Using genetics and proteomics data to identify proteins causally related to COVID-19, healthspan and lifespan: a Mendelian randomization study

Jie V. Zhao[1,2,*], Minhao Yao[3,*], Zhonghua Liu[4]

[1]School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong SAR, China
[2]State Key Laboratory of Pharmaceutical Biotechnology, The University of Hong Kong, Hong Kong SAR, China
[3]Department of Statistics and Actuarial Science, The University of Hong Kong, Hong Kong SAR, China
[4]Department of Biostatistics, Mailman School of Public Health, Columbia University, New York, NY 10032, USA
[*]Equal contribution

Correspondence to: Jie V. Zhao, Zhonghua Liu; **email:** janezhao@hku.hk, zl2509@cumc.columbia.edu
Keywords: proteomics, healthspan, lifespan, COVID-19
Received: August 30, 2023 **Accepted:** January 24, 2024 **Published:** April 3, 2024

ABSTRACT

Background: COVID-19 pandemic poses a heavy burden on public health and accounts for substantial mortality and morbidity. Proteins are building blocks of life, but specific proteins causally related to COVID-19, healthspan and lifespan have not been systematically examined.

Methods: We conducted a Mendelian randomization study to assess the effects of 1,361 plasma proteins on COVID-19, healthspan and lifespan, using large GWAS of severe COVID-19 (up to 13,769 cases and 1,072,442 controls), COVID-19 hospitalization (32,519 cases and 2,062,805 controls) and SARS-COV2 infection (122,616 cases and 2,475,240 controls), healthspan (n = 300,477) and parental lifespan (~0.8 million of European ancestry).

Results: We identified 35, 43, and 63 proteins for severe COVID, COVID-19 hospitalization, and SARS-COV2 infection, and 4, 32, and 19 proteins for healthspan, father's attained age, and mother's attained age. In addition to some proteins reported previously, such as SFTPD related to severe COVID-19, we identified novel proteins involved in inflammation and immunity (such as ICAM-2 and ICAM-5 which affect COVID-19 risk, CXCL9, HLA-DRA and LILRB4 for healthspan and lifespan), apoptosis (such as FGFR2 and ERBB4 which affect COVID-19 risk and FOXO3 which affect lifespan) and metabolism (such as PCSK9 which lowers lifespan). We found 2, 2 and 3 proteins shared between COVID-19 and healthspan/lifespan, such as CXADR and LEFTY2, shared between severe COVID-19 and healthspan/lifespan. Three proteins affecting COVID-19 and seven proteins affecting healthspan/lifespan are targeted by existing drugs.

Conclusions: Our study provided novel insights into protein targets affecting COVID-19, healthspan and lifespan, with implications for developing new treatment and drug repurposing.

INTRODUCTION

Increasing lifespan and promoting healthy living into old age are among top priorities of health care systems all over the world. Although lifespan is steadily improving, an increasing proportion of the population is affected by multiple chronic diseases. Moreover, in recent years, the global COVID-19 pandemic poses a heavy burden on public health and accounts for substantial mortality and morbidity [1, 2]. According to the estimation of World Health Organization, around 14.83 million excess deaths globally were due to COVID-19 [1].

Proteins are the building blocks of life, important for etiology of disease and often used as drug targets for treatment. However, the specific proteins involved in the complex ageing process and COVID-19 risk are largely unclear. The dysregulation in inflammatory and immunomodulatory response has been thought to play a key role in multiple diseases and disorders [3], as well as in COVID-19 [4]. Several proteins related to inflammation and immune function, such as CXCL14 [5] and soluble HLA-G (sHLA-G) [6], have been identified by comparing patients with and without COVID-19. However, these associations may be confounded by some other characteristics, such as socioeconomic position. Moreover, proteins in other pathways, such as cardiometabolism, may also affect COVID-19 and longevity. Therefore, a comprehensive examination of proteins using advanced causal inference methods robust to confounding is needed.

Mendelian randomization (MR) uses genetic variants as instrument variables (IVs) to assess the causal role of proteins in COVID-19 and lifespan. As the genotypes are determined at conception and randomly allocated, MR estimates are thus not subject to residual confounding [7]. Using MR, some studies explored the associations of proteins with COVID-19 [8, 9] or with healthspan [10], but not both. In addition, the genetic instruments of these proteins are generally based on relatively smaller genome-wide association study (GWAS) (~3,301 people) [8–10]. Recently, the UK Biobank Pharma Proteomics Project (UKB-PPP), a collaboration between the UK Biobank (UKB) and thirteen biopharmaceutical companies has measured the plasma proteomic profiles in 54,306 UKB participants. In this study, we conducted MR analysis using genetic instruments from much larger GWAS of proteins and applied to large GWAS of COVID-19 (severe COVID-19, COVID-19 hospitalization and SARS-COV2 infection), healthspan and lifespan (mother's and father's attained age), to identify proteins causally related to COVID-19, healthspan and lifespan. We included both COVID-19 and healthspan and lifespan in the outcome, because COVID-19 which occurred in recent years reflects a new threat to longevity, whilst healthspan and lifespan reflect overall morbidity and mortality.

METHODS

Study design

We used proteome-wide MR study to identify proteins causally related to COVID-19, healthspan and lifespan. To better understand the function of these proteins, we checked the function of each selected protein. To assess the drug repurposing opportunity,

we also checked whether these proteins are targeted by existing drugs. The flow chart of the study design was shown in Figure 1.

