

Genetic and Phenotypic Evidence for the Causal Relationship Between Aging and COVID-19

Kejun Ying^{1,2,3†}, Ranran Zhai^{1†}, Timothy V. Pyrkov⁴, Marco Mariotti², Peter O. Fedichev^{4,5}, Xia Shen^{*1,6,7} & Vadim N. Gladyshev^{*2}

¹ Biostatistics Group, State Key Laboratory of Biocontrol, School of Life Sciences, Sun Yat-sen University, Guangzhou, China

² Division of Genetics, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA 02115, USA

³ T. H. Chan School of Public Health, Harvard University, Boston, USA

⁴ Gero LLC PTE, Singapore

⁵ Moscow Institute of Physics and Technology, Dolgoprudny, Moscow Region, Russia

⁶ Usher Institute of Population Health Sciences and Informatics, University of Edinburgh, Edinburgh, United Kingdom

⁷ Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden

* Correspondence to: xia.shen@ed.ac.uk and vgladyshev@rics.bwh.harvard.edu

† These authors contributed equally to this work

Abstract

Epidemiological studies revealed that the elderly and those with co-morbidities are most susceptible to COVID-19. To understand how genetics affects the risk of COVID-19, we conducted a multi-instrument Mendelian randomization analysis and found that the genetic variation that supports a longer life is significantly associated with the lower risk of COVID-19 infection. The odds ratio is 0.31 (95% CI: 0.18 to 0.52; $P = 9.7 \times 10^{-6}$) per additional 10 years of life, and 0.53 (95% CI: 0.43 to 0.65; $P = 2.3 \times 10^{-9}$) per unit higher log odds of surviving to the 90th percentile. On the other hand, there was no association between COVID-19 susceptibility and healthspan (the lifespan free of the top seven age-related morbidities). We further applied aging clock models and detected an association between biological age acceleration and future incidence and severity of COVID-19 infection for all subjects as well as for individuals free of chronic disease. Biological age acceleration was also significantly associated with the risk of death in COVID-19 patients. Finally, a bivariate genomic scan for age-related COVID-19 infection identified a key contribution of the Notch signaling pathway. Our analysis suggests that Notch2 expression is associated with a higher risk of COVID-19 infection, providing a druggable target. More generally, interventions that reduce biological age have the opportunity to reduce the risk of COVID-19.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

37 **Introduction**

38 The coronavirus disease 2019 (COVID-19), caused by severe acute respiratory coronavirus 2
39 (SARS-CoV-2), first emerged in late 2019 and has led to an unprecedented global health crisis¹.
40 Notably, the aging population is particularly at risk of COVID-19², e.g. in Italy, 88% of the
41 individuals tested positive for COVID-19 were 40 years or older³. A recent report based on
42 epidemiological data from multiple countries showed that 69% of infections in people over 70
43 are symptomatic, whereas this number drops to 21% for 10-19-year-olds⁴. Unsurprisingly,
44 elderly people are also more likely to die from COVID-19, and the case fatality rate for
45 COVID-19 grows exponentially with age³. As observational evidence implies a strong link
46 between COVID-19 and age, COVID-19 can be considered a disease of aging³, and multiple
47 clinical trials using potential lifespan-extending drugs (e.g., metformin, rapamycin, and
48 senolytics) to protect the elderly from COVID-19 have been proposed⁵⁻⁷. Although
49 observational data on metformin seems promising^{8,9}, it is unclear if other lifespan-extending
50 drugs should be prioritized in clinical trials, since the evidence of any causal link between
51 lifespan and COVID-19 susceptibility is still missing.

52 Mendelian Randomization (MR) is a genetic instrumental variable approach that
53 assesses the causal effect of exposure of interest on an outcome, by ascertaining on genetic
54 variants, e.g., single nucleotide polymorphisms (SNPs), strongly associated with the exposure
55 phenotype. Since the alleles of the genetic variants are naturally randomly allocated at
56 conception, when the genetic effects on the outcome are only mediated through the exposure,
57 the causal effect inferred by MR is, in analogy to randomized clinical trials (RCTs), free of any
58 environmental confounders and reverse causation. Although RCTs are considered a gold
59 standard for establishing causal relationships, MR can provide valuable insights into causality
60 when it is not feasible to perform an RCT or before an RCT is performed¹⁰.

61 In this study, we performed a multi-SNP MR analysis to elucidate whether and how
62 aging is associated with COVID-19. We considered four lifespan-related traits (parental
63 lifespan, healthspan, longevity, and healthy aging, the combination of these three traits), four
64 measures of epigenetic age acceleration, and four leading genetic risk factors associated with
65 earlier death in humans (Alzheimer Disease (AD), cardiovascular disease (CVD), type 2
66 diabetes (T2D), and smoking as exposures and evaluated their causal effects on COVID-19
67 infection and related phenotypes. To support the argument, we also estimated the biological
68 age acceleration (BAA) in COVID-19 patients from UK Biobank (UKBB) cohort and observed
69 a significant association between the phenotypic indicators of aging progress (aging clocks)

70 and the risk and case fatality rate of COVID-19. To provide functional insight into how aging
71 contributes to a higher risk of COVID-19, we further conducted a bivariate genomic scan to
72 highlight the loci contribute to both aging and COVID-19 risk. The pathway enrichment of
73 these loci points to the Notch signaling pathway and specifically to Notch2, whose expression
74 supports an elevated risk of COVID-19 infection.

75 **Methods**

76 **GWAS data for lifespan-related traits and diseases**

77 We studied four lifespan-related traits with publicly available GWAS summary statistics:

78 The parental Lifespan GWAS includes unrelated, European-ancestry subjects (a total of
79 512,047 mother and 500,193 father lifespans), 60% of which were complete. The statistics for
80 every cohort was calculated by fitting Cox survival models to mother and father survival
81 respectively, taking account of 10 principal components, study-specific covariates, and
82 individual sex. Note that in the GWAS setting, parental lifespan is the same phenotype as
83 general lifespan of individuals. This is due to the mathematical property that the genetic effect
84 on a parental phenotype is simply half that on the individual's phenotype itself. Therefore, the
85 parental lifespan GWAS is a general lifespan GWAS with weaker power. But thanks to the
86 large sample size especially in UK Biobank, such a GWAS is powerful enough to uncover
87 some genetic architecture¹¹.

88 The longevity GWAS includes unrelated, European-ancestry subjects who had a lifespan
89 above the 90th survival percentile (N = 11,262) or whose age at the last follow-up visit (or age
90 at death) was before the 60th percentile age (N = 25,483). The statistics for each cohort were
91 calculated using logistic regression and then combined using a fixed-effect meta-analysis¹².

92 The healthspan GWAS contains 300,477 unrelated, British-ancestry individuals from
93 UKBB. The statistics were calculated by fitting Cox-Gompertz survival models, and the events
94 are defined as the first incidence of dementia, congestive heart failure, diabetes, chronic
95 obstructive pulmonary disease, stroke, cancer, myocardial infarction, or demise¹³.

96 The summary association statistics of healthy aging is from the meta-analysis of
97 healthspan, lifespan, and longevity summary statistics using MANOVA¹⁴, while accounting
98 for correlations between studies due to sample overlap and correlation amongst the traits.
99 Summary association statistics were calculated for the 7,320,282 SNPs shared between studies.
100 These statistics represent the significance of each SNP affecting one or more of the traits,
101 giving a P-value against the null hypothesis that effect sizes are zero in all studies^{14,15}.

102 We investigated four additional traits genetically correlated with lifespan, using the
103 published case-control studies: Alzheimer’s disease¹⁶, coronary artery disease¹⁷, type 2
104 diabetes¹⁸, and smoking¹⁹ (Table S1).

105 We also included GWAS for age acceleration measured by four epigenetic clocks,
106 including Hannum age, Horvath age, PhenoAge, and GrimAge²⁰. The epigenetic age was
107 measured on 34,449 healthy individuals of European ancestry.

108 GWAS data for 22 common diseases were from a community-based study, Genetic
109 Epidemiology Research on Adult Health and Aging (GERA)²¹, and was analyzed in the
110 original GSMR study²². There were 60,586 individuals of European ancestry in the GERA
111 data. There is an additional trait “disease count”, which represents the number of diseases
112 affecting each individual and the summary statistics of these diseases were adjusted with age,
113 gender, and the first 20 PCs.

114 We used 1000 Genomes Phase 3 reference (released in 2014 October) to map variants in
115 the GWAS results to rsIDs by chromosome, position, and alleles. Only the autosomal SNPs
116 available in the 1000 Genomes reference panel were used, and the 1000 Genomes European
117 ancestry reference was used to estimate the linkage disequilibrium (LD) among these SNPs.
118 Duplicated rsIDs in the data were removed prior to the analysis.

