim 40h (Hospitalized) – Hazardous CD4 Naïve Stim 16h (Hospitalized) – Hazardous CD4 Naïve Stim 40h (Hospitalized) – Hazardous Esophagus mucosa (Hospitalized) – Hazardous Esophagus muscularis (Hospitalized) – Hazardous Liver (Hospitalized) – Hazardous Lung (Hospitalized) – Hazardous Skin not sun exposed (Hospitalized) – Hazardous T naïve Stim 40h (Hospitalized) – Hazardous	MR and colocalization for pQTLs, but not for eQTLs (Pietzner et al. 2022)
<i>RALGDS</i>	T effect memory Stim 16h (Severe and Susceptible) – Protective	None
<i>WNT3</i>	Adrenal Gland (Hospitalized) – Hazardous Artery Aorta (Severe and Hospitalized) – Protective Colon Transverse (Severe) – Hazardous Nerve Tibial (Severe) – Hazardous Thyroid (Severe and Hospitalized) – Hazardous	Fine mapped (Wu et al. 2021)

488

489

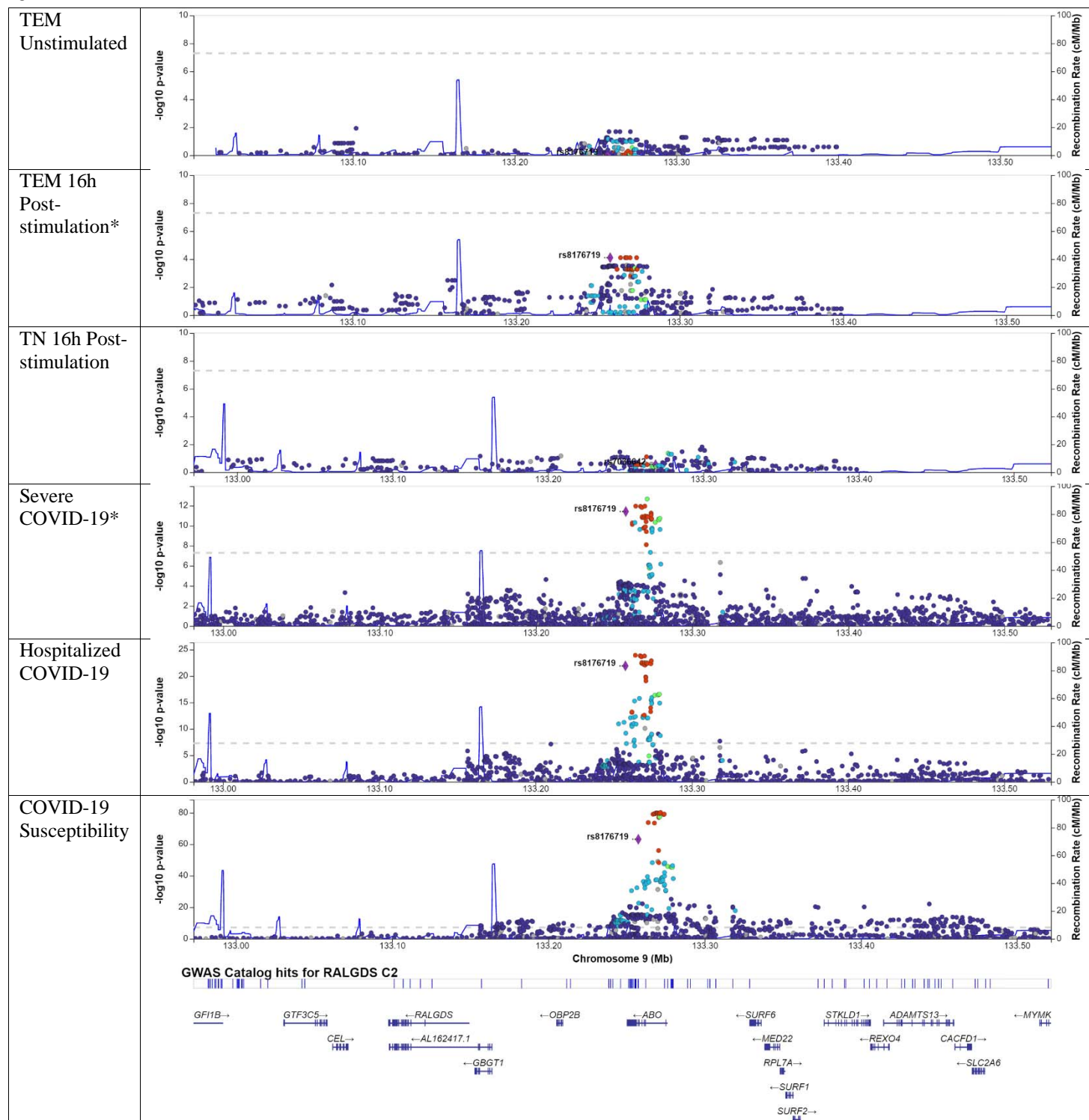
Colocalization of expression transcripts with COVID-19 outcomes is dependent on cell-type and cell-state.



490 **Fig. 3.**

Colocalization of expression transcripts with COVID-19 outcomes is dependent on cell-type and cell-state.