Proteomics data from UK Biobank

The proteomic profiling measured 1,472 protein analytes and captured 1,463 unique proteins using the Olink Explore 1536 platform in 54,306 plasma samples in the UK Biobank. The blood samples were from a randomised subset of 46,673 UKB participants at baseline visit, 6,385 individuals at baseline selected by the UKB-PPP consortium and 1,268 individuals who participated in the COVID-19 repeat imaging study at multiple visits. No batch effects, plate effects or abnormalities in protein coefficients of variation were observed [11]. We used significant ($p < 3.4 \times 10^{-11}$) and independent ($r^2 < 0.01$) single nucleotide polymorphisms (SNPs) as instruments provided by the study of Sun et al. [1]. The details of the samples and sample selection in UKB-PPP were shown in Supplementary Methods. Up to 1,361 proteins with genetic instruments available in at least one of the outcome datasets were included in the analysis. The genetic instruments for proteins used in the MR analysis are provided in Supplementary Table 1.

Genetic associations with COVID-19

We obtained genetic associations of these genetic instruments with severe COVID-19 (up to 13,769 cases and 1,072,442 controls), hospitalization (up to 32,519 cases and 2,062,805 controls) and SARS-COV2 infection (up to 122,616 cases and 2,475,240 controls), from the latest GWAS results provided by the *COVID-19 host genetics initiative* [12]. Severe COVID-19 was defined as (1) hospitalization with laboratory confirmed SARS-CoV-2 infection based on RNA and/or serology and (2) hospitalization with COVID19 as the primary reason for admission, and (3) followed by death or respiratory support. COVID-19 hospitalization was defined as hospitalization with laboratory confirmed SARS-CoV-2 infection due to corona-related symptoms. COVID-19 infection was defined as (1) laboratory confirmed SARS-CoV-2 infection (RNA and/or serology based), or (2) physician diagnosis of COVID-19, or (3) self-report as COVID-19 positive. The GWAS was adjusted for age, age square, sex and the interaction of age and sex [12].

Genetic associations with healthspan and lifespan (mother's and father's attained age)

We obtained summary statistics for healthspan from a large GWAS conducted in 300,477 people of British ancestry in the UK Biobank. Healthspan was defined

as the age of the first incidence of congestive Heart Failure (CHF), Myocardial Infarction (MI), Chronic Obstructive Pulmonary Disease (COPD), stroke, dementia, diabetes, cancer, and death [13]. The GWAS used Cox-Gompertz survival models, with adjustment for sex, the first genetic principal components variables, assessment centre and genotyping batch [13].

We obtained genetic associations with parental survival (attained age) from a GWAS of parental longevity in European descent UK Biobank participants ($n = 415,311$ for father's attained age, $n = 412,937$ for mother's attained age) [14]. The GWAS used a Cox proportional hazards model to estimate offspring genetic variant effects on parental survival, stratified by sex, and adjusted for age and 10 principal components of ancestry. As the effect sizes obtained using genetic data from offspring are half of the actual

variant effect size in the parent, they were doubled to reflect the expected genetic effects in parents [14].

In both GWAS, the summary statistics from the GWAS were reported as log hazard ratios, for ease of understanding, these were converted to years of life by inverting the sign and multiplied by 10 [13, 14].

Statistical analysis

MR estimates were based on the genetic association with each type of COVID-19 (severe COVID-19, COVID-19 hospitalization or SARS-COV2 infection) divided by the genetic association with each protein), i.e., the Wald ratio estimates. Similarly, we obtained the MR estimates for the association of each protein with healthspan and lifespan. The genetic variant specific estimates were meta-analysed using inverse variance
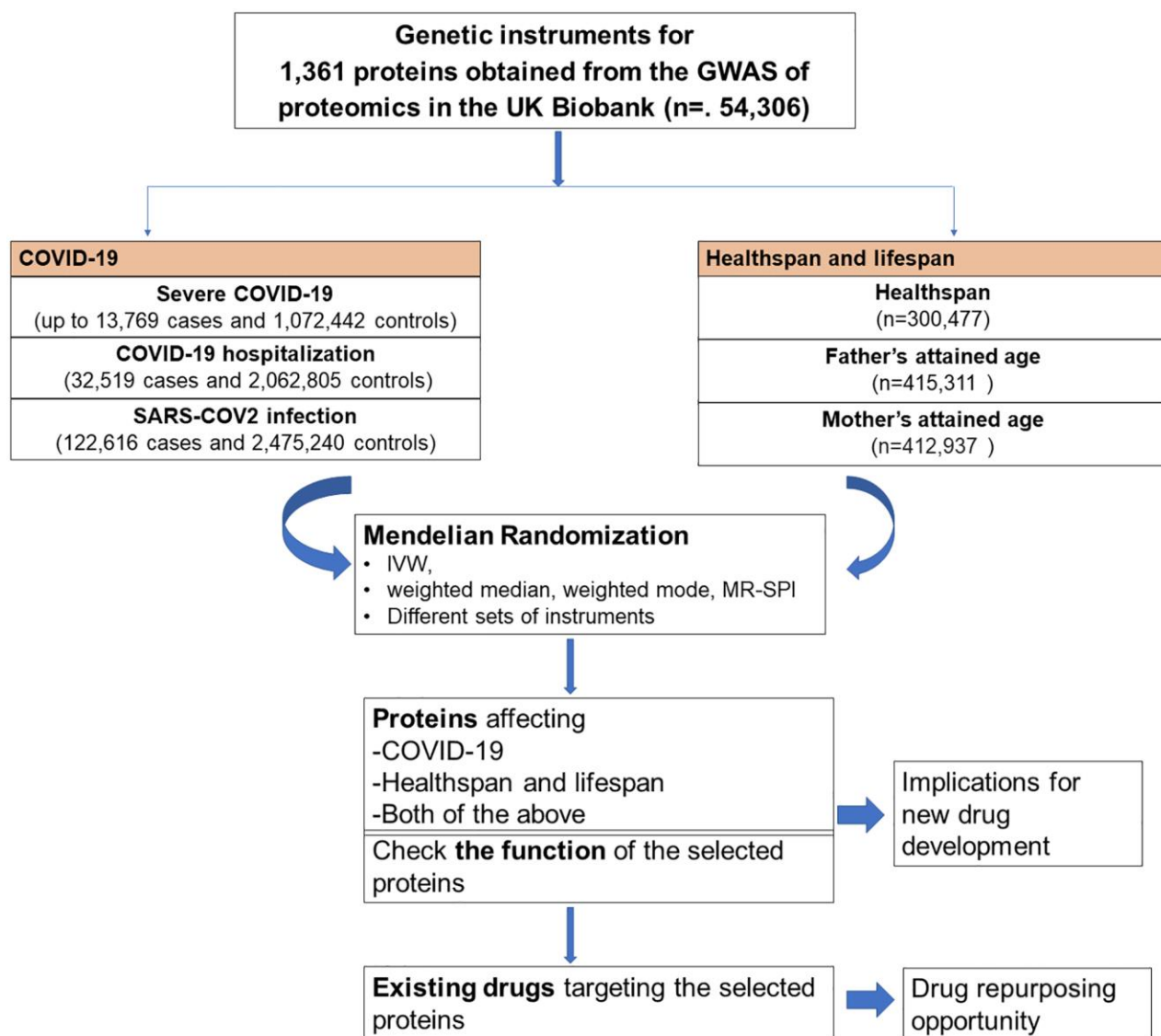