119 **COVID-19-related traits**

120 To extensively evaluate the genetic effects on COVID-19 risk, we used GWAS summary
121 statistics data from 12 COVID-19-related traits (Table S1). The GWAS results for SARS-
122 COV-2 infection are from the National Institute of Health, Genome-Wide Repository of
123 Associations Between SNPs and Phenotypes (NIH-GRASP), released in August 2020, which
124 includes 1,503 positive cases and 11,409 negative or 457,747 UK Biobank controls with
125 European ancestry. The GWAS summary statistics for severe COVID-19 with respiratory
126 failure is from a genome-wide association study performed in 1,610 cases and 2,205 controls
127 in Italy and Spain²³. The rest of the 9 traits are from the COVID-19 Host Genetics Initiative
128 (HGI), with the sample size varies from 1,332 to 1,079,768²⁴. Among them, 3 are were from
129 HGI release 2 (May 2020), including COVID-19 Hospitalization (versus non-hospitalized
130 COVID-19), susceptibility (affected versus unaffected population), and COVID-19 predicted
131 by flu-like symptoms; Other 6 traits are were in HGI release 3 (June 2020), including very
132 severe respiratory confirmed COVID-19 (versus the general population), COVID-19 infection
133 (versus negative control or population), hospitalized COVID-19 (versus not hospitalized

134 COVID-19 or population), and predicted COVID from self-reported symptoms (versus
135 predicted or self-reported non-COVID-19).

136 **Expression quantitative trait loci (eQTLs) and age-related gene expression in blood**

137 Blood eQTL data were obtained from the eQTLGen Consortium (31,684 whole blood
138 samples)²⁵. Only the significant near-independent eQTLs (FDR-q < 0.05, $r^2 < 0.05$) were used
139 in the Mendelian Randomization analysis.

140 The age-related transcriptomic change in whole blood was obtained from a large-scale
141 meta-analysis²⁶, which include six European-ancestry studies (n = 7,074 samples) and detected
142 roughly half of the genes in the human genome (n = 11,908). The direction and P-value of age-
143 related differential expression were directly obtained from the published dataset.

144 **Mendelian Randomization analysis**

145 Mendelian randomization is a method that uses genetic variants as instrumental variables to
146 determine whether an observational association between a risk factor and an outcome is
147 consistent with a potential causal effect²⁷. The multi-SNP MR analysis was implemented using
148 GSMR (Generalized Summary-data-based Mendelian randomization) in GCTA²².

149 As instruments for each exposure (four lifespan-related traits, four risk factors, and four
150 epigenetic age acceleration traits), we selected near-independent SNPs ($r^2 < 0.1$) with genome-
151 wide significant ($P < 5 \times 10^{-8}$) association with the exposure; For the expression of Notch1-4 in
152 whole blood, we selected significant near-independent eQTLs (FDR-q < 0.05, $r^2 < 0.05$); For
153 22 diseases from GERA community-based study, we selected SNPs with suggestive genome-
154 wide significance ($P < 1 \times 10^{-6}$) as instruments and performed a separate analysis due to the
155 limited case number in the community-based study. Full list of the genetic instruments are
156 shown in Supplementary data 1. GSMR includes a HEIDI-outlier filter to remove potential
157 pleiotropic SNPs that have effects on the exposures and the outcomes via different pathways.
158 We set its p-value threshold to 0.01 and tested the remaining SNPs for association with the
159 COVID-19-related traits. The required minimum number of instrumental SNPs for each
160 exposure in the analysis is lowered to 1.

161 **Bivariate genomic scan and functional Annotation**

162 To identify genetic variants associated with aging-related COVID-19 risk, we meta-analyzed
163 UKBB COVID-19 infection (with population control) and healthy aging (with the sign of effect
164 size reversed) summary statistics, while accounting for correlations between studies due to
165 sample overlap and correlation between the traits, as implemented in MultiABEL v1.1-610^{14,28}.

166 Summary association statistics were calculated for the 7,318,649 SNPs shared between studies.
167 These statistics represent the significance and consistency of each SNP affecting one or both
168 of the traits (e.g. the SNPs that significantly contribute to both aging and COVID-19 risk in the
169 same direction will have a smaller P-value). Therefore, we refer to this bivariate genomic scan
170 result as the aging-related COVID-19 throughout this study.

171 We then used the summary statistics of aging-related COVID-19 and performed
172 functional annotation for all SNPs in genomic areas identified by lead SNPs ($P < 1 \times 10^{-6}$, 250
173 Kb apart) using FUMA (Functional Mapping and Annotation)²⁹. The annotated genes were
174 then used for functional enrichment analysis using the default setting on the FUMA platform.

175 **Genetic correlation analysis**

176 We estimated the genetic correlations between lifespan-related traits, risk factors, epigenetic
177 age acceleration, and COVID-19 using LD score regression (LDSC) and high-definition
178 likelihood (HDL) methods^{30,31}. SNPs that are imperfectly imputed (INFO < 0.9) or with low
179 frequency (MAF < 0.05) were removed to reduce statistical noise. LDSC was performed using
180 LDSC software v1.0.1 (<https://github.com/bulik/ldsc>); and the HDL was performed using R
181 package "HDL" v1.3.8 (<https://github.com/zhenin/HDL>).

182 **Biological age estimation for UKBB cohorts**

183 All-cause mortality increases exponentially with age, and hence log-linear risk predictors from
184 proportional hazards models can provide natural composite organism state representations
185 characterizing the progress of aging based on biological and physiological measurements. We
186 used two such biological age measures: Phenotypic Age based on blood biochemistry³², and
187 Dynamic Organism State Indicator (DOSI) based on widely available Complete Blood Counts
188 (CBC) only³³. The latter is a proxy for the frailty index and is derived from the blood markers
189 only, whereas the Phenotypic Age additionally employs the explicit age. We also used physical
190 activity (the number of steps per day averaged over the week), which is also associated with
191 all-cause mortality and hence can also be viewed as a measure of biological aging³⁴.

192 We investigated the association of the incidence of COVID-19 with biological aging
193 acceleration (BAA, which is the difference between the biological age of a person and the
194 average biological age in the cohort of individuals of the same age and sex) using logistic
195 regression. Chronological age and biological sex were used as additional covariates in the
196 analysis.

197 Following UKBB recommendations, we used the “result” label from the table “COVID-
198 19 test results table” as the proxy of disease severity. This implies that mostly those individuals
199 that showed characteristic COVID-19 symptoms were selected for testing. We investigated
200 BAA associations with the incidence of COVID-19 and its associated fatality using all
201 available cases (All) and separately cohorts of individuals who have (Frail) or do not have (Not
202 Frail) major chronic diseases (from the list including all kinds of cancer, angina pectoris,
203 coronary heart disease, heart attack, heart failure, hypertension, stroke, diabetes, arthritis,
204 bronchitis, and emphysema) at the time of infection.

205 **Results**

206 We applied GSMR to test for potential causal associations between four lifespan-related traits
207 and COVID-19, including lifespan, longevity (i.e. surviving to the 90th percentile), healthspan
208 (time to a first major age-related disease), and healthy aging (multivariate meta-analysis of all
209 three traits combined) (Table S1). We employed summary-level GWAS data^{11-13,15} and selected
210 near-independent SNPs at a genome-wide significance level as genetic instruments for each
211 trait. The HEIDI-outlier filter was used to detect and eliminate genetic instruments with
212 pleiotropic effects on both exposure and outcome, as described by Zhu et al.²². For the
213 outcomes, we used 12 different sets of GWAS summary statistic data for COVID-19-related
214 traits from case-control studies (Table S1).

215 Strikingly, our GSMR analysis showed that long lifespan, longevity, as well as healthy
216 aging were protective against COVID-19 infection based on UKBB reporting (Fig. 1 A-E,
217 Table 1). The estimated odds ratio for lifespan was 0.31 (95% CI: 0.18 to 0.52; $P = 9.7 \times 10^{-6}$),
218 indicating that the risk of COVID-19 infection is decreased by 69% with approximately
219 every additional 10 years of life¹¹. For longevity, the odds ratio was 0.53 (95% CI: 0.43 to 0.65;
220 $P = 2.3 \times 10^{-9}$), and for healthy aging, the odds ratio was 0.12 (95% CI: 0.05 to 0.32; $P = 1.6$
221 $\times 10^{-5}$). As UKBB reporting on COVID-19 infection is biased toward severe cases of
222 hospitalized subjects, this measure might also be related to COVID-19 severity and mortality³⁵.
223 We observed a similar effect of these three traits on COVID-19 susceptibility in the HGI dataset
224 with overlapping confidential interval of estimated effect size, suggesting the association
225 between lifespan-related traits and COVID-19 infection is robust across different cohorts (Fig.
226 1A, Table 1). However, none of the lifespan-related traits showed a significant protective effect
227 on COVID-19 with a severe respiratory disorder or respiratory failure, possibly due to the small
228 case number of severe COVID-19 (Table S1). We also estimated the causal effect of genetically

229 proxied epigenetic age acceleration on the risk of COVID-19²⁰, while none of them shown to
230 have a significant effect on COVID-19 after Bonferroni correction for 144 tests (Fig. S1).

231 The loci for AD, CVD, T2D, cancer, and smoking (or lung cancer) explained the most
232 genetic variance of lifespan, as reported by Timmer et al.¹¹. To investigate whether these risk
233 factors contribute to the plausible causal association between lifespan and COVID-19, we
234 conducted an MR analysis of late-onset AD, coronary artery disease (CAD, including
235 myocardial infarction, percutaneous transluminal coronary angioplasty, coronary artery bypass
236 grafting, chronic ischemic heart disease, and angina), T2D, and smoking (the number of
237 cigarettes smoked per day) as exposures (Fig. S1 A-C, Supplementary data 2). Only the late-
238 onset AD was found to significantly increase the risk of COVID-19 after Bonferroni correction
239 for 144 tests ($P = 4.3 \times 10^{-6}$), with the odds ratio of 1.14 (95% CI: 1.08 to 1.21), suggesting the
240 benefit of a longer lifespan on the risk of COVID-19 may be partially mediated by less severe
241 or later occurring Alzheimer's disease.