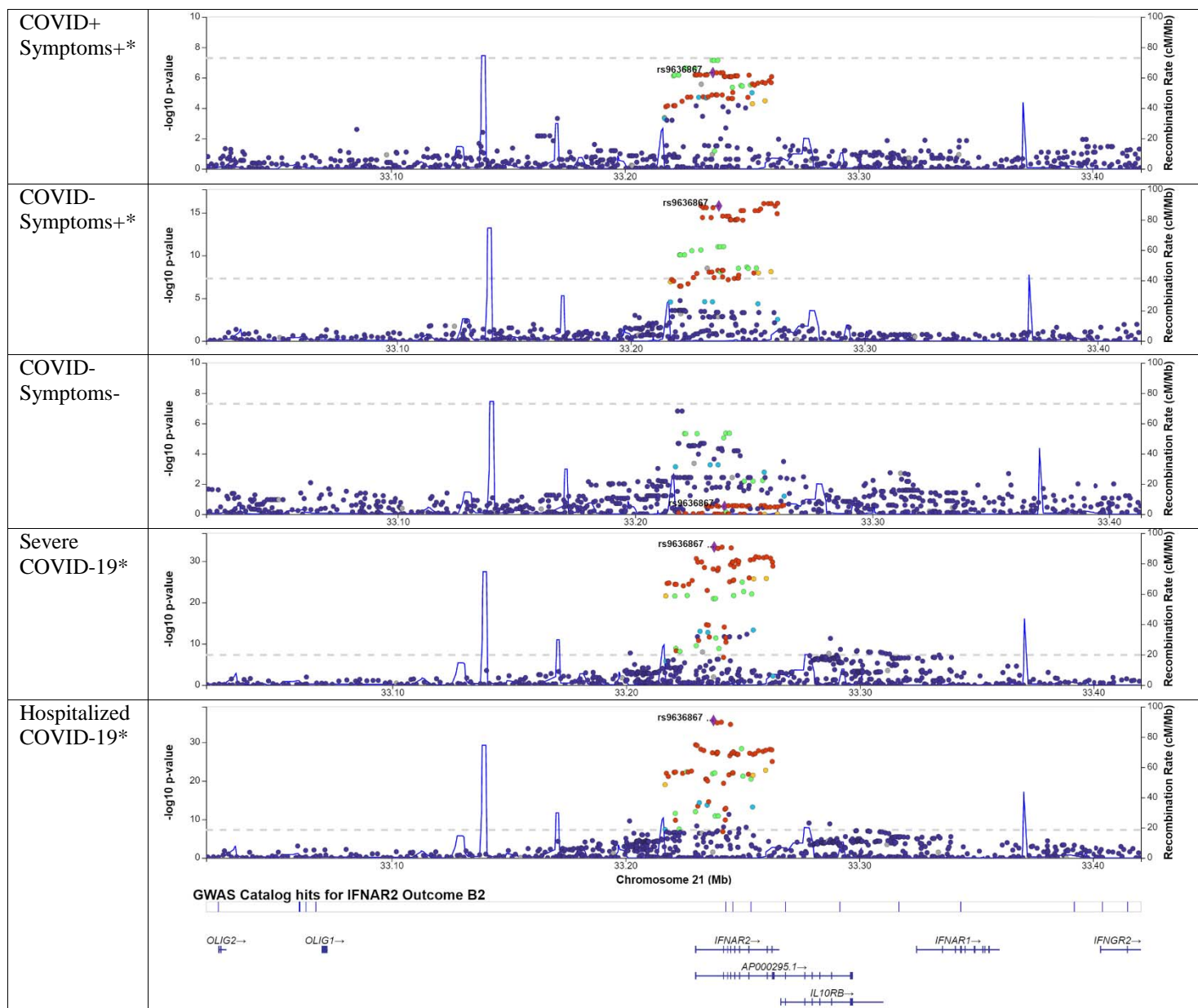
491



492 Fig 4.

493

Colocalization of expression transcripts with COVID-19 outcomes is dependent on cell-type and cell-state.

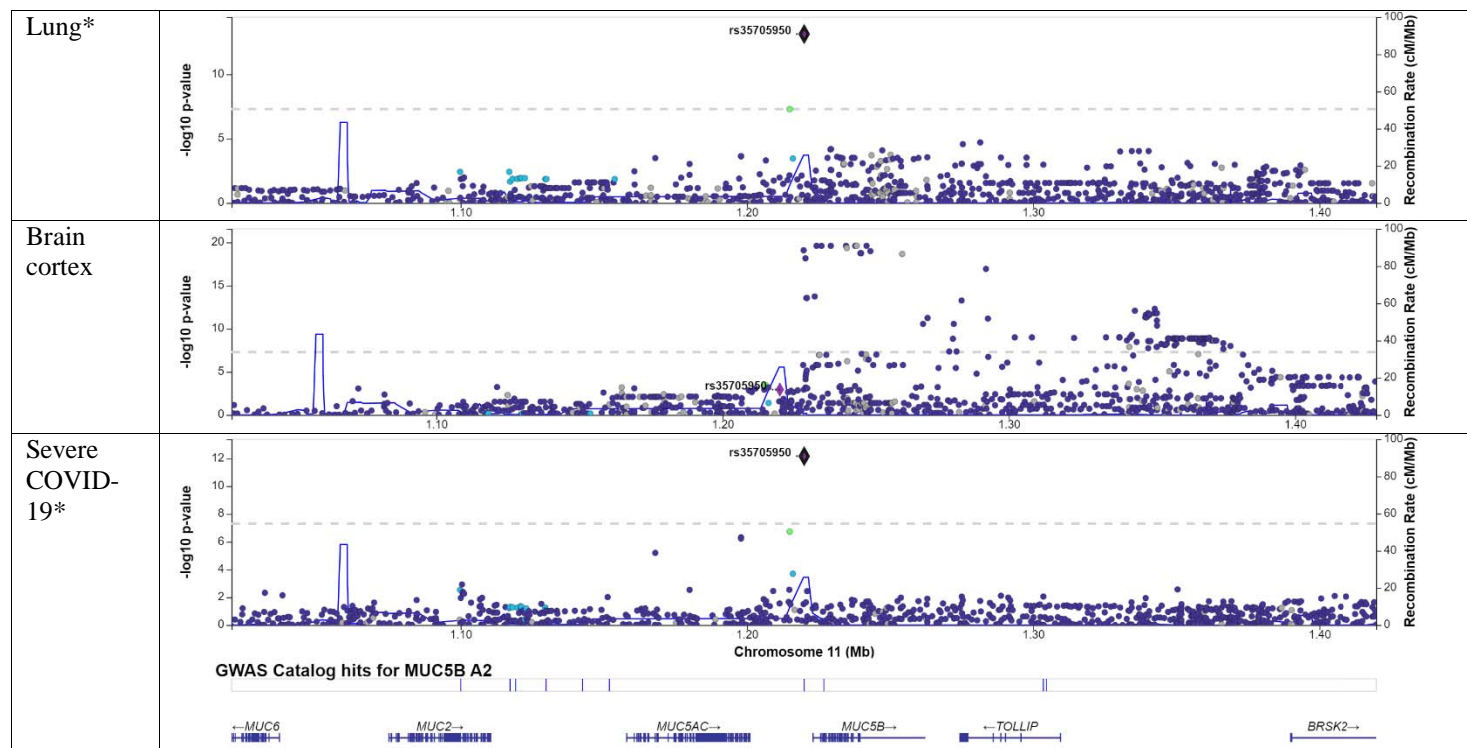


494 Fig 5.

495

Colocalization of expression transcripts with COVID-19 outcomes is dependent on cell-type and cell-state.