Figure 1. Flow chart of the study design.

weighting (IVW). Multiplicative random effects model was used when three or more genetic variants were used as instruments, and fixed effects model was used when less than three genetic variants were used as instruments. To account for multiple testing, we used Bonferroni corrected significance (p-value $< 0.05/1{,}361 = 3.7 \times 10^{-5}$) as the cut-off. Heterogeneity test was also conducted for the identified protein-outcome associations. In sensitivity analysis, for proteins with three or more genetic variants as instruments, we used different MR methods under different assumptions from IVW, including weighted median, weighted mode and MR-SPI. The weighted median method can provide consistent estimates even when up to 50% of the information comes from invalid genetic variants [15]. The weighted mode is based on the assumption that a plurality of genetic variants are valid instruments, i.e., no larger subset of invalid instruments estimating the same causal parameter than the subset of valid instruments exists [16]. Also based on the plurality assumption, MR-SPI can automatically select genetic variants as valid instruments and provide robust inference in finite samples [17]. Given that some proteins lack cis-SNPs as instruments, we used cis- and trans-SNPs as instruments in the main analysis and cis-SNPs in the sensitivity analysis.

Identification of drug repurposing opportunities

To have a better understanding of the biological function of each identified protein, we also looked into their functions in STRING [18], a database with comprehensive information on protein network and function, and UNIPROT (https://www.uniprot.org/). To identify potential drug repurposing opportunity, we checked whether the proteins are targeted by currently available drugs, by searching in DrugBank, a publicly available resource with drug and drug target information on over 13,000 drugs (https://www.drugbank.ca/).

All statistical analysis were performed using R (Foundation for Statistical Computing, Vienna, Austria; Version 4.1.1) and "TwoSampleMR", "MendelianRandomization", "ggplot2", "MR.SPI" R packages.

Availability of data and materials

All the data used in the study are publicly available. The data sources have been specified in the methods.

RESULTS

Proteins causally related to COVID-19

Figure 2 and Supplementary Tables 2–4 showed the associations of all proteins with three types of COVID-19, healthspan and lifespan (mother's and father's attained age). Among these proteins, we selected proteins with Bonferroni-corrected significance. Figure 3 showed the selected proteins affecting the risk of severe COVID-19

Figure 2. Volcano plot on the associations of proteins with severe COVID-19, COVID-19 hospitalization and SARS-COV2 infection.

(Figure 3A), COIVD-19 hospitalization (Figure 3B) and SARS-COV2 infection (Figure 3C). We identified 35 proteins associated with severe COVID (Figure 3A), 43 proteins associated with COVID-19 hospitalization (Figure 3B), and 63 proteins associated with SARS-COV2 infection (Figure 3C). There are 24 proteins shared by the three traits, including ADGRG2, AMY2B, CCL15, CD109, CD209, CD34, CX3CL1, FGF19, GOLM2, ICAM5, ISLR2, KLK1, LEFTY2, NRCAM, PECAM1, PODXL, PTPRM, REG1A, REG1B, SCG2, SEMA4C, SFTPD, TDGF1, and VAMP5. Among them, 8 proteins increased the risk of severe COVID-19, COVID-19 hospitalization and SARS-COV2 infection, 16 decreased the risk of COVID-19. Supplementary Table 5 showed the function of proteins affecting severe COVID-19, COVID-19 hospitalization and/or SARS-COV2 infection. These proteins had functions involved in inflammation and immunity, such as SFTPD, ICAM-2, ICAM-5, CD209, CD58, CCL15, CCL28, and MNDA; apoptosis, such as FGFR2 and ERBB4; and metabolism such as AMY2A and AMY2B.