242 To further evaluate the risk factors for COVID-19 infection and severity, we conducted
243 a separate MR analysis using GWAS data of 22 common diseases from GERA²¹ (Fig. S3,
244 Table S3). None of the diseases reached the significance threshold after Bonferroni correction
245 ($P = 0.05/276 = 2 \times 10^{-4}$). Among the nominally significant associations, asthma, dyslipidemia,
246 hernia abdominopelvic cavity, peptic ulcer, hypertensive disease, age-related macular
247 degeneration, and allergic rhinitis were the risk factors for COVID-19 infection; hypertensive
248 disease, irritable bowel syndrome, peripheral vascular disease, age-related macular
249 degeneration, and varicose veins were the risk factors for COVID-19 severity (Fig. S3,
250 Supplementary data 2). Interestingly, the disease count, a trait that represents the number of
251 comorbidities affecting each individual, was the risk factor for severe cases of COVID-19
252 (hospitalization). The odds ratio was 3.57 (95% CI: 1.04 to 12.28, $P = 0.04$), suggesting a
253 subject with roughly every two (one standard deviation) more comorbidities has a 3.6-fold
254 higher risk of having a severe case of COVID-19 (Table S2).

255 Healthspan is defined as the age period free of major age-related morbidities. In the
256 healthspan GWAS study, the top seven age-related morbidities were included (see Method)¹³.
257 In our analysis, healthspan did not show a significant effect on COVID-19-related traits (Fig
258 1A). This is unlikely to be due to the power of healthspan GWAS since there were 17 near-
259 independent genome-wide significant SNPs ($P < 5 \times 10^{-8}$), which is more than lifespan and
260 longevity. Therefore, we hypothesized that the strong protective effect of longevity against
261 COVID-19 may not be explained by the delayed appearance of age-related morbidities, but
262 rather by decelerated biological age that extends lifespan.

263 To address this hypothesis, we assessed in parallel the three different risk-based
264 biological age predictions computed for the subjects in the UKBB cohort using blood
265 biochemistry (Phenotypic Age), Complete Blood Counts (DOSI), and physical activity
266 measurements³²⁻³⁴ (Fig. 2A). We found that COVID-19 incidence in all UKBB datasets was
267 significantly associated with the BAA of Phenotypic Age, DOSI, and decreased physical
268 activity (Fig. 2B-E, Table 2). The estimated odds ratio of COVID-19 infection is 1.28 (95%
269 CI: 1.25 to 1.31; $P = 8.4 \times 10^{-82}$) and 1.31 (95% CI: 1.26 to 1.38; $P = 9.5 \times 10^{-32}$) for every ten
270 years higher biological age measured by Phenotypic Age and DOSI, respectively. Phenotypic
271 Age and DOSI were also significantly associated with COVID-19 incidence and case fatality
272 independently from the BAA association with chronic diseases, i.e., separately in cohorts of
273 UKBB individuals having (Frail) or not (Not frail) chronic age-related health conditions (Fig.
274 2E, Table 2).

275 We also observed elevated BAA levels of all measures of biological age (Fig. 2B-D, Fig.
276 S2A-D) in cohorts of individuals died from COVID-19 compared to those tested (and most
277 probably suffering from the disease), and, separately, in cohorts of those tested versus the rest
278 of UKBB (and presumed free of the disease). The number of UKBB subjects with data fields
279 required for the Phenotypic Age and DOSI was comparable, and we found that Phenotypic Age
280 comparisons produced a better statistical power. The number of UKBB subjects with physical
281 activity metrics was small, but the association of BAA in the form of physical activity deficit
282 and the incidence of COVID-19 was significant.

283 To gain the mechanistic insight of how aging and COVID-19 intertwined at the genetic
284 level, we performed a bivariate genomic scan using the GWAS of healthy aging and UKBB
285 COVID-19 infection, to identify the genetic variants that contribute to both aging and the risk
286 COVID-19 infection i.e., aging-related COVID-19 risk (Fig. S4, see Method). We identified
287 twenty bivariate loci at genome-wide significance ($P < 5 \times 10^{-8}$), where the null hypothesis is
288 no association with healthy aging and COVID-19 infection (Fig. S4). The summary statistics
289 of aging-related COVID-19 risk were then annotated using FUMA and a functional enrichment
290 analysis in 2868 canonical pathways (including gene sets from BIOCARTA, KEGG, PID,
291 REACTOME, and WikiPathways) and 7350 Gene Ontology (GO) biological processes was
292 performed. We find significant enrichment ($P_{\text{adjusted}} < 0.05$) in 67 canonical pathways and 26
293 biological processes. The canonical pathways with the strongest enrichment include pre-Notch
294 expression and processing ($P = 3.0 \times 10^{-8}$), signaling by Notch ($P = 3.6 \times 10^{-7}$), and oxidative
295 stress-induced senescence ($P = 1.4 \times 10^{-6}$) (Fig. 3A, Supplementary Data 3). Top enriched

296 biological processes are immune system development ($P = 2.3 \times 10^{-7}$) and myeloid cell
297 differentiation ($P = 2.4 \times 10^{-6}$), among others (Fig. 3B, Supplementary Data 3).

298 The Notch pathway is an evolutionally conserved signaling pathway, which is suggested
299 to be involved in both age-related inflammation and the development of age-related disease³⁶.
300 Moreover, the Notch signaling is related to the entry of SARS-CoV-2 through the positive
301 regulation of the host proteins that promote the entrance of the virus into the cell (e.g. FURIN
302 and ACE2)³⁷. In humans, there are four paralogs in the Notch family (Notch1-4)³⁸. We
303 hypothesized that the Notch signaling is a mediator for aging-related COVID-19 infection and
304 its effect may be related to the expression of Notch. This hypothesis was investigated with MR
305 of blood eQTLs of Notch1-4 from eQTLgen³⁹, against five traits that representing COVID-19
306 infection. We found that the expression of Notch2 significantly ($P < 0.05/20$) increases the risk
307 of COVID-19 infection (Fig. 3 B-D). The odds ratio estimated from HGI COVID-19 GWAS
308 was 1.31 (95% CI: 1.1 to 1.55; $P = 0.002$) per standard deviation higher Notch2 expression in
309 blood. We also observed a similar odds ratio estimate with overlapping 95% CI in other four
310 COVID-19 infection traits. This result suggests a causal role of Notch2 and more generally
311 Notch signaling in COVID-19 infection. To determine the age-related expression of Notch, we
312 examined the dataset of Peters et al.²⁶, which contains associations of genes with age in humans,
313 estimated from 7,074 whole blood samples. Among Notch1-4, only the Notch2 significantly
314 ($P = 0.007$) increased during aging, suggesting that the age-related increase of COVID-19 risk
315 is partially mediated through the increase of Notch2 expression. Moreover, we found that
316 Notch2 is primarily expressed in monocytes based on DICE (database of immune cell
317 expression, eQTLs, and epigenomics) project (Fig. S5)⁴⁰, which is also implicated in aging and
318 COVID-19⁴¹. This finding offers a new druggable target for preventing COVID-19 in the
319 elderly and is worthy of further experimental studies.

320 Finally, we estimated the genetic correlations between lifespan-related traits and
321 COVID-19 using LD score regression and high-definition likelihood (HDL) methods (Fig. 4,
322 supplementary data 4)^{30,31}. However, among 12 COVID-19-related traits, we only observed the
323 the positive genetic correlation between COVID-19 infection and AD at nominal significance
324 level, suggesting the case sample sizes of COVID-19 GWAS studies were too small to support
325 a sufficient power for estimating genetic correlations among these traits. Future genetic
326 analyses utilizing larger sample sizes should provide opportunities to improve these estimates.

327 Discussion

328 In this study, we explored a potential causal relationship between aging and the risk of COVID-
329 19 by conducting a multi-instrument MR analysis using four different lifespan-related traits as
330 exposures and eleven COVID-19-related traits as outcomes. We found that genetically proxied
331 longer lifespan and longevity were significantly associated with the decreased risk of COVID-
332 19 ($P = 9.7 \times 10^{-6}$ and 2.3×10^{-9} , respectively), and further analyses revealed a key role of an
333 elevated biological age and severity of chronic age-related diseases in this association. One of
334 the contributing factors is likely the immune response. The competence of the immune system
335 declines as people age, which is known as "immunosenescence"⁴². The hallmarks of
336 immunosenescence include an impaired response to new antigens, unsustained memory
337 responses, increased autoimmune responses, and prolonged inflammation. As a result, elderly
338 subjects are more susceptible to infectious diseases, including COVID-19, and have a poor
339 response to vaccines^{42,43}. On the other hand, it has been reported that the circulating immune
340 cells in centenarians possess unique characteristics that sustain immune responses to
341 infections⁴⁴. Moreover, the offspring of centenarians were shown to have a lower level of
342 inflammation⁴⁵, suggesting that the benefits on the immune system in centenarians are
343 heritable. Therefore, a better immunological profile in people with pro-longevity genetics may
344 help support the causal effect of longevity on COVID-19 we observed.