496

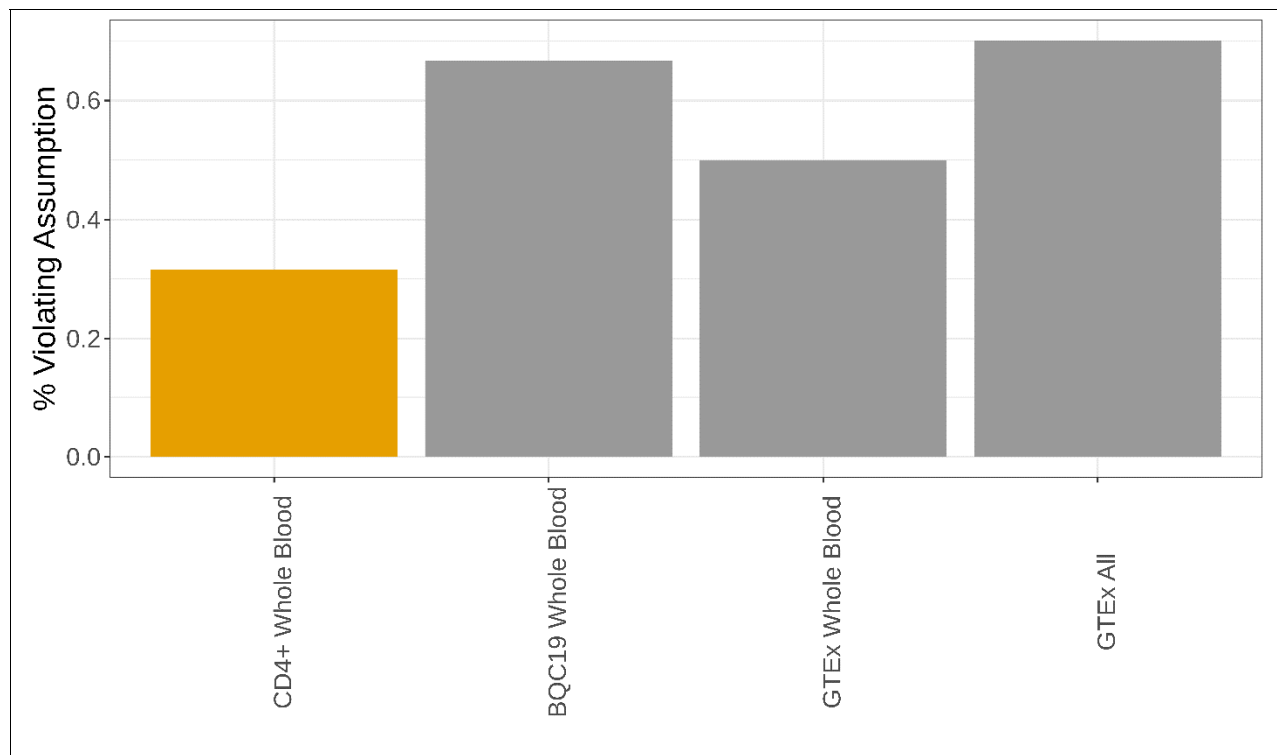


497 **Fig 6.**

498

499

Colocalization of expression transcripts with COVID-19 outcomes is dependent on cell-type and cell-state.



500 Fig. 7.
501

Colocalization of expression transcripts with COVID-19 outcomes is dependent on cell-type and cell-state.

502 **Supplemental Table 2.**

Gene	Colocalizing cell/tissue	Existing evidence
<i>ABO</i>	Adrenal gland (Severe) - Hazardous Artery tibial (Severe) – Hazardous Breast Mammary Tissue (Susceptible) – Hazardous Esophagus mucosa (Susceptible) - Hazardous Esophagus muscularis (Severe) - Hazardous Heart Left Ventricle (Severe and Hospitalized) – Hazardous Liver (Hospitalized) - Hazardous Muscle skeletal (Severe) - Hazardous Testis (All) - Hazardous Whole blood COVID+ Symptoms+ (Hospitalized) - Protective	MR in non-GTEX and GTEX eQTL and pQTL datasets for lung and whole blood (Baranova et al. 2022a; Hernandez Cordero et al. 2021)
<i>ADAM15</i>	T central memory Stim 5d (Susceptible) - Hazardous	MR with colocalization in GTEX did not observe colocalization (Wang et al. 2022)
<i>ATP11A</i>	Artery Aorta (Hospitalized) - Protective	Fine mapping (Kousathanas et al. 2022)
<i>ATP5PO</i>	Lung (Hospitalized) - Hazardous	Associational (Li et al. 2022)
<i>CNPY4</i>	Testis (Severe) - Protective	None
<i>COPS6</i>	Esophagus muscularis (Hospitalized) - Protective	None
<i>DLX3</i>	Lung (Hospitalized) - Hazardous	MR with colocalization in lung (Hernandez Cordero et al. 2021)
<i>DPP9</i>	Skin sun exposed lower leg (All) - Protective	Associational data in blood (Sharif-Zak et al. 2022)
<i>FAM55C</i>	Esophagus muscularis (Susceptible) - Hazardous	None
<i>GBA</i>	Nerve Tibial (Susceptible) - Hazardous Thyroid (Susceptible) - Hazardous	None
<i>GBAP1</i>	Breast Mammary Tissue (Susceptible) - Hazardous Cultured Fibroblasts (Susceptible) - Hazardous Colon Sigmoid (Susceptible) - Hazardous Lung (Susceptible) - Hazardous Nerve Tibial (Susceptible) - Hazardous Small Intestine Terminal Ileum (Susceptible) - Hazardous Stomach (Susceptible) - Hazardous	MR with colocalization for splice isoform (Nakanishi et al. 2022)
<i>HIP1</i>	Breast mammary tissue (Severe) - Hazardous Nerve tibial (Severe) - Hazardous Whole blood COVID- Symptoms- (Severe) - Protective	None
<i>IFNAR2</i>	Cells cultured fibroblasts (Severe and hospitalized) - Protective Esophagus mucosa (Susceptible) - Protective Whole blood COVID- Symptoms+ (Hospitalized) - Protective Whole blood COVID+ Symptoms+ (Severe) - Protective	MR with colocalization for GTEX v.7 eQTLs and pQTLs for whole-blood and lung (Baranova et al. 2021; Fricke-Galindo et al. 2022; Krishnamoorthy et al. 2023; Liu et al. 2021; Pairo-Castineira et al. 2021)
<i>IL10RB</i>	Cells Cultured fibroblasts (All) - Protective Esophagus Gastroesophageal junction (All) - Hazardous Skeletal muscle (Susceptible) - Hazardous Nerve tibial (All) - Hazardous	MR with colocalization for eQTLs and pQTLs, did not study excitatory neurons (Gaziano et al. 2021)
<i>KCNC3</i>	Tibial nerve (Severe and Hospitalized) – Hazardous	TWAS (Krishnamoorthy et al. 2023)
<i>LINC01301</i>	Breast mammary tissue (Hospitalized) – Protective Skin not sun exposed (Hospitalized) - Protective	TWAS (Pairo-Castineira et al. 2022)
<i>LINC02210</i>	Artery tibial (Severe) – Protective	MR with colocalization (Baranova et al. 2021)
<i>Lnc-INTS12-1</i>	Artery aorta (Severe) – Protective Thyroid (Severe) - Protective	None
<i>Lnc-THBS3-1</i>	Whole blood COVID- Symptoms- (Susceptible) - Protective	None
<i>MAMSTR</i>	Skin Sun Exposed (Hospitalized) – Protective	TWAS (Krishnamoorthy et al. 2023)