In sensitivity analyses, we found that the estimates were generally robust to different MR methods, i.e., weighted median, weighted mode and MR-SPI (Supplementary Tables 6–11). The associations were also consistent when using cis-SNPs as instruments (Supplementary Tables 12–14).

Proteins causally related to healthspan and lifespan

Figure 4 and Supplementary Tables 15–17 showed the associations of all proteins with healthspan and lifespan (mother's and father's attained age). Figure 5 showed

the selected proteins associated with healthspan (Figure 5A), father's attained age (Figure 5B) and mother's attained age (Figure 5C). We identified 4 proteins related to healthspan (Figure 5A), 32 proteins related to father's attained age (Figure 5B), 19 proteins related to mother's attained age (Figure 5C). Supplementary Table 18 showed the function of proteins affecting healthspan and/or lifespan. The proteins are also involved in inflammation and immunity, such as CXCL9, HLA-DRA, LILRB4, IL19 and TNFRSF8, apoptosis, such as FOXO3, and involved in metabolism, such as PCSK9 and LDLR. Ten proteins play a role in both maternal and paternal ageing, including CDH1, CPE, CXCL9, F3, LAIR1, LEFTY2, LGALS9, LILRB4, POLR2F and RP2. The analysis was robust to weighted median, weighted mode and MR-SPI (Supplementary Tables 19–24). The associations were also consistent when using cis-SNPs as instruments (Supplementary Tables 25–27). Heterogeneity test suggested large heterogeneity for some protein-outcome associations (Supplementary Tables 28–33), such as AMY2B and severe COVID-19, in which case estimates from more robust methods are more valid.

Proteins related to both COVID-19 and ageing

According to Figures 3 and 5, we found two proteins shared between severe COVID-19 and healthspan/lifespan, including CXADR, and LEFTY2. Two proteins shared between COVID-19 hospitalization and healthspan/lifespan, including CXADR and LEFTY2. Three proteins were shared between SARS-COV2 infection and healthspan/lifespan, including CDH17, CXADR and LEFTY2.
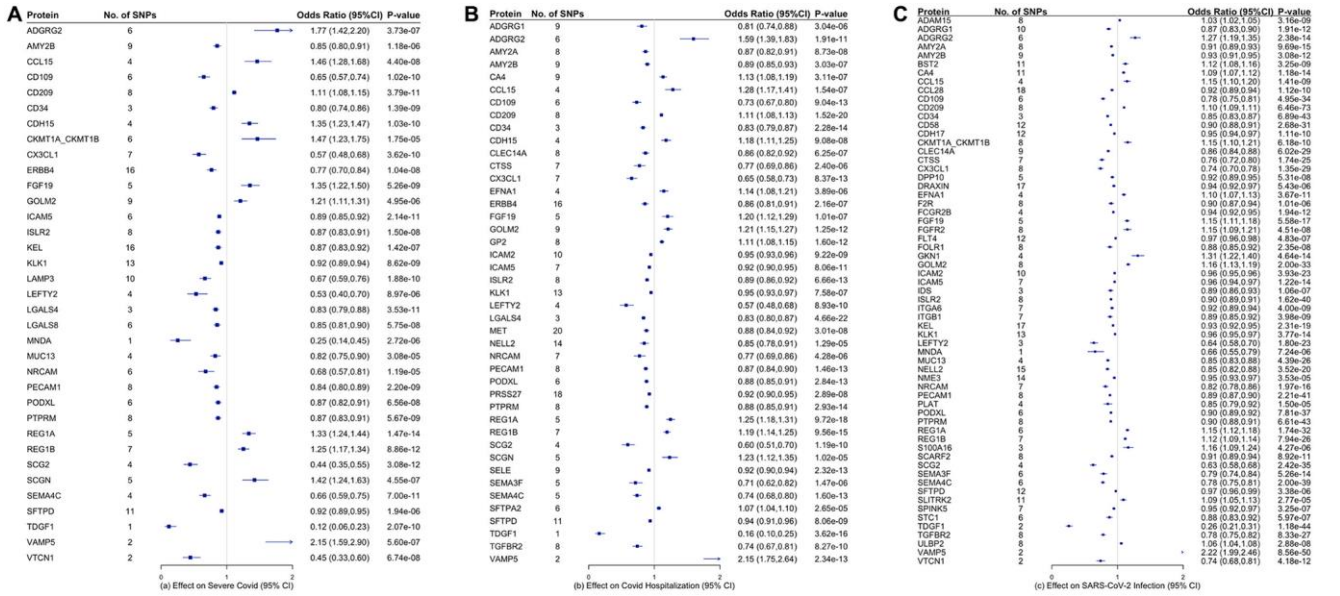

Figure 3. Proteins that are significantly associated with (**A**) severe COVID-19, (**B**) COVID-19 hospitalization, and (**C**) COVID-19 infection. The inverse-variance weighted (IVW) estimates are presented in log-odds ratio with the corresponding 95% confidence intervals.