345 Notch pathway is an evolutionally conserved signaling pathway involved in age-related
346 inflammation and disease³⁶. Moreover, Notch signaling is related to the entry of SARS-CoV-
347 2 through the positive regulation of host proteins that promote entrance of the virus into the
348 cell³⁷. The entry of SARS-CoV-2 is mediated by binding of its S (spike) glycoprotein to the
349 Angiotensin Converting Enzyme 2 (ACE2)⁴⁶. Therefore, upregulation of ACE2 could
350 potentially increase the risk of viral infection. ADAM17 (A Disintegrin And Metalloproteases
351 17) is a metalloprotease that supports the shedding of ACE2 on the cell membrane⁴⁷. It is
352 negatively regulated by Notch signaling, whereas downregulation of ADAM17 significantly
353 reduced the shedding of ACE2³⁷. Besides ADAM17, a proteolytic cut of the S protein mediated
354 by furin after S glycoprotein binds to ACE2 is also required for the entry of SARS-CoV-2 into
355 the cell. The expression of furin is positively regulated by Notch signaling, and furin is also
356 involved in the maturation of Notch precursor³⁷. This evidence is in line with our finding that
357 Notch signaling plays an important role in aging-related COVID-19.

358 Notch2 is one of the four Notch paralogs in mammals. In our MR analysis, we found
359 evidence of the causal relationship between Notch2 expression and COVID-19 infection (Fig.

360 3 B-D). A previous study suggests that Notch2 promotes goblet cell metaplasia in the lung,
361 which is the hallmark of airway diseases⁴⁸. Moreover, the goblet cells are the major source of
362 ACE2 in the lung and were indicated to play a significant role in enabling COVID-19 infections.
363 Therefore, the increased Notch2 expression during aging might play a causal role in the
364 increased risk of COVID-19 infection in the elderly. We observed a relatively large effect size
365 (31% increased risk of infection for every 1 standard deviation higher Notch2 expression),
366 suggesting the Notch2 might provide a desirable target for the prevention and treatment of
367 COVID-19, as well as identifying population with a higher potential risk of infection. Further
368 experimental and clinical study on Notch2 and COVID-19 is needed to validate the causal
369 relationship.

370 Aging manifests as progressive remodeling of the organism, and hence a great number
371 of biological measurements are associated with age. Several sets of physiological and
372 biological indices have been proposed for quantification of aging progression in the form of a
373 single number – the biological age^{49,50}, or frailty index^{51,52}. One popular approach is to regress
374 relevant variables to predict chronological age and thus produce the “biological age”
375 prediction. Popular Hannum's and Horvath's methylation age-clock models, as well as other
376 clocks, are the widely used examples of such an approach^{53,54}.

377 An interesting alternative is to produce the log-linear all-cause mortality estimate with a
378 proportional hazard model and treat the resulting value as a measure of biological age.
379 Phenotypic Age from blood biochemistry markers³², DOSI from CBC³³, averaged physical
380 activity levels³⁴, and more sophisticated machine learning algorithms used to predict the risk
381 of death from physical activity time series of wearable devices⁵⁵, or even self-reported health
382 questionnaires, are all examples of this approach⁵⁶. All reliable biological age predictors are
383 associated with chronic disease burden, unhealthy lifestyles such as smoking (both overall and
384 in disease-free population), and future incidence of chronic diseases in healthy subjects<sup>32-
385 34,49,50,52,57,58</sup>. In our work, we also established the association of BAA with the risk of non-
386 chronic diseases, such as COVID-19 and the corresponding case fatality in the UKBB cohort
387 independent of disease burden. The association was significant for BAA measures obtained
388 from blood biochemistry (Phenotypic Age)³², CBC (DOSI)³³ and mean physical activity
389 (number of steps per day recorded by wearable devices over a week-long period of time³⁴; the
390 number of UKBB subjects with physical activity measurements was too low for separate BAA
391 characterization in frail and non-frail cohorts).

392 Decreased physical activity was associated with an increased risk of COVID-19 in the
393 UKBB cohort. This observation may be interesting on its own since the widespread lockdown

394 measures brought about a dramatic (up to 27.3%, which is 1,432 steps per day, within 30 days)
395 decline of average physical activity⁵⁹. Our association study suggests a more than 10% risk
396 increase corresponding to 1.5 thousand steps per day loss. There are feedback loop effects of
397 decreased mobility on BAA measures, and as such, the associated risk adjustments must be
398 taken into account in advanced epidemiological models of lockdown effects.

399 One advantage of our study design is that all of the BAA predictors were measured prior
400 to the pandemic. Therefore, the association between BAA and the risk of COVID-19 (and
401 probably other dangerous infectious diseases) is free of reverse causation (like in MR) and
402 likely to be causal if there are no other confounders. Thus, our research supports the idea of the
403 pro-active application of anti-aging (that is BAA-reducing) drugs in a prophylaxis mode to
404 protect the biomarker-defined vulnerable individuals. And, reversely, a significant reduction
405 of BA by an experimental drug in a clinical trial (probably as early as phase I) could warrant
406 further clinical studies in elderly subjects.

407 The association of BAA with case fatality was weaker (only Phenotypic Age BAA
408 exhibited a significant effect). This can be explained by the considerably smaller number of
409 UKBB subjects involved in the statistical analysis (346 of dead individuals compared to 11,619
410 tested (and presumed sick) and 459,872 overall subjects in UKBB). The case fatality rate
411 increases exponentially with age, and therefore it would be reasonable to expect the association
412 of BAA with the risk of death in COVID-19 patients³. We expect future studies to corroborate
413 our findings. Whether or not this association is causative could not be established in our study.
414 The age-dependent severity of COVID-19 has been demonstrated in many epidemiological
415 studies^{3,60}. However, we did not observe a significant association between longevity and
416 COVID-19 severity-related traits. This could be largely due to a small number of cases
417 included in GWAS analyses, varying from 536 to 1,610. Moreover, there is an underlying
418 selection bias toward symptomatic cases in the GWAS of COVID-19 positive cases, especially
419 in the UKBB study, which is overrepresented by severe and hospitalized cases³⁵. Therefore,
420 our results on UKBB may also be interpreted as a protective effect of longevity on the severe
421 form of COVID-19 infection.

422 It has been reported that people with comorbidities are more likely to suffer from
423 COVID-19 and have poor prognosis⁶¹. The most prevalent comorbidities were hypertension,
424 diabetes, and cardiovascular disease^{3,61}. Based on our Mendelian randomization analysis,
425 genetically predicted CAD, one of the most common types of cardiovascular disease, increases
426 COVID-19 susceptibility and the chance of being hospitalized after infection at a nominal
427 significance level, which is consistent with the result of recent study⁶². However, we didn't find

428 evidence of the causal effect of T2D and smoking on COVID-19-related traits. There are two
429 possible explanations: (1) the observed link between T2D and COVID-19 in epidemiological
430 studies is confounded by other factors; (2) the power of current COVID-19 GWAS results is
431 limited to reveal its potential causal link to T2D/smoking. Notably, a nominally significant
432 association ($P = 0.03$) between the life-time smoking index and risk of severe COVID-19 was
433 reported⁶³, suggesting a potential causal role of smoking on COVID-19. Future genetic
434 analyses with larger sample sizes or clinical experiments are required to fully identify the
435 relationship between smoking and COVID-19. The genetically predicted late-onset AD was
436 also shown to be significantly associated with a higher risk of infection based on both UKBB
437 and HGI report, which is not observed in another recent MR analysis⁶⁴. This is possibly due to
438 the larger case number (35,274 compared with 17,008) was included in late-onset AD GWAS,
439 thus increased the power of the MR analysis.

440 There are multiple clinical trials proposed to employ potential lifespan-extending drugs
441 to protect the elderly from COVID-19, based on promising observational data on metformin⁵⁻
442 ⁹. However, epidemiological studies are prone to confounding, reverse causation, and various
443 biases, and therefore are an unreliable indicator of the causal associations. MR is a method that
444 utilizes genetic instruments that are robustly associated with exposures, and thus generate more
445 reliable evidence in predicting novel interventions⁶⁵. In our MR analysis, we found evidence
446 for the causal relationship between longevity and decreased COVID-19. The following analysis
447 of genetic risk factors and phenotypic measurements suggests that this causal effect is likely to
448 be mediated by the decelerated rate of aging, which can be captured by biological age
449 measurements. Therefore, our finding can directly support lifespan-extending drugs as a
450 potential measure of COVID-19 when one of the following assumptions holds: (1) the selected
451 anti-aging drugs extend lifespan through a mechanism that mimics the genetics of longevity;
452 and (2) the selected anti-aging drugs could slow down or reverse the aging process measured
453 by biological age models (e.g., phenotypic age).

454 While the first assumption is hard to test, recent studies suggest that some anti-aging
455 interventions can slow down and even reverse the biological age measured by biological age
456 models⁶⁶. For example, a cocktail treatment of recombinant human growth hormone,
457 dehydroepiandrosterone, and metformin reversed the immunosenescent trend, and the
458 biological age measured by several biological age models (including PhenoAge) was reversed
459 by 2.5 years on average after 12 months of treatment⁶⁶. Thus, it could be worthwhile
460 prioritizing established anti-aging drugs in COVID clinical trials. This may accelerate COVID
461 drug development and save costs.

462

463 **Acknowledgements**

464 XS was in receipt of a Swedish Research Council (Vetenskapsrådet) Starting Grant (No. 2017-
465 02543). This research was also supported by NIA grants (to VNG). PF and TP were supported
466 by Gero PTE LLC (Singapore). We thank HGI and NIH-GRASP for the timely release of
467 COVID-19 GWAS summary statistics. We also thank Miss. Hanna Liu for assistance in part
468 of the analysis.