Colocalization of expression transcripts with COVID-19 outcomes is dependent on cell-type and cell-state.

<i>MUC1</i>	Nerve tibial (Severe and Hospitalized) - Hazardous	MR from GTEx v.7 against HGI release 6 found signal in blood (Krishnamoorthy et al. 2023)
<i>MUC5B</i>	Lung (Severe) - Protective	MR only for lung eQTLs GTEx release 7 (Liu et al. 2021)
<i>NAPSA</i>	CD4 Naïve Stim 40h (Susceptible) - Hazardous Esophageal mucosa (Severe and Hospitalized) - Protective T effector memory Stim 40h (All) - Hazardous T naïve (Hospitalized and Susceptible) - Hazardous Testis (Hospitalized) - Protective	MR with colocalization for eQTLs, did not study CD4+ T cells (Degenhardt et al. 2022)
<i>NAPSB</i>	Artery Coronary (Susceptible) - Protective	None
<i>NTN5</i>	Colon sigmoid (Hospitalized) - Protective Esophageal Gastroesophageal junction (Hospitalized) - Protective Esophagus mucosa (Severe and Hospitalized) - Protective	None
<i>OAS1</i>	Adipose subcutaneous (Hospitalized) - Protective Artery tibial (Susceptible) - Protective Esophagus mucosa (Hospitalized) - Hazardous Nerve tibial (Susceptible) - Hazardous Skin not sun exposed (Severe and Susceptible) - Protective Skin sun exposed (All) - Protective	MR with colocalization of eQTL and pQTL (Zhou et al. 2021)
<i>PDCL3P4</i>	Brain Caudate Basal Ganglia (Susceptible)	Colocalization (D'Antonio et al. 2021)
<i>RAB2A</i>	Adipose subcutaneous (Hospitalized) - Hazardous Artery aorta (Hospitalized) - Hazardous Artery tibial (Hospitalized) - Hazardous CD4 Memory Stim 40h (Hospitalized) - Hazardous CD4 Naïve Stim 16h (Hospitalized) - Hazardous CD4 Naïve Stim 40h (Hospitalized) - Hazardous Esophagus mucosa (Hospitalized) - Hazardous Esophagus muscularis (Hospitalized) - Hazardous Liver (Hospitalized) - Hazardous Lung (Hospitalized) - Hazardous Skin not sun exposed (Hospitalized) - Hazardous T naïve Stim 40h (Hospitalized) - Hazardous	MR and colocalization for pQTLs, but not for eQTLs (Pietzner et al. 2022)
<i>RALGDS</i>	T effect memory Stim 16h (Severe and Susceptible) - Protective	None
<i>RASIP1</i>	Esophageal Gastroesophageal junction (Severe and Hospitalized) - Protective Esophagus muscularis (Severe) - Protective	None
<i>SLC22A31</i>	Heart left ventricle (Severe and Hospitalized) – Hazardous	PheWAS linked (Covid-19 Host Genetics Initiative 2022)
<i>SLC6A20</i>	Breast mammary tissue (Severe and Hospitalized) – Hazardous	GWAS (Severe Covid et al. 2020)
<i>WNT3</i>	Adrenal Gland (Hospitalized) - Hazardous Artery Aorta (Severe and Hospitalized) - Protective Colon Transverse (Severe) - Hazardous Nerve Tibial (Severe) - Hazardous Thyroid (Severe and Hospitalized) - Hazardous	Fine mapped (Wu et al. 2021)

503

504 **References:**

505

506 Baranova A, Cao H, Chen J, Zhang F (2022a) Causal Association and Shared Genetics Between Asthma
507 and COVID-19. *Front Immunol* 13: 705379. doi: 10.3389/fimmu.2022.705379

508 Baranova A, Cao H, Teng S, Zhang F (2023) A phenome-wide investigation of risk factors for severe
509 COVID-19. *J Med Virol* 95: e28264. doi: 10.1002/jmv.28264

510 Baranova A, Cao H, Zhang F (2021) Unraveling Risk Genes of COVID-19 by Multi-Omics Integrative

511 Analyses. *Front Med (Lausanne)* 8: 738687. doi: 10.3389/fmed.2021.738687

Colocalization of expression transcripts with COVID-19 outcomes is dependent on cell-type and cell-state.