Drug repurposing opportunity

Table 1 shows the proteins targeted by existing drugs. In this study, we found three proteins that affect the risk of COVID-19 in MR analysis have already been targeted by drugs currently used for the treatment of epilepsy, glaucoma, rheumatoid arthritis and neoplasm. Seven proteins that affect healthspan or lifespan have been targeted by drugs currently used for the treatment of cardiovascular disease, atherosclerosis, and neoplasm. This suggests that these existing drugs have the repurposing potential for lowering the risk of COVID-19 or improving healthspan/lifespan.

DISCUSSION

Using MR to minimize confounding bias, we not only confirmed some proteins reported in previous studies, such as SFTPD lowering the risk of severe COVID-19 [8], but also identified several novel proteins which are associated with the risk of COVID-19, with functions involved in inflammation and immunity, apoptosis and metabolism. We also identified novel proteins related to healthspan and lifespan, and some of them are also involved in inflammation and immunity, such as CXCL9, HLA-DRA, LILRB4, IL19 and TNFRSF8, in apoptosis, such as FOXO3, and involved in metabolism,
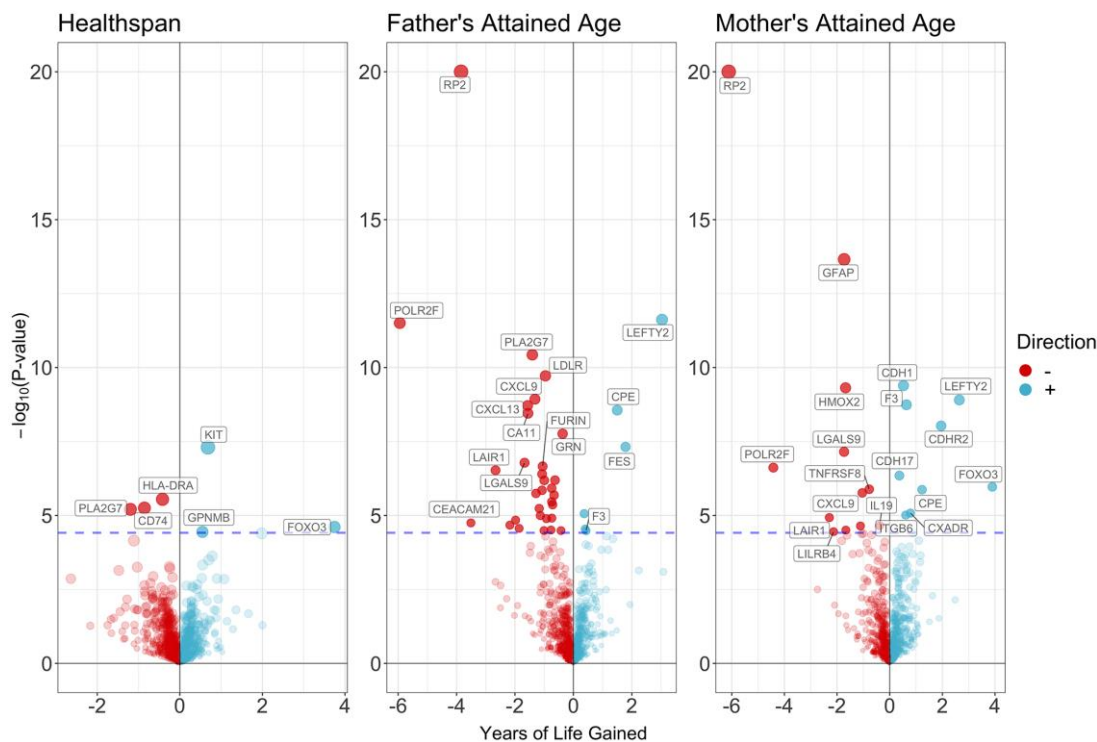

Figure 4. Volcano plot on the associations of proteins with healthspan and lifespan.
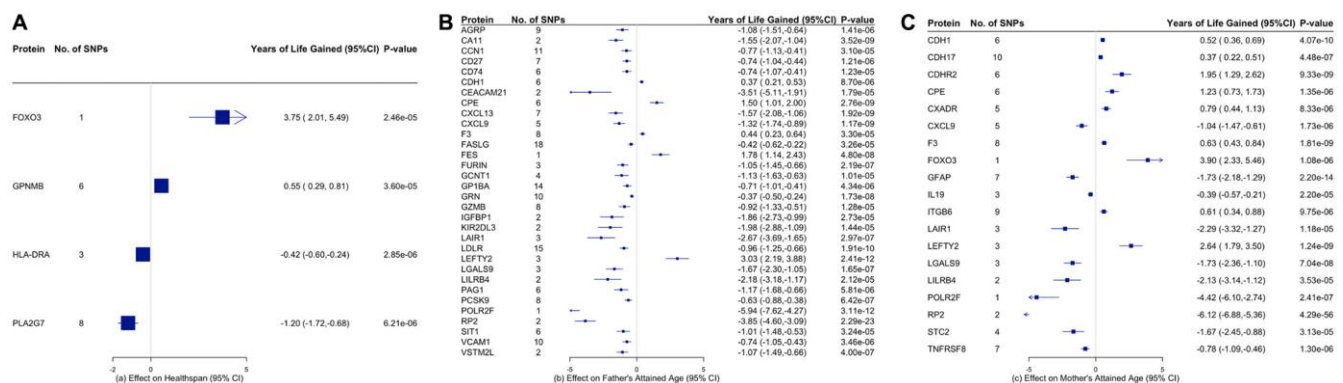

Figure 5. Proteins that are significantly associated with (**A**) healthspan, (**B**) father's attained age, and (**C**) mother's attained age. The inverse-variance weighted (IVW) estimates are presented in years of life gained with the corresponding 95% confidence intervals.