469

470 **Author contributions**

471 XS initiated the study; VNG, XS, and POF supervised the study; KY, RZ, and TVP performed
472 data analyses; KY, VNG, POF, and XS wrote the manuscript; All authors contributed to
473 manuscript writing.

474

475 **Competing interest statement**

476 The authors declare no competing financial interests.

477

478 **Data availability statement**

479 GWAS summary statistics used in this study are publicly available (for URLs, see Table S1).
480 The individual-level phenotype data are available by application from the UK Biobank
481 (<http://www.ukbiobank.ac.uk/>). The bivariate GWAS summary statistics of aging-related
482 COVID-19 generated in this study are available at
483 https://www.dropbox.com/s/bt6mtmttzgnhtfo/combined_ukbbCOVID_meta.txt?dl=0.

484

485 **Code availability statement**

486 The GSMR analysis was performed using GCTA 1.93.1beta available at
487 <https://cns.genomics.com/software/gcta/>. The genetic correlation analysis was performed using
488 LDSC v1.0.1 available at <https://github.com/bulik/ldsc>, and HDL v1.3.8 available at
489 <https://github.com/zhenin/HDL>.

490 **References**

491

- 492 1 Zhu, N. *et al.* A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N Engl J Med* 382, 727-
493 733, doi:10.1056/NEJMoa2001017 (2020).
- 494 2 Koff, W. C. & Williams, M. A. Covid-19 and Immunity in Aging Populations - A New Research Agenda. *N*
495 *Engl J Med*, doi:10.1056/NEJMp2006761 (2020).
- 496 3 Santesmasses, D. *et al.* COVID-19 is an emergent disease of aging. *medRxiv*,
497 doi:10.1101/2020.04.15.20060095 (2020).
- 498 4 Davies, N. G. *et al.* Age-dependent effects in the transmission and control of COVID-19 epidemics. *Nat*
499 *Med*, 1-7, doi:10.1038/s41591-020-0962-9 (2020).
- 500 5 Sargiacomo, C., Sotgia, F. & Lisanti, M. P. COVID-19 and chronological aging: senolytics and other anti-
501 aging drugs for the treatment or prevention of corona virus infection? *Aging (Albany NY)* 12, 6511-6517,
502 doi:10.18632/aging.103001 (2020).

- 503 6 Zhavoronkov, A. Geroprotective and senoremediative strategies to reduce the comorbidity, infection rates,
504 severity, and lethality in gerophilic and gerolavic infections. *Aging (Albany NY)* 12, 6492-6510,
505 doi:10.18632/aging.102988 (2020).
- 506 7 Omarjee, L. *et al.* Targeting T-cell senescence and cytokine storm with rapamycin to prevent severe
507 progression in COVID-19. *Clin Immunol* 216, 108464, doi:10.1016/j.clim.2020.108464 (2020).
- 508 8 Bramante, C. *et al.* Observational Study of Metformin and Risk of Mortality in Patients Hospitalized with
509 Covid-19. *medRxiv*, doi:10.1101/2020.06.19.20135095 (2020).
- 510 9 Luo, P. *et al.* Metformin Treatment Was Associated with Decreased Mortality in COVID-19 Patients with
511 Diabetes in a Retrospective Analysis. *Am J Trop Med Hyg* 103, 69-72, doi:10.4269/ajtmh.20-0375 (2020).
- 512 10 Roberts, R. Mendelian Randomization Studies Promise to Shorten the Journey to FDA Approval. *JACC*
513 *Basic Transl Sci* 3, 690-703, doi:10.1016/j.jacbts.2018.08.001 (2018).
- 514 11 Timmers, P. R. *et al.* Genomics of 1 million parent lifespans implicates novel pathways and common
515 diseases and distinguishes survival chances. *Elife* 8, doi:10.7554/eLife.39856 (2019).
- 516 12 Deelen, J. *et al.* A meta-analysis of genome-wide association studies identifies multiple longevity genes. *Nat*
517 *Commun* 10, 3669, doi:10.1038/s41467-019-11558-2 (2019).
- 518 13 Zenin, A. *et al.* Identification of 12 genetic loci associated with human healthspan. *Commun Biol* 2, 41,
519 doi:10.1038/s42003-019-0290-0 (2019).
- 520 14 Shen, X. *et al.* Multivariate discovery and replication of five novel loci associated with Immunoglobulin G
521 N-glycosylation. *Nat Commun* 8, 447, doi:10.1038/s41467-017-00453-3 (2017).
- 522 15 Timmers, P., Wilson, J. F., Joshi, P. K. & Deelen, J. Multivariate genomic scan implicates novel loci and
523 haem metabolism in human ageing. *Nat Commun* 11, 3570, doi:10.1038/s41467-020-17312-3 (2020).
- 524 16 Kunkle, B. W. *et al.* Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and
525 implicates Abeta, tau, immunity and lipid processing. *Nat Genet* 51, 414-430, doi:10.1038/s41588-019-
526 0358-2 (2019).
- 527 17 Nelson, C. P. *et al.* Association analyses based on false discovery rate implicate new loci for coronary artery
528 disease. *Nat Genet* 49, 1385-1391, doi:10.1038/ng.3913 (2017).
- 529 18 Xue, A. *et al.* Genome-wide association analyses identify 143 risk variants and putative regulatory
530 mechanisms for type 2 diabetes. *Nat Commun* 9, 2941, doi:10.1038/s41467-018-04951-w (2018).
- 531 19 Liu, M. *et al.* Association studies of up to 1.2 million individuals yield new insights into the genetic etiology
532 of tobacco and alcohol use. *Nat Genet* 51, 237-244, doi:10.1038/s41588-018-0307-5 (2019).
- 533 20 McCartney, D. L. *et al.* Genome-wide association studies identify 137 loci for DNA methylation biomarkers
534 of ageing. doi:10.1101/2020.06.29.133702 (2020).
- 535 21 Banda, Y. *et al.* Characterizing Race/Ethnicity and Genetic Ancestry for 100,000 Subjects in the Genetic
536 Epidemiology Research on Adult Health and Aging (GERA) Cohort. *Genetics* 200, 1285-1295,
537 doi:10.1534/genetics.115.178616 (2015).
- 538 22 Zhu, Z. *et al.* Causal associations between risk factors and common diseases inferred from GWAS summary
539 data. *Nat Commun* 9, 224, doi:10.1038/s41467-017-02317-2 (2018).
- 540 23 Ellinghaus, D. *et al.* Genomewide Association Study of Severe Covid-19 with Respiratory Failure. *N Engl J*
541 *Med*, doi:10.1056/NEJMoa2020283 (2020).
- 542 24 COVID-19 Host Genetics Initiative. The COVID-19 Host Genetics Initiative, a global initiative to elucidate
543 the role of host genetic factors in susceptibility and severity of the SARS-CoV-2 virus pandemic. *Eur J Hum*
544 *Genet* 28, 715-718, doi:10.1038/s41431-020-0636-6 (2020).
- 545 25 Vösa, U. *et al.* Unraveling the polygenic architecture of complex traits using blood eQTL metaanalysis.
546 *bioRxiv*, doi:10.1101/447367 (2018).
- 547 26 Peters, M. J. *et al.* The transcriptional landscape of age in human peripheral blood. *Nat Commun* 6, 8570,
548 doi:10.1038/ncomms9570 (2015).
- 549 27 Emdin, C. A., Khera, A. V. & Kathiresan, S. Mendelian Randomization. *JAMA* 318, 1925-1926,
550 doi:10.1001/jama.2017.17219 (2017).
- 551 28 Ning, Z. *et al.* Beyond power: Multivariate discovery, replication, and interpretation of pleiotropic loci using
552 summary association statistics. *bioRxiv*, doi:10.1101/022269 (2019).
- 553 29 Watanabe, K., Taskesen, E., van Bochoven, A. & Posthuma, D. Functional mapping and annotation of
554 genetic associations with FUMA. *Nat Commun* 8, 1826, doi:10.1038/s41467-017-01261-5 (2017).
- 555 30 Ning, Z., Pawitan, Y. & Shen, X. High-definition likelihood inference of genetic correlations across human
556 complex traits. *Nat Genet*, 1-6, doi:10.1038/s41588-020-0653-y (2020).
- 557 31 Bulik-Sullivan, B. *et al.* An atlas of genetic correlations across human diseases and traits. *Nat Genet* 47,
558 1236-1241, doi:10.1038/ng.3406 (2015).
- 559 32 Levine, M. E. *et al.* An epigenetic biomarker of aging for lifespan and healthspan. *Aging (Albany NY)* 10,
560 573-591, doi:10.18632/aging.101414 (2018).