- 512 Baranova A, Cao H, Zhang F (2022b) Severe COVID-19 increases the risk of schizophrenia. *Psychiatry*
513 Res 317: 114809. doi: 10.1016/j.psychres.2022.114809
- 514 Connally NJ, Nazeen S, Lee D, Shi H, Stamatoyannopoulos J, Chun S, Cotsapas C, Cassa CA, Sunyaev
515 SR (2022) The missing link between genetic association and regulatory function. *Elife* 11. doi:
516 10.7554/eLife.74970
- 517 Covid-19 Host Genetics Initiative (2021) Mapping the human genetic architecture of COVID-19. *Nature*
518 600: 472-477. doi: 10.1038/s41586-021-03767-x
- 519 Covid-19 Host Genetics Initiative (2022) A first update on mapping the human genetic architecture of
520 COVID-19. *Nature* 608: E1-E10. doi: 10.1038/s41586-022-04826-7
- 521 D'Antonio M, Nguyen JP, Arthur TD, Matsui H, Initiative C-HG, D'Antonio-Chronowska A, Frazer KA
522 (2021) SARS-CoV-2 susceptibility and COVID-19 disease severity are associated with genetic
523 variants affecting gene expression in a variety of tissues. *Cell Rep* 37: 110020. doi:
524 10.1016/j.celrep.2021.110020
- 525 Davies NM, Holmes MV, Davey Smith G (2018) Reading Mendelian randomisation studies: a guide,
526 glossary, and checklist for clinicians. *BMJ* 362: k601. doi: 10.1136/bmj.k601
- 527 De Biasi S, Meschiari M, Gibellini L, Bellinazzi C, Borella R, Fianza L, Gozzi L, Iannone A, Lo Tartaro
528 D, Mattioli M, Paolini A, Menozzi M, Milic J, Franceschi G, Fantini R, Tonelli R, Sita M, Sarti
529 M, Trenti T, Brugioni L, Cicchetti L, Facchinetti F, Pietrangelo A, Clini E, Girardis M, Guaraldi
530 G, Mussini C, Cossarizza A (2020) Marked T cell activation, senescence, exhaustion and skewing
531 towards TH17 in patients with COVID-19 pneumonia. *Nat Commun* 11: 3434. doi:
532 10.1038/s41467-020-17292-4
- 533 Degenhardt F, Ellinghaus D, Juzenas S, Lerga-Jaso J, Wendorff M, Maya-Miles D, Uellendahl-Werth F,
534 ElAbd H, Ruhlemann MC, Arora J, Ozer O, Lenning OB, Myhre R, Vadla MS, Wacker EM,
535 Wienbrandt L, Blandino Ortiz A, de Salazar A, Garrido Chercoles A, Palom A, Ruiz A, Garcia-
536 Fernandez AE, Blanco-Grau A, Mantovani A, Zanella A, Holten AR, Mayer A, Bandera A,
537 Cherubini A, Protti A, Aghemo A, Gerussi A, Ramirez A, Braun A, Nebel A, Barreira A, Lleo A,
538 Teles A, Kildal AB, Biondi A, Caballero-Garralda A, Ganna A, Gori A, Gluck A, Lind A, Tanck
539 A, Hinney A, Carreras Nolla A, Fracanzani AL, Peschuck A, Cavallero A, Dyrhol-Riise AM,
540 Ruello A, Julia A, Muscatello A, Pesenti A, Voza A, Rando-Segura A, Solier A, Schmidt A,
541 Cortes B, Mateos B, Nafria-Jimenez B, Schaefer B, Jensen B, Bellinghausen C, Maj C, Ferrando
542 C, de la Horra C, Quereda C, Skurk C, Thibeault C, Scollo C, Herr C, Spinner CD, Gassner C,
543 Lange C, Hu C, Paccapelo C, Lehmann C, Angelini C, Cappadona C, Azuure C, Covicat study
544 group AS, Bianco C, Cea C, Sancho C, Hoff DAL, Galimberti D, Prati D, Haschka D, Jimenez D,
545 Pestana D, Toapanta D, Muniz-Diaz E, Azzolini E, Sandoval E, Binatti E, Scarpini E, Helbig ET,
546 et al. (2022) Detailed stratified GWAS analysis for severe COVID-19 in four European
547 populations. *Hum Mol Genet* 31: 3945-3966. doi: 10.1093/hmg/ddac158
- 548 Forgetta V, Jiang L, Vulpescu NA, Hogan MS, Chen S, Morris JA, Grinek S, Benner C, Jang DK, Hoang
549 Q, Burt N, Flannick JA, McCarthy MI, Fauman E, Greenwood CMT, Maurano MT, Richards JB
550 (2022) An effector index to predict target genes at GWAS loci. *Hum Genet*. doi: 10.1007/s00439-
551 022-02434-z
- 552 Fricke-Galindo I, Martinez-Morales A, Chavez-Galan L, Ocana-Guzman R, Buendia-Roldan I, Perez-
553 Rubio G, Hernandez-Zenteno RJ, Veronica-Aguilar A, Alarcon-Dionet A, Aguilar-Duran H,
554 Gutierrez-Perez IA, Zaragoza-Garcia O, Alanis-Ponce J, Camarena A, Bautista-Becerril B, Nava-
555 Quiroz KJ, Mejia M, Guzman-Guzman IP, Falfan-Valencia R (2022) IFNAR2 relevance in the
556 clinical outcome of individuals with severe COVID-19. *Front Immunol* 13: 949413. doi:
557 10.3389/fimmu.2022.949413
- 558 Gaziano L, Giambartolomei C, Pereira AC, Gaulton A, Posner DC, Swanson SA, Ho YL, Iyengar SK,
559 Kosik NM, Vujkovic M, Gagnon DR, Bento AP, Barrio-Hernandez I, Ronnblom L, Hagberg N,
560 Lundtoft C, Langenberg C, Pietzner M, Valentine D, Gustincich S, Tartaglia GG, Allara E,
561 Surendran P, Burgess S, Zhao JH, Peters JE, Prins BP, Angelantonio ED, Devineni P, Shi Y,
562 Lynch KE, DuVall SL, Garcon H, Thomann LO, Zhou JJ, Gorman BR, Huffman JE, O'Donnell

Colocalization of expression transcripts with COVID-19 outcomes is dependent on cell-type and cell-state.