Table 1. Proteins targeted by existing drugs with potential for drug repurposing.

Protein	Findings in this MR	Drug(s) targeting this protein	Disease of treatment	Action type
CA4	Increases the risk of COVID-19 hospitalization and SARS-COV2 infection	METHAZOLAMIDE	Open-angle glaucoma	Inhibitor
		ACETAZOLAMIDE	Epilepsy	Inhibitor
		TOPIRAMATE	Migraine disorder	Inhibitor
		SULTHIAME	Epilepsy	Inhibitor
		DICHLORPHENAMIDE	Glaucoma	Inhibitor
FGFR2	Increases the risk of SARS-COV2 infection	NINTEDANIB	Pulmonary fibrosis; neoplasm	Inhibitor
		REGORAFENIB	Metastatic colorectal cancer	Inhibitor
		ERDAFITINIB	Neoplasm	Inhibitor
TDGF1	Lowers the risk of three COVID-19 related traits	BIIB-015	Neoplasm	Binding agent
CD74	Lowers father's lifespan	MILATUZUMAB	Multiple myeloma; chronic lymphocytic leukemia	Antagonist
GP1BA	Lowers father's lifespan	ANFIBATIDE	Myocardial infarction	Antagonist
GPNMB	Increases healthspan	GLEMBATUMUMAB VEDOTIN; GLEMBATUMUMAB	Breast cancer; melanoma	Binding agent
KIR2DL3	Lowers father's lifespan	IPH-2101; LIRILUMAB	Multiple myeloma	Inhibitor
PCSK9	Lowers father's lifespan	PCSK9 inhibitors, such as EVOLOCUMAB and INCLISIRAN	Cardiovascular disease	Inhibitor
PLA2G7	Lowers healthspan	DARAPLADIB	Atherosclerosis	Inhibitor
		RILAPLADIB	Atherosclerosis; Alzheimer's disease	Inhibitor
TNFRSF8	Lowers mother's lifespan	BRENTUXIMAB VEDOTIN	Lymphoma	Inhibitor

such as PCSK9 and LDLR. This finding implies that drugs targeting on these proteins may be used for disease prevention and treatment. For example, PCSK9 inhibitors, which have been used to treat hyperlipidemia, increase years of life. The identification of these proteins deepened the understanding of molecular mechanisms and provided new targets, with relevance to new drug development. Notably, we found 3 proteins affecting COVID-19 and 7 proteins affecting healthspan or lifespan are targeted by existing drugs, suggesting a great potential of drug repurposing. In the following discussion, we discussed in detail these proteins involved in these functions, as well as proteins targeted by existing drugs.

Proteins involved in inflammation and immune function

In the analysis of COVID-19, our findings are consistent with a previous MR study showing SFTPD is related to lower severe COVID-19 [8]. The protein is part of the innate immune response, the protein and the gene encoding this protein protects the lung against inhaled microorganisms and chemicals [19]. SFTPD also interacts with COVID-19 spike proteins [20]. Partly consistent with a previous MR study showing sICAM-2 lowers the risk of severe COVID-19 [21]

and in line with another MR study showing ICAM5 lowers the risk of severe COVID-19 [22], we found ICAM-2 and ICAM-5 lower the risk of COVID-19 hospitalization. ICAM-2 and ICAM-5 both belong to the Ig-like cell adhesion molecule family which may play a role in lymphocyte recirculation by blocking LFA-1-dependent cell adhesion. They mediate adhesive interactions important for antigen-specific immune response, NK-cell mediated clearance, lymphocyte recirculation, and other cellular interactions important for immune response and surveillance. The gene, *ICAM5,* is also related to severe COVID-19 [23].

Moreover, we identified several other inflammation-related proteins that are linked to COVID-19, such as CD209, CD58, CCL15, CCL28, and MNDA. CD209 is a pathogen-recognition receptor expressed on the surface of immature dendritic cells (DCs) and involved in initiation of primary immune response. *In vitro* experiment shows that CD209 also serves as alternative receptors for SARS-CoV-2 in disease-relevant cell types, including the vascular system [24]. This is in line with our findings that CD209 increased the risk of severe COVID-19, COVID-19 hospitalization and SARS-COV2 infection. CD58 is involved in activation of NK and T cells, a reduction in the expression of *CD58* results in reduced activation of NK and cytotoxic

T cells and may play a role in COVID-19 [25]. This may explain why we found CD58 lowers the risk of severe COVID-19 in our MR study. CCL15, a chemokine involved in leukocyte trafficking, was identified as predictor for severe COVID-19 [26]. CCL28 displays strong homing capabilities for B and T cells and orchestrates the trafficking and functioning of lymphocytes [27]. CCL28 may be used as an indicator for mucosal immune responses in people with SARS-COV2 infection [28]. In our study, we found CCL28 lowers the risk of SARS-COV2 infection. MNDA plays a role in the granulocyte/monocyte cell-specific response to interferon. It is required for INFα production from human blood cells in response to viruses [29], and may contribute to the immune response to SARS-CoV-2 [30].