- 561 33 Pyrkov, T. V. *et al.* Longitudinal analysis of blood markers reveals progressive loss of resilience and
562 predicts ultimate limit of human lifespan. *bioRxiv*, doi:10.1101/618876 (2020).
- 563 34 Pyrkov, T. V. *et al.* Quantitative characterization of biological age and frailty based on locomotor activity
564 records. *Aging (Albany NY)* 10, 2973-2990, doi:10.18632/aging.101603 (2018).
- 565 35 Griffith, G. *et al.* Collider bias undermines our understanding of COVID-19 disease risk and severity.
566 *medRxiv*, 2020.2005.2004.20090506, doi:10.1101/2020.05.04.20090506 (2020).
- 567 36 Balistreri, C. R., Madonna, R., Melino, G. & Caruso, C. The emerging role of Notch pathway in ageing:
568 Focus on the related mechanisms in age-related diseases. *Ageing Res Rev* 29, 50-65,
569 doi:10.1016/j.arr.2016.06.004 (2016).
- 570 37 Rizzo, P. *et al.* COVID-19 in the heart and the lungs: could we "Notch" the inflammatory storm? *Basic Res*
571 *Cardiol* 115, 31, doi:10.1007/s00395-020-0791-5 (2020).
- 572 38 Bray, S. J. Notch signalling in context. *Nat Rev Mol Cell Biol* 17, 722-735, doi:10.1038/nrm.2016.94 (2016).
- 573 39 Vösa, U. *et al.* Unraveling the polygenic architecture of complex traits using blood eQTL metaanalysis.
574 doi:10.1101/447367 (2018).
- 575 40 Schmiedel, B. J. *et al.* Impact of Genetic Polymorphisms on Human Immune Cell Gene Expression. *Cell*
576 175, 1701-1715 e1716, doi:10.1016/j.cell.2018.10.022 (2018).
- 577 41 Pence, B. D. Severe COVID-19 and aging: are monocytes the key? *Geroscience* 42, 1051-1061,
578 doi:10.1007/s11357-020-00213-0 (2020).
- 579 42 Goronzy, J. J. & Weyand, C. M. Understanding immunosenescence to improve responses to vaccines. *Nat*
580 *Immunol* 14, 428-436, doi:10.1038/ni.2588 (2013).
- 581 43 Alpert, A. *et al.* A clinically meaningful metric of immune age derived from high-dimensional longitudinal
582 monitoring. *Nat Med* 25, 487-495, doi:10.1038/s41591-019-0381-y (2019).
- 583 44 Hashimoto, K. *et al.* Single-cell transcriptomics reveals expansion of cytotoxic CD4 T cells in
584 supercentenarians. *Proc Natl Acad Sci U S A* 116, 24242-24251, doi:10.1073/pnas.1907883116 (2019).
- 585 45 Arai, Y. *et al.* in *EBioMedicine* Vol. 2 1549-1558 (2015).
- 586 46 Hoffmann, M. *et al.* SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a
587 Clinically Proven Protease Inhibitor. *Cell* 181, 271-280 e278, doi:10.1016/j.cell.2020.02.052 (2020).
- 588 47 Lambert, D. W. *et al.* Tumor necrosis factor-alpha convertase (ADAM17) mediates regulated ectodomain
589 shedding of the severe-acute respiratory syndrome-coronavirus (SARS-CoV) receptor, angiotensin-
590 converting enzyme-2 (ACE2). *J Biol Chem* 280, 30113-30119, doi:10.1074/jbc.M505111200 (2005).
- 591 48 Danahay, H. *et al.* Notch2 is required for inflammatory cytokine-driven goblet cell metaplasia in the lung.
592 *Cell Rep* 10, 239-252, doi:10.1016/j.celrep.2014.12.017 (2015).
- 593 49 Jylhava, J., Pedersen, N. L. & Hagg, S. Biological Age Predictors. *EBioMedicine* 21, 29-36,
594 doi:10.1016/j.ebiom.2017.03.046 (2017).
- 595 50 Lara, J. *et al.* A proposed panel of biomarkers of healthy ageing. *BMC Med* 13, 222, doi:10.1186/s12916-
596 015-0470-9 (2015).
- 597 51 Cardoso, A. L. *et al.* Towards frailty biomarkers: Candidates from genes and pathways regulated in aging
598 and age-related diseases. *Ageing Res Rev* 47, 214-277, doi:10.1016/j.arr.2018.07.004 (2018).
- 599 52 Kojima, G., Iliffe, S. & Walters, K. Frailty index as a predictor of mortality: a systematic review and meta-
600 analysis. *Age Ageing* 47, 193-200, doi:10.1093/ageing/afx162 (2018).
- 601 53 Horvath, S. DNA methylation age of human tissues and cell types. *Genome Biol* 14, R115, doi:10.1186/gb-
602 2013-14-10-r115 (2013).
- 603 54 Hannum, G. *et al.* Genome-wide methylation profiles reveal quantitative views of human aging rates. *Mol*
604 *Cell* 49, 359-367, doi:10.1016/j.molcel.2012.10.016 (2013).
- 605 55 Pyrkov, T. V. *et al.* Extracting biological age from biomedical data via deep learning: too much of a good
606 thing? *Sci Rep* 8, 5210, doi:10.1038/s41598-018-23534-9 (2018).
- 607 56 Pyrkov, T. V., Fedichev, P. O. & Moskalev, A. Biological Age is a Universal Marker of Aging, Stress, and
608 Frailty. *Healthy Ageing and Longevity* 10, 23-36, doi:10.1007/978-3-030-24970-0_3
609 info:doi/10.1007/978-3-030-24970-0_3 (2019).
- 610 57 F. Hillary, R. *et al.* *Epigenetic clocks predict prevalence and incidence of leading causes of death and*
611 *disease burden* (Cold Spring Harbor Laboratory, 2020).
- 612 58 Li, X. *et al.* Longitudinal trajectories, correlations and mortality associations of nine biological ages across
613 20-years follow-up. *Elife* 9, doi:10.7554/eLife.51507 (2020).
- 614 59 Tison, G. H. *et al.* Worldwide Effect of COVID-19 on Physical Activity: A Descriptive Study. *Ann Intern*
615 *Med*, doi:10.7326/M20-2665 (2020).
- 616 60 Bi, Q. *et al.* Epidemiology and transmission of COVID-19 in 391 cases and 1286 of their close contacts in
617 Shenzhen, China: a retrospective cohort study. *Lancet Infect Dis* 20, 911-919, doi:10.1016/S1473-
618 3099(20)30287-5 (2020).

- 619 61 Yang, J. *et al.* Prevalence of comorbidities and its effects in patients infected with SARS-CoV-2: a
620 systematic review and meta-analysis. *Int J Infect Dis* 94, 91-95, doi:10.1016/j.ijid.2020.03.017 (2020).
621 62 Wu, L., Zhu, J. & Wu, C. Mendelian randomization analysis to characterize causal association between
622 coronary artery disease and COVID-19. *medRxiv*, doi:10.1101/2020.05.29.20117309 (2020).
623 63 Mark, P. J. *et al.* Cardiometabolic traits, sepsis and severe covid-19 with respiratory failure: a Mendelian
624 randomization investigation. *medRxiv*, doi:10.1101/2020.06.18.20134676 (2020).
625 64 Liu, D. *et al.* Association between Alzheimer's disease and COVID-19: A bidirectional Mendelian
626 randomization. *medRxiv*, doi:10.1101/2020.07.27.20163212 (2020).
627 65 Davey Smith, G. & Hemani, G. Mendelian randomization: genetic anchors for causal inference in
628 epidemiological studies. *Hum Mol Genet* 23, R89-98, doi:10.1093/hmg/ddu328 (2014).
629 66 Fahy, G. M. *et al.* Reversal of epigenetic aging and immunosenescent trends in humans. *Aging Cell* 18,
630 e13028, doi:10.1111/accel.13028 (2019).

631

632 **Tables and Figures**

633 **Table 1. Mendelian randomization estimates for the association between lifespan-related traits and risk of**
634 **COVID-19**

<i>Exposure</i>	<i>Outcome</i>	<i>OR</i>	<i>95% CI</i>	<i>P</i>
<i>Healthy aging</i>	HGI_covid_susceptibility	0.33	0.13 - 0.85	2.2e-02
	UKBB_covid_vs_neg	0.25	0.09 - 0.69	7.0e-03
	UKBB_covid_vs_pop	0.12	0.05 - 0.32	1.6e-05
<i>Lifespan</i>	HGI_covid_susceptibility	0.45	0.27 - 0.77	3.2e-03
	HGI_covid_vs_neg	0.61	0.39 - 0.97	3.5e-02
	UKBB_covid_vs_neg	0.44	0.26 - 0.77	3.6e-03
	UKBB_covid_vs_pop	0.31	0.18 - 0.52	9.7e-06
<i>Longevity</i>	HGI_covid_susceptibility	0.68	0.56 - 0.82	8.5e-05
	UKBB_covid_vs_neg	0.58	0.47 - 0.72	5.0e-07
	UKBB_covid_vs_pop	0.53	0.43 - 0.65	2.3e-09

635 Only the associations that reached nominal significance ($P < 0.05$) are shown

636 **Table 2. Association between biological age acceleration and the risk of COVID-19**
637

<i>Biological age measurement</i>	<i>Outcome</i>	<i>OR</i>	<i>95% CI</i>	<i>P</i>
<i>Phenotypic Age</i>	COVID19 incidence (All)	1.28	1.25 - 1.31	8.4e-82
	Case fatality (All)	1.19	1.04 - 1.35	1.1e-02
	COVID19 incidence (Not frail)	1.12	1.04 - 1.2	1.9e-03
	Case fatality (Not frail)	1.72	1.17 - 2.51	5.4e-03
	COVID19 incidence (Frail)	1.26	1.23 - 1.3	3.7e-62
<i>DOSI</i>	COVID19 incidence (All)	1.31	1.26 - 1.38	9.5e-32
	COVID19 incidence (Not frail)	1.09	1.01 - 1.19	3.6e-02
	Case fatality (Not frail)	2.44	1.45 - 4.06	7.7e-04
	COVID19 incidence (Frail)	1.35	1.28 - 1.42	3.6e-28
<i>Physical Activity</i>	COVID19 incidence (All)	0.95	0.93 - 0.96	9.1e-19