- 563 CJ, Tsao PS, Beckham JC, Pyarajan S, Muralidhar S, Huang GD, Ramoni R, Beltrao P, Danesh J,
564 Hung AM, Chang KM, Sun YV, Joseph J, Leach AR, Edwards TL, Cho K, Gaziano JM,
565 Butterworth AS, Casas JP, Initiative VAMVPC-S (2021) Actionable druggable genome-wide
566 Mendelian randomization identifies repurposing opportunities for COVID-19. *Nat Med* 27: 668-
567 676. doi: 10.1038/s41591-021-01310-z
- 568 GTEx Consortium (2020) The GTEx Consortium atlas of genetic regulatory effects across human tissues.
569 *Science* 369: 1318-1330. doi: 10.1126/science.aaz1776
- 570 Hernandez Cordero AI, Li X, Milne S, Yang CX, Bosse Y, Joubert P, Timens W, van den Berge M,
571 Nickle D, Hao K, Sin DD (2021) Multi-omics highlights ABO plasma protein as a causal risk
572 factor for COVID-19. *Hum Genet* 140: 969-979. doi: 10.1007/s00439-021-02264-5
- 573 Kanai M, Elzur R, Zhou W, Global Biobank Meta-analysis I, Daly MJ, Finucane HK (2022) Meta-
574 analysis fine-mapping is often miscalibrated at single-variant resolution. *Cell Genom* 2. doi:
575 10.1016/j.xgen.2022.100210
- 576 Kanai M, Ulirsch JC, Karjalainen J, Kurki M, Karczewski KJ, Fauman E, Wang QS, Jacobs H, Aguet F,
577 Ardlie KG, Kerimov N, Alasoo K, Benner C, Ishigaki K, Sakaue S, Reilly S, Kamatani Y,
578 Matsuda K, Palotie A, Neale BM, Tewhey R, Sabeti PC, Okada Y, Daly MJ, Finucane HK (2021)
579 Insights from complex trait fine-mapping across diverse populations. medRxiv. doi:
580 10.1101/2021.09.03.21262975
- 581 Kousathanas A, Pairo-Castineira E, Rawlik K, Stuckey A, Odhams CA, Walker S, Russell CD,
582 Malinauskas T, Wu Y, Millar J, Shen X, Elliott KS, Griffiths F, Oosthuyzen W, Morrice K,
583 Keating S, Wang B, Rhodes D, Klaric L, Zechner M, Parkinson N, Siddiq A, Goddard P,
584 Donovan S, Maslove D, Nichol A, Semple MG, Zainy T, Maleady-Crowe F, Todd L, Salehi S,
585 Knight J, Elgar G, Chan G, Arumugam P, Patch C, Rendon A, Bentley D, Kingsley C, Kosmicki
586 JA, Horowitz JE, Baras A, Abecasis GR, Ferreira MAR, Justice A, Mirshahi T, Oetjens M, Rader
587 DJ, Ritchie MD, Verma A, Fowler TA, Shankar-Hari M, Summers C, Hinds C, Horby P, Ling L,
588 McAuley D, Montgomery H, Openshaw PJM, Elliott P, Walsh T, Tenesa A, Gen OI, and Me i,
589 Initiative C-HG, Fawkes A, Murphy L, Rowan K, Ponting CP, Vitart V, Wilson JF, Yang J,
590 Bretherick AD, Scott RH, Hendry SC, Moutsianas L, Law A, Caulfield MJ, Baillie JK (2022)
591 Whole-genome sequencing reveals host factors underlying critical COVID-19. *Nature* 607: 97-
592 103. doi: 10.1038/s41586-022-04576-6
- 593 Krishnamoorthy S, Li GH, Cheung CL (2023) Transcriptome-wide summary data-based Mendelian
594 randomization analysis reveals 38 novel genes associated with severe COVID-19. *J Med Virol*
595 95: e28162. doi: 10.1002/jmv.28162
- 596 Kundu R, Narean JS, Wang L, Fenn J, Pillay T, Fernandez ND, Conibear E, Koycheva A, Davies M,
597 Tolosa-Wright M, Hakki S, Varro R, McDermott E, Hammett S, Cutajar J, Thwaites RS, Parker
598 E, Rosadas C, McClure M, Tedder R, Taylor GP, Dunning J, Lalvani A (2022) Cross-reactive
599 memory T cells associate with protection against SARS-CoV-2 infection in COVID-19 contacts.
600 *Nat Commun* 13: 80. doi: 10.1038/s41467-021-27674-x
- 601 Lamontagne M, Berube JC, Obeidat M, Cho MH, Hobbs BD, Sakornsakolpat P, de Jong K, Boezen HM,
602 International CGC, Nickle D, Hao K, Timens W, van den Berge M, Joubert P, Laviolette M, Sin
603 DD, Pare PD, Bosse Y (2018) Leveraging lung tissue transcriptome to uncover candidate causal
604 genes in COPD genetic associations. *Hum Mol Genet* 27: 1819-1829. doi: 10.1093/hmg/ddy091
- 605 Li P, Ke Y, Shen W, Shi S, Wang Y, Lin K, Guo X, Wang C, Zhang Y, Zhao Z (2022) Targeted
606 screening of genetic associations with COVID-19 susceptibility and severity. *Front Genet* 13:
607 1073880. doi: 10.3389/fgene.2022.1073880
- 608 Liu D, Yang J, Feng B, Lu W, Zhao C, Li L (2021) Mendelian randomization analysis identified genes
609 pleiotropically associated with the risk and prognosis of COVID-19. *J Infect* 82: 126-132. doi:
610 10.1016/j.jinf.2020.11.031
- 611 Mathew D, Giles JR, Baxter AE, Oldridge DA, Greenplate AR, Wu JE, Alanio C, Kuri-Cervantes L,
612 Pampena MB, D'Andrea K, Manne S, Chen Z, Huang YJ, Reilly JP, Weisman AR, Ittner CAG,
613 Kuthuru O, Dougherty J, Nzingha K, Han N, Kim J, Pattekar A, Goodwin EC, Anderson EM,

Colocalization of expression transcripts with COVID-19 outcomes is dependent on cell-type and cell-state.