Meanwhile, in the analysis for healthspan and lifespan, we also found several proteins which play a role in inflammation and immune function and have not been reported in previous MR studies, including CXCL9, HLA-DRA, LILRB4, IL19 and TNFRSF8. CXCL9 was involved in T cell trafficking. In animal experiments, it increased with aging, which can be prevented by calorie restriction, an established approach of increasing longevity [31]. Consistently, we found CXCL9 was related to shorter years of life. HLA-DRA was expressed on the surface of various antigen presenting cells such as B lymphocytes, dendritic cells, and monocytes/macrophages, and plays a central role in the immune system. The expression of HLA-DRA was higher in older people [32]. LILRB4 plays an important role in adaptive immunity, and increases with age in animal experiments [33]. In our study, we further suggested that increased HLA-DRA and LILRB4 lower healthspan and lifespan, respectively. TNFRSF8 may play a role in the regulation of cellular growth, transformation of activated lymphoblasts. Its regulating gene, *TNFRSF8*, has also been shown to relate to ageing [34]. IL19 promotes the production of IL6, which increased with age and its *in vitro* synthesis is prevented by diet restriction [35]. In our study, for the first time, we clearly showed that TNFRSF8 and IL19 possibly lowered lifespan.

Proteins involved in apoptosis

Apoptosis is an important process in aging. Consistent with previous studies [36], we found FOXO3 affects healthspan and maternal lifespan. The *FOXO3* gene functions as a trigger for apoptosis through expression of genes necessary for cell death. Notably, in the proteins related to COVID-19, we also found several proteins involved in apoptosis, such as FGFR2 and ERBB4. FGFR2 promotes gastric cancer progression by inhibiting the expression of Thrombospondin4 via

PI3K-Akt-Mtor pathway [37]. Drug targeting FGFR2 has been hypothesized to be used as treatments for COVID-19 [38] but this has not been tested in trials. ERBB4 is a tyrosine-protein kinase that plays an essential role as cell surface receptor for neuregulins and EGF family members and regulates cell proliferation, differentiation, migration and apoptosis. ERBB4 was downregulated after the coronavirus infection, and was upregulated by using the Wortmannin, which may inhibit the pathological cycle and development of SARS-CoV-2 in the human hosts [39].

Notably, we also found LEFTY2 contributes to both COVID-19 and lifespan. LEFTY2 lowers the risk of all COVID-19 phenotypes and increases paternal and maternal lifespan. *LEFTY2* encodes a secreted ligand of the transforming growth factor-beta (TGF-beta) superfamily of proteins, which acts as a crucial regulator of cell growth, proliferation, differentiation and apoptosis [40]. More studies are needed to clarify the pathways underlying the effects of LEFTY2, to provide insights for clinical practice and healthcare.

Proteins involved in metabolism

Metabolic syndrome is a known risk factor for COVID-19 and leads to multiple chronic diseases and mortality. Interestingly, we identified several proteins regulating metabolism. For example, we found AMY2B lowers the risk of severe COVID-19 and AMY2A lowers the risk of hospitalization. AMY2A and AMY2B both play a role in carbohydrate metabolism; the former (AMY2A) was FDA approved drug target. Evidently, the levels of AMY2A and AMY2B were relatively higher in survivors of COVID-19 compared to those who died of COVID-19 [41].

Among the proteins for aging, we found PLA2G7, PCSK9 and LDLR, which are all related to lipid metabolism, lower healthspan and/or lifespan. PLA2G7 has been associated with atherosclerosis, diabetes, and cardiovascular disease, and considered as a potential target that is a nexus between the immune, metabolic, and cardiovascular pathways of aging [42, 43]. PCSK9 and LDLR play an important role in lipid metabolism, and lead to higher risk of IHD, the leading cause of mortality. PCSK9 inhibitors, the drugs targeting PCSK9, are getting arising attention. In this novel study, we showed that PCSK9 inhibitors not only lower the cardiovascular risk, but also should be considered as a treatment for aging.

Proteins related to hormone regulation

Hormones have been thought to play a vital role in COVID-19 and ageing. In our study, we also found

proteins regulating hormone metabolism affected COVID-19 and healthspan/lifespan. For example, NELL2, which was involved in the regulation of hypothalamic GNRH secretion and the control of puberty, lowers the risk of COVID-19 hospitalization and SARS-COV2 infection. In line with our findings, *NELL2* is downregulated in patients with severe and mild COVID-19 in comparison to controls [44]. In proteins related to lifespan, we found IGFBP-1 lowers paternal lifespan. IGFBP-1 binds with IGF-1, which increased the risk of prostate cancer [45] and cardiometabolic diseases [46]. We also found AgRP lowers paternal lifespan. The main action of AgRP involves its antagonistic binding to melanocortin receptors 3 and 4, which are normally targeted by alpha Melanocyte Stimulating Hormone [47]. AgRP-deficiency was thought to lead to increased lifespan [48]. In our study, we confirmed its role in lifespan.

Drug repurposing opportunity

Importantly, we found that several proteins are targeted by existing drugs, which provided novel insights into drug repurposing. For example, we found that PLA2G7 which lowers healthspan is targeted by PLA2G7 inhibitors which are currently used in the treatment for atherosclerosis and Alzheimer disease. This suggests that PLA2G7 inhibitors may be repurposed to increase healthspan. Similarly, we suggested several drugs, such as FGFR2 inhibitors which have been used for the treatment of neoplasm can be considered to lower the risk of COVID-19. These novel findings are worthwhile to be tested in future randomized controlled trials.