638 Only the associations that reached nominal significance ($P < 0.05$) are shown.

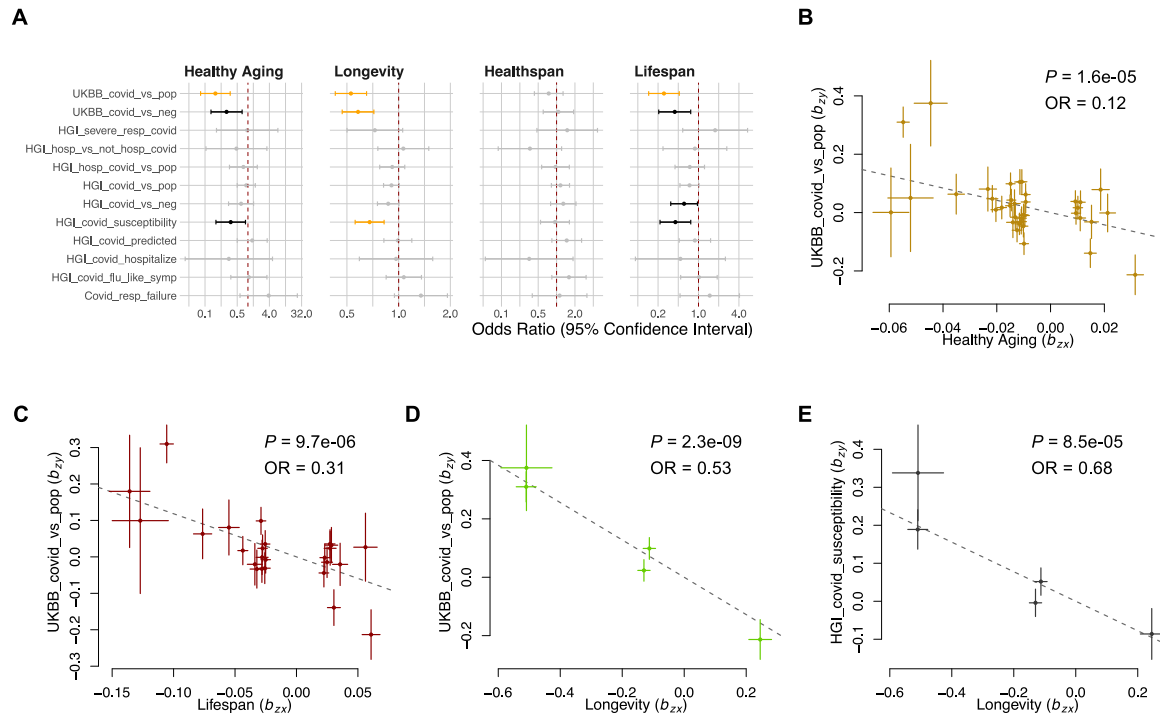
639

640 **Table 3. Mendelian randomization estimates for the association between Notch2 expression and risk of**
641 **COVID-19**
642

<i>Exposure</i>	<i>Outcome</i>	<i>OR</i>	<i>95% CI</i>	<i>P</i>
<i>NOTCH2</i>	HGI_covid_vs_neg	1.30	1.03 - 1.65	0.03020
	HGI_covid_vs_pop	1.31	1.1 - 1.55	0.00197
	UKBB_covid_vs_neg	1.46	1.08 - 1.99	0.01540
	UKBB_covid_vs_pop	1.43	1.07 - 1.91	0.01460

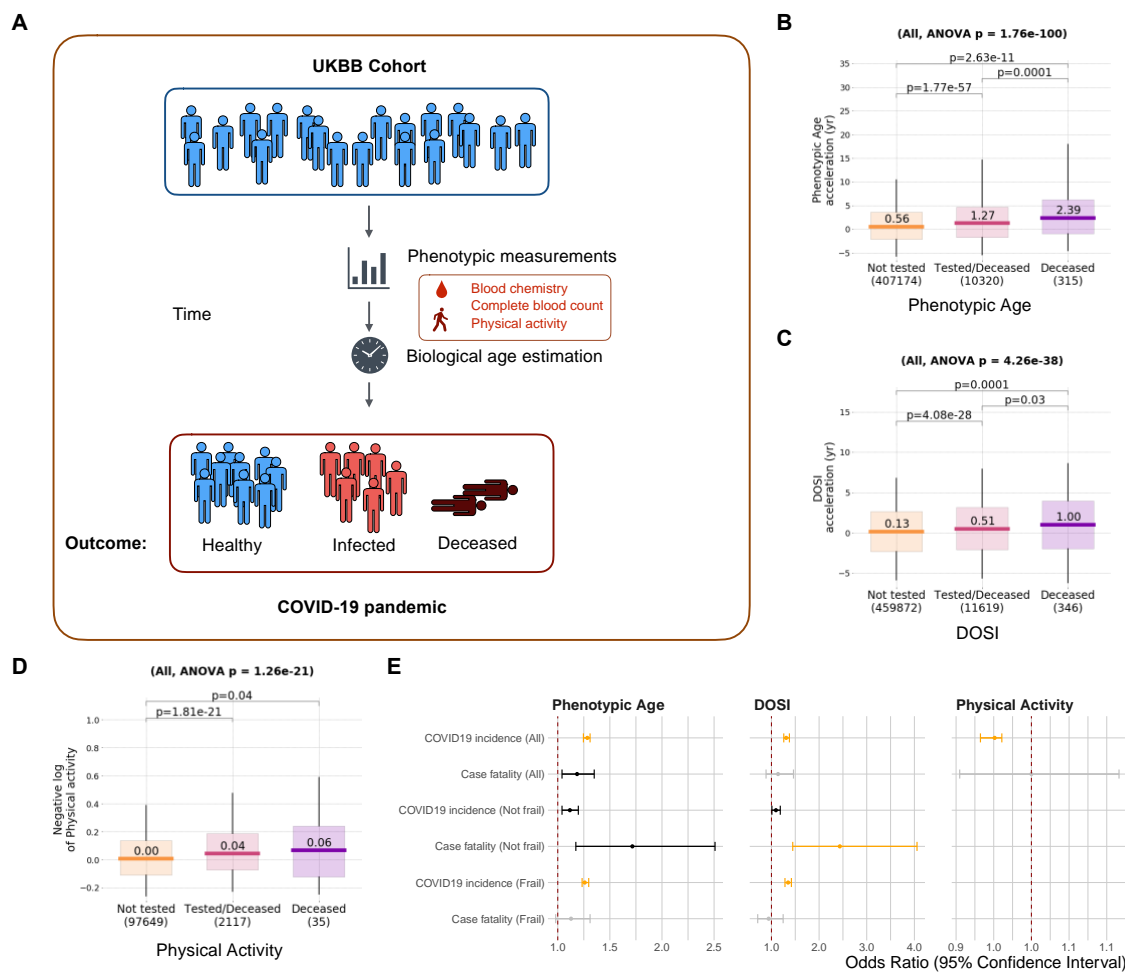
643 Only the associations that reached nominal significance ($P < 0.05$) are shown.