- 614 Weirick ME, Gouma S, Arevalo CP, Bolton MJ, Chen F, Lacey SF, Ramage H, Cherry S,
615 Hensley SE, Apostolidis SA, Huang AC, Vella LA, Unit UCP, Betts MR, Meyer NJ, Wherry EJ
616 (2020) Deep immune profiling of COVID-19 patients reveals distinct immunotypes with
617 therapeutic implications. *Science* 369. doi: 10.1126/science.abc8511
- 618 Mehandru S, Merad M (2022) Pathological sequelae of long-haul COVID. *Nat Immunol* 23: 194-202.
619 doi: 10.1038/s41590-021-01104-y
- 620 Merad M, Blish CA, Sallusto F, Iwasaki A (2022) The immunology and immunopathology of COVID-
621 19. *Science* 375: 1122-1127. doi: 10.1126/science.abm8108
- 622 Nakanishi T, Farjoun Y, Willett J, Allen RJ, Guillen-Guio B, Zhou S, Richards JB (2022) Alternative
623 splicing in the lung influences COVID-19 severity and respiratory diseases. medRxiv. doi:
624 10.1101/2022.10.18.22281202
- 625 Ong EZ, Chan YFZ, Leong WY, Lee NMY, Kalimuddin S, Haja Mohideen SM, Chan KS, Tan AT,
626 Bertoletti A, Ooi EE, Low JGH (2020) A Dynamic Immune Response Shapes COVID-19
627 Progression. *Cell Host Microbe* 27: 879-882 e2. doi: 10.1016/j.chom.2020.03.021
- 628 Pairo-Castineira E, Clohisey S, Klaric L, Bretherick AD, Rawlik K, Pasko D, Walker S, Parkinson N,
629 Fourman MH, Russell CD, Furniss J, Richmond A, Gountouna E, Wrobel N, Harrison D, Wang
630 B, Wu Y, Meynert A, Griffiths F, Oosthuyzen W, Kousathanas A, Moutsianas L, Yang Z, Zhai R,
631 Zheng C, Grimes G, Beale R, Millar J, Shih B, Keating S, Zechner M, Haley C, Porteous DJ,
632 Hayward C, Yang J, Knight J, Summers C, Shankar-Hari M, Klenerman P, Turtle L, Ho A,
633 Moore SC, Hinds C, Horby P, Nichol A, Maslove D, Ling L, McAuley D, Montgomery H, Walsh
634 T, Pereira AC, Renieri A, Gen OI, Investigators IC, Initiative C-HG, and Me I, Investigators B,
635 Gen CI, Shen X, Ponting CP, Fawkes A, Tenesa A, Caulfield M, Scott R, Rowan K, Murphy L,
636 Openshaw PJM, Semple MG, Law A, Vitart V, Wilson JF, Baillie JK (2021) Genetic
637 mechanisms of critical illness in COVID-19. *Nature* 591: 92-98. doi: 10.1038/s41586-020-03065-
638 y
- 639 Pairo-Castineira E, Rawlik K, Klaric L, Kousathanas A, Richmond A, Millar J, Russell CD, Malinauskas
640 T, Thwaites R, Stuckey A, Odhams CA, Walker S, Griffiths F, Oosthuyzen W, Morrice K,
641 Keating S, Nichol A, Semple MG, Knight J, Shankar-Hari M, Summers C, Hinds C, Horby P,
642 Ling L, McAuley D, Montgomery H, Openshaw PJM, Walsh T, Tenesa A, Scott RH, Caulfield
643 MJ, Moutsianas L, Ponting CP, Wilson JF, Vitart V, Pereira AC, Luchessi A, Parra E, Cruz-
644 Guerrero R, Carracedo A, Fawkes A, Murphy L, Rowan K, Law A, Hendry SC, Baillie JK (2022)
645 GWAS and meta-analysis identifies multiple new genetic mechanisms underlying severe Covid-
646 19. medRxiv. doi: 10.1101/2022.03.07.22271833
- 647 Pietzner M, Chua RL, Wheeler E, Jechow K, Willett JDS, Radbruch H, Trump S, Heidecker B, Zeberg H,
648 Heppner FL, Eils R, Mall MA, Richards JB, Sander LE, Lehmann I, Lukassen S, Wareham NJ,
649 Conrad C, Langenberg C (2022) ELF5 is a potential respiratory epithelial cell-specific risk gene
650 for severe COVID-19. *Nat Commun* 13: 4484. doi: 10.1038/s41467-022-31999-6
- 651 Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, Boehnke M, Abecasis GR, Willer
652 CJ (2010) LocusZoom: regional visualization of genome-wide association scan results.
653 *Bioinformatics* 26: 2336-7. doi: 10.1093/bioinformatics/btq419
- 654 Randolph HE, Fiege JK, Thielen BK, Mickelson CK, Shiratori M, Barroso-Batista J, Langlois RA,
655 Barreiro LB (2021) Genetic ancestry effects on the response to viral infection are pervasive but
656 cell type specific. *Science* 374: 1127-1133. doi: 10.1126/science.abg0928
- 657 Rocheleau G, Forrest IS, Duffy A, Bafna S, Dobbyn A, Verbanck M, Won HH, Jordan DM, Do R (2022)
658 A tissue-level phenome-wide network map of colocalized genes and phenotypes in the UK
659 Biobank. *Commun Biol* 5: 849. doi: 10.1038/s42003-022-03820-z
- 660 Schmiedel BJ, Rocha J, Gonzalez-Colin C, Bhattacharyya S, Madrigal A, Ottensmeier CH, Ay F,
661 Chandra V, Vijayanand P (2021) COVID-19 genetic risk variants are associated with expression
662 of multiple genes in diverse immune cell types. *Nat Commun* 12: 6760. doi: 10.1038/s41467-
663 021-26888-3

Colocalization of expression transcripts with COVID-19 outcomes is dependent on cell-type and cell-state.