644



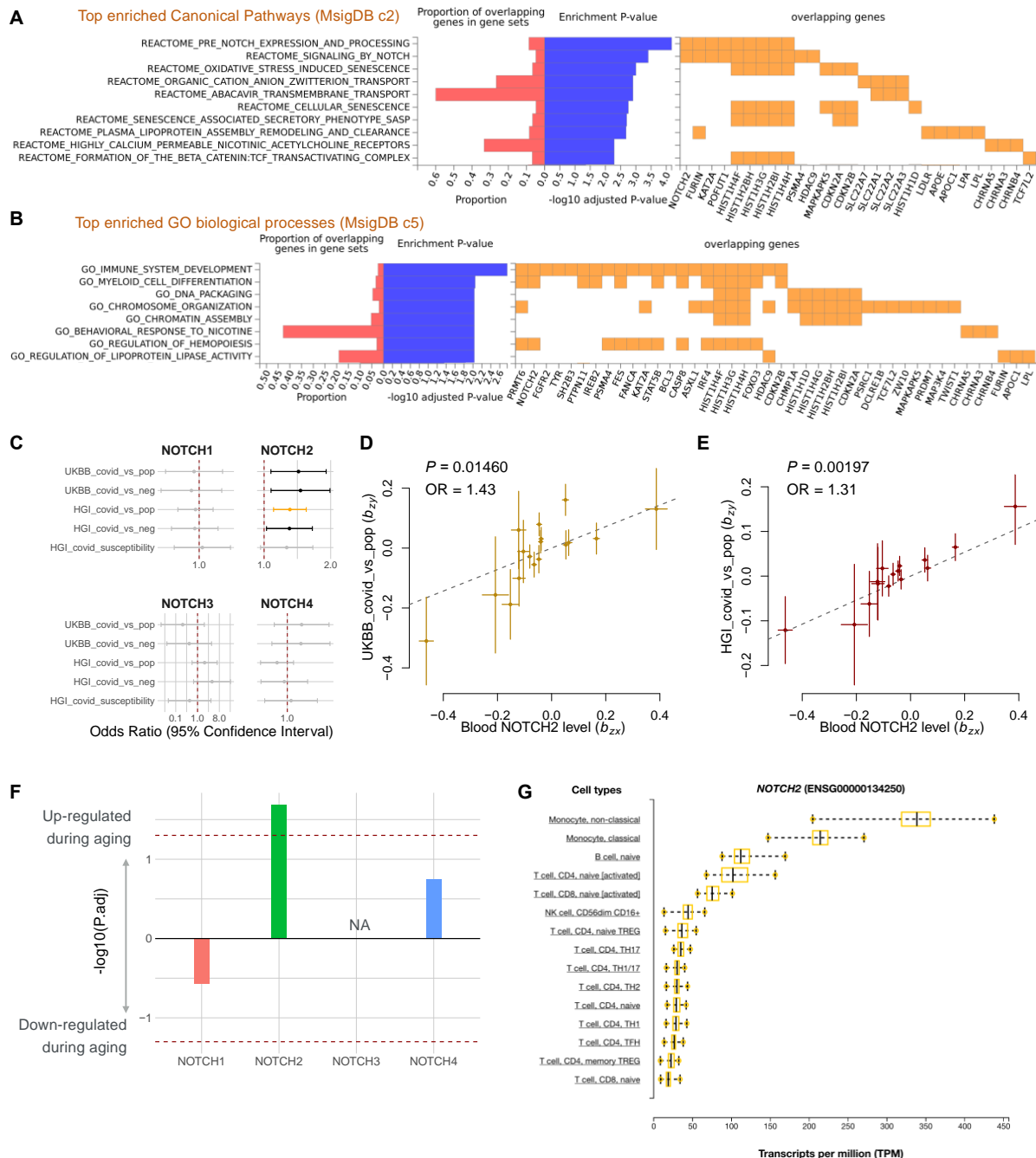
645

646 **Figure 1. Mendelian Randomization analysis investigating the association of lifespan-related traits with**
 647 **the risk of COVID-19.** **A.** The forest plot showing Mendelian randomization estimates for the causal effect of
 648 lifespan-related traits on the risk of COVID-19. Error bars show the 95% confidential interval. Significant
 649 effects after correcting for 144 tests (12 exposures and 12 outcomes, $P < 0.05/144$) are in orange. Nominally
 650 significant effects ($P < 0.05$) are in black. **B-E.** Plots of effect sizes of all genetic instruments from GWAS for
 651 healthy aging (**B**), lifespan (**C**), and longevity (**D**) (x-axis) versus those for UKBB COVID-19 (y-axis); and
 652 longevity (x-axis) versus HGI COVID-19 (y-axis) (**E**). Error bars represent standard errors.



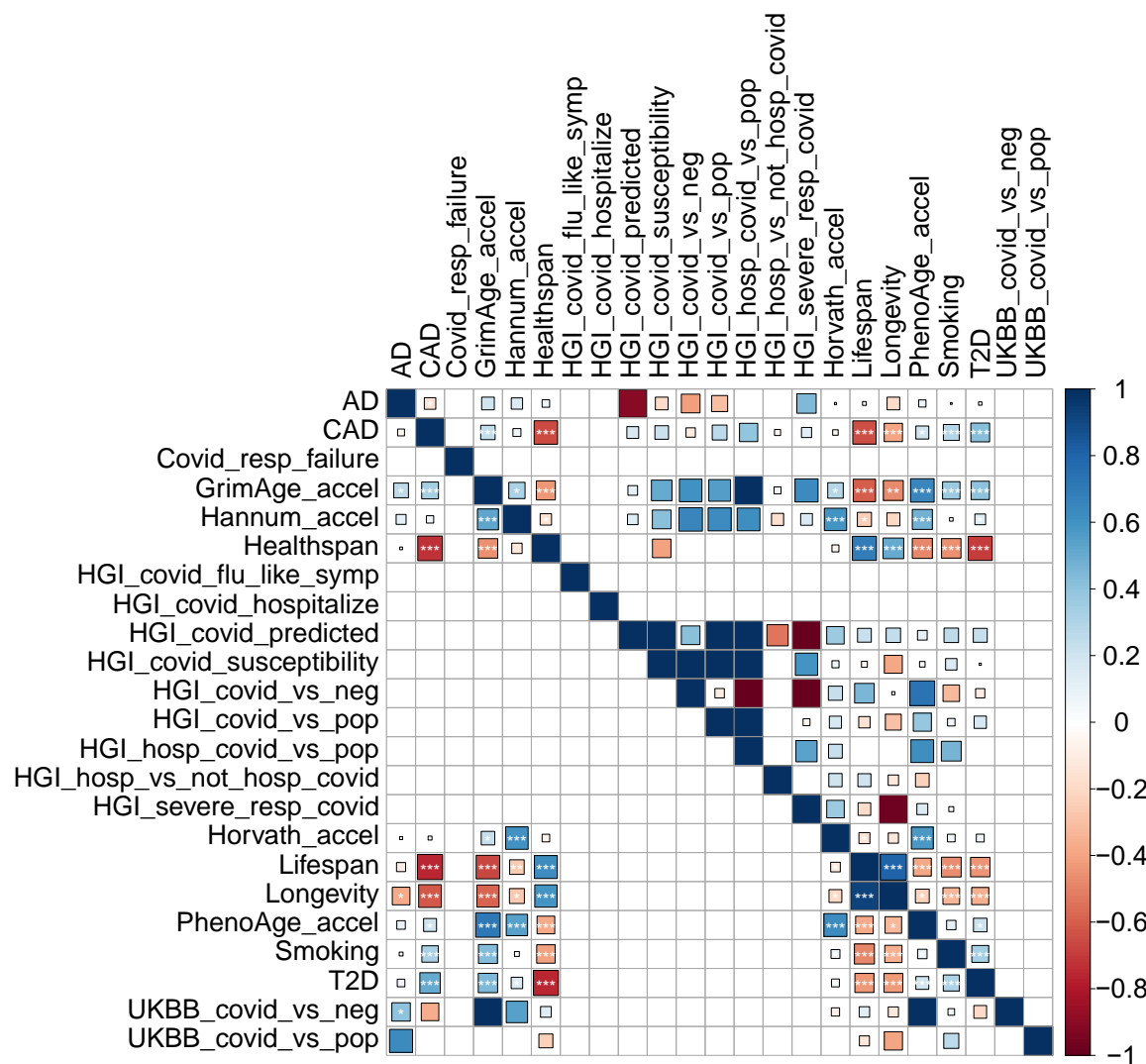
653

654 **Figure 2. Analysis of association of biological age acceleration with the risk of COVID-19.** A. Schematic
 655 representation of analysis of the biological age acceleration in UKBB cohort. B-D. the box plots showing the
 656 distribution of biological age acceleration measured by phenotypic age (B), DOSI (C), and negative log physical
 657 activity (D) in UKBB subjects that were not tested, tested, or died with COVID-19 infection. E. The forest plot
 658 showing the predicted effect of biological age acceleration on the risk of COVID-19 in different category. Error
 659 bars show the 95% confidential interval. Significant effects ($P < 0.001$) are in orange. Nominally significant
 660 effects ($P < 0.05$) are in black. Odds ratio for Phenotypic Age and Dynamic Organism State Index (DOSI) is
 661 given per 10-yr biological age acceleration. Odds ratio for physical activity is given per increase of 1000
 662 steps/day.



663

664 **Figure 3. Bivariate genomic scan identifying Notch signaling as age-related COVID-19 risk.** **A, B.** Gene set
 665 enrichment analysis of aging-related COVID-19. Top significantly enriched ($P_{adjusted} < 0.05$) canonical pathways
 666 (**A**) and GO biological processes (**B**) are shown. **C.** Forest plot showing Mendelian randomization estimates for
 667 the causal effect of blood Notch2 expression on the risk of COVID-19 infection. Error bars show the 95%
 668 confidential interval. Significant effects after correcting for 20 tests ($P < 0.0025$) are in orange. Nominally
 669 significant effects ($P < 0.05$) are in black. Error bars show the 95% confidential interval. **D, E.** Plots of effect
 670 sizes of all genetic instruments from blood eQTLs for Notch2 (x-axis) versus those for UKBB (**D**) and HGI (**E**)
 671 COVID-19 infection. Error bars represent standard errors. **F.** Bar plot showing the age-related
 672 differential expression of Notch1-4 in blood. The y-axis represents the $-\log_{10}(P_{adj}) \times \text{sign}$ of changing
 673 direction, i.e. positive value represents an age-related increase. **G.** Expression levels of Notch2 in transcripts per
 674 million (TPM, x-axis) from the DICE database. Cell types (y-axis) are sorted based on the median expression
 675 level within the cohort from highest to lowest. Boxes indicate 25%-75% interquartile ranges, and whiskers
 676 indicate minimum to maximum.



677

678 **Figure 4. Genetic correlation estimates from HDL and LDSC among phenotypes.** Lower triangle: HDL
 679 estimates; upper triangle: LDSC estimate. the areas of the squares represent the absolute value of corresponding
 680 genetic correlations. The genetic correlation that couldn't be estimated are in blank. P values are corrected using
 681 Bonferroni correction for 253 tests, * $P_{\text{nominal}} < 0.05$, ** $P_{\text{adjusted}} < 0.05$, *** $P_{\text{adjusted}} < 0.01$.