- 664 Severe Covid GG, Ellinghaus D, Degenhardt F, Bujanda L, Buti M, Albillos A, Invernizzi P, Fernandez J,
665 Prati D, Baselli G, Asselta R, Grimsrud MM, Milani C, Aziz F, Kassens J, May S, Wendorff M,
666 Wienbrandt L, Uellendahl-Werth F, Zheng T, Yi X, de Pablo R, Chercoles AG, Palom A, Garcia-
667 Fernandez AE, Rodriguez-Frias F, Zanella A, Bandera A, Protti A, Aghemo A, Lleo A, Biondi A,
668 Caballero-Garralda A, Gori A, Tanck A, Carreras Nolla A, Latiano A, Fracanzani AL, Peschuck
669 A, Julia A, Pesenti A, Voza A, Jimenez D, Mateos B, Nafria Jimenez B, Quereda C, Paccapelo C,
670 Gassner C, Angelini C, Cea C, Solier A, Pestana D, Muniz-Diaz E, Sandoval E, Paraboschi EM,
671 Navas E, Garcia Sanchez F, Ceriotti F, Martinelli-Boneschi F, Peyvandi F, Blasi F, Tellez L,
672 Blanco-Grau A, Hemmrich-Stanisak G, Grasselli G, Costantino G, Cardamone G, Foti G, Aneli
673 S, Kurihara H, ElAbd H, My I, Galvan-Femenia I, Martin J, Erdmann J, Ferrusquia-Acosta J,
674 Garcia-Etxebarria K, Izquierdo-Sanchez L, Bettini LR, Sumoy L, Terranova L, Moreira L,
675 Santoro L, Scudeller L, Mesonero F, Roade L, Ruhlemann MC, Schaefer M, Carrabba M,
676 Riveiro-Barciela M, Figuera Basso ME, Valsecchi MG, Hernandez-Tejero M, Acosta-Herrera M,
677 D'Angio M, Baldini M, Cazzaniga M, Schulzky M, Cecconi M, Wittig M, et al. (2020)
678 Genomewide Association Study of Severe Covid-19 with Respiratory Failure. *N Engl J Med* 383:
679 1522-1534. doi: 10.1056/NEJMoa2020283
- 680 Sharif-Zak M, Abbasi-Jorjandi M, Asadikaram G, Ghoreishi ZA, Rezazadeh-Jabalbarzi M, Afsharipur A,
681 Rashidinejad H, Khajepour F, Jafarzadeh A, Arefinia N, Kheyrkhah A, Abolhassani M (2022)
682 CCR2 and DPP9 expression in the peripheral blood of COVID-19 patients: Influences of the
683 disease severity and gender. *Immunobiology* 227: 152184. doi: 10.1016/j.imbio.2022.152184
- 684 Skrivankova VW, Richmond RC, Woolf BAR, Yarmolinsky J, Davies NM, Swanson SA, VanderWeele
685 TJ, Higgins JPT, Timpson NJ, Dimou N, Langenberg C, Golub RM, Loder EW, Gallo V,
686 Tybjaerg-Hansen A, Davey Smith G, Egger M, Richards JB (2021) Strengthening the Reporting
687 of Observational Studies in Epidemiology Using Mendelian Randomization: The STROBE-MR
688 Statement. *JAMA* 326: 1614-1621. doi: 10.1001/jama.2021.18236
- 689 Smith GD, Ebrahim S (2003) 'Mendelian randomization': can genetic epidemiology contribute to
690 understanding environmental determinants of disease? *Int J Epidemiol* 32: 1-22. doi:
691 10.1093/ije/dyg070
- 692 Soskic B, Cano-Gamez E, Smyth DJ, Ambridge K, Ke Z, Matte JC, Bossini-Castillo L, Kaplanis J,
693 Ramirez-Navarro L, Lorenc A, Nakic N, Esparza-Gordillo J, Rowan W, Wille D, Tough DF,
694 Bronson PG, Trynka G (2022) Immune disease risk variants regulate gene expression dynamics
695 during CD4(+) T cell activation. *Nat Genet*. doi: 10.1038/s41588-022-01066-3
- 696 Tan LY, Komarasamy TV, Rmt Balasubramaniam V (2021) Hyperinflammatory Immune Response and
697 COVID-19: A Double Edged Sword. *Front Immunol* 12: 742941. doi:
698 10.3389/fimmu.2021.742941
- 699 Wang Y, Guga S, Wu K, Khaw Z, Tzoumkas K, Tombleson P, Comeau ME, Langefeld CD,
700 Cunninghame Graham DS, Morris DL, Vyse TJ (2022) COVID-19 and systemic lupus
701 erythematosus genetics: A balance between autoimmune disease risk and protection against
702 infection. *PLoS Genet* 18: e1010253. doi: 10.1371/journal.pgen.1010253
- 703 Wu L, Zhu J, Liu D, Sun Y, Wu C (2021) An integrative multiomics analysis identifies putative causal
704 genes for COVID-19 severity. *Genet Med* 23: 2076-2086. doi: 10.1038/s41436-021-01243-5
- 705 Yoshiji S, Butler-Laporte G, Lu T, Willett JDS, Su CY, Nakanishi T, Morrison DR, Chen Y, Liang K,
706 Hultstrom M, Ilboudo Y, Afrasiabi Z, Lan S, Duggan N, DeLuca C, Vaezi M, Tselios C, Xue X,
707 Bouab M, Shi F, Laurent L, Munter HM, Afilalo M, Afilalo J, Mooser V, Timpson NJ, Zeberg H,
708 Zhou S, Forgetta V, Farjoun Y, Richards JB (2023) Proteome-wide Mendelian randomization
709 implicates nephronectin as an actionable mediator of the effect of obesity on COVID-19 severity.
710 *Nat Metab* 5: 248-264. doi: 10.1038/s42255-023-00742-w
- 711 Zheng J, Haberland V, Baird D, Walker V, Haycock PC, Hurler MR, Gutteridge A, Erola P, Liu Y, Luo S,
712 Robinson J, Richardson TG, Staley JR, Elsworth B, Burgess S, Sun BB, Danesh J, Runz H,
713 Maranville JC, Martin HM, Yarmolinsky J, Laurin C, Holmes MV, Liu JZ, Estrada K, Santos R,
714 McCarthy L, Waterworth D, Nelson MR, Smith GD, Butterworth AS, Hemani G, Scott RA,

Colocalization of expression transcripts with COVID-19 outcomes is dependent on cell-type and cell-state.

715 Gaunt TR (2020) Phenome-wide Mendelian randomization mapping the influence of the plasma
716 proteome on complex diseases. Nat Genet 52: 1122-1131. doi: 10.1038/s41588-020-0682-6
717 Zhou S, Butler-Laporte G, Nakanishi T, Morrison DR, Afilalo J, Afilalo M, Laurent L, Pietzner M,
718 Kerrison N, Zhao K, Brunet-Ratnasingham E, Henry D, Kimchi N, Afrasiabi Z, Rezk N, Bouab
719 M, Petitjean L, Guzman C, Xue X, Tselios C, Vulesevic B, Adeleye O, Abdullah T, Almamlouk
720 N, Chen Y, Chasse M, Durand M, Paterson C, Normark J, Frithiof R, Lipcsey M, Hultstrom M,
721 Greenwood CMT, Zeberg H, Langenberg C, Thysell E, Pollak M, Mooser V, Forgetta V,
722 Kaufmann DE, Richards JB (2021) A Neanderthal OAS1 isoform protects individuals of
723 European ancestry against COVID-19 susceptibility and severity. Nat Med 27: 659-667. doi:
724 10.1038/s41591-021-01281-1
725