Strengths and limitations

Using MR enables us to minimize unmeasured confounding bias in cohort studies and case-control studies. In this study, we used by far the largest available GWAS of proteomics, COVID-19, lifespan and healthspan, which improves the power to identify proteins causally related to these outcomes. Despite, we also acknowledge a few limitations. First, MR is based on three core assumptions, i.e., the relevance, independence, and exclusion-restriction assumption [49]. To satisfy these assumptions, we used genetic variants strongly associated with these proteins, with Bonferroni-corrected significance. We also used multiple sensitivity analysis methods robust to pleiotropy and checked the directions of associations from different genetic instruments (trans plus cis SNPs versus cis SNPs). In the situation where heterogeneity test suggested large heterogeneity, estimates from methods robust to pleiotropy are more valid than IVW. Considering that the measurements of proteins at baseline were conducted before the occurrence of the

outcomes, reverse causality is not a main concern, so we did not perform bi-directional MR analysis. Second, MR studies of COVID-19 might be subject to survivor bias (selection bias), i.e., people might have died of COVID-19 or other diseases before recruitment. So, the causal effects might be under-estimated. Third, population stratification might affect MR estimates. However, the GWAS data used in this study were derived from people largely of European ancestry. Meanwhile, as the study was based on people of European ancestry, the findings may not be generalizable to other ancestries. Fourth, the samples for proteins and outcomes both included UK Biobank. The partly overlapping in samples may bias two-sample MR estimates [2], but a simulation study shows two-sample MR can be safely conducted in a single large dataset [3], such as UK Biobank. So, the sample overlapping may not be a main concern, but the estimates need to be interpreted more cautiously. Finally, as we used summary statistics, we cannot assess the potential nonlinear association of these proteins with COVID-19, healthspan and lifespan.

CONCLUSIONS

In this study, using genetic instruments from by far the largest GWAS of proteomics, as well as by far the largest GWAS of COVID-19, healthspan and lifespan, we identified multiple proteins affecting COVID-19 and aging. We not only confirmed proteins identified in previous studies, such as SFTPD which affects COVID-19 and FOXO3 which affects healthspan, notably we identified several proteins with functions involved in inflammation and immune function, apoptosis, metabolism and hormone regulation. These findings provide novel insights into disease pathogenesis and pinpoint targets for therapeutic development or drug repurposing. For example, our findings of PCSK9 shortening lifespan and PLA2G7 shortening healthspan suggested that PCSK9 inhibitors, such as Evolocumab, and drugs inhibiting PLA2G7, such as Darapladib, may be beneficial for increasing lifespan and healthspan.

AUTHOR CONTRIBUTIONS

The manuscript was drafted by JVZ. Data analysis was conducted by MY, with the help of ZL and JVZ. JVZ developed the study conception, JVZ and ZL designed the analysis, and interpreted the results. All authors read and approved the final manuscript.

ACKNOWLEDGMENTS

and lifespan. The authors would like to thank all participants in the study and investigators for sharing the valuable data.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest related to this study.

ETHICAL STATEMENT

The study is an analysis using publicly available summary data that does not require ethical approval.

FUNDING

No funding was used for this paper.

REFERENCES

1. Msemburi W, Karlinsky A, Knutson V, Aleshin-Guendel S, Chatterji S, Wakefield J. The WHO estimates of excess mortality associated with the COVID-19 pandemic. Nature. 2023; 613:130–7. https://doi.org/10.1038/s41586-022-05522-2 PMID:36517599

2. Aburto JM, Kashyap R, Schöley J, Angus C, Ermisch J, Mills MC, Dowd JB. Estimating the burden of the COVID-19 pandemic on mortality, life expectancy and lifespan inequality in England and Wales: a population-level analysis. J Epidemiol Community Health. 2021; 75:735–40. https://doi.org/10.1136/jech-2020-215505 PMID:33468602

3. Furman D, Campisi J, Verdin E, Carrera-Bastos P, Targ S, Franceschi C, Ferrucci L, Gilroy DW, Fasano A, Miller GW, Miller AH, Mantovani A, Weyand CM, et al. Chronic inflammation in the etiology of disease across the life span. Nat Med. 2019; 25:1822–32. https://doi.org/10.1038/s41591-019-0675-0 PMID:31806905

4. Blanco-Melo D, Nilsson-Payant BE, Liu WC, Uhl S, Hoagland D, Møller R, Jordan TX, Oishi K, Panis M, Sachs D, Wang TT, Schwartz RE, Lim JK, et al. Imbalanced Host Response to SARS-CoV-2 Drives Development of COVID-19. Cell. 2020; 181:1036–45.e9. https://doi.org/10.1016/j.cell.2020.04.026 PMID:32416070

5. Bi X, Liu W, Ding X, Liang S, Zheng Y, Zhu X, Quan S, Yi X, Xiang N, Du J, Lyu H, Yu D, Zhang C, et al. Proteomic and metabolomic profiling of urine uncovers immune responses in patients with COVID-19. Cell Rep. 2022; 38:110271. https://doi.org/10.1016/j.celrep.2021.110271 PMID:35026155

6. Al-Bayatee NT, Ad'hiah AH. Soluble HLA-G is upregulated in serum of patients with severe COVID-19. Hum Immunol. 2021; 82:726–32. https://doi.org/10.1016/j.humimm.2021.07.007 PMID:34304938

7. Lawlor DA, Harbord RM, Sterne JA, Timpson N, Davey Smith G. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. Stat Med. 2008; 27:1133–63. https://doi.org/10.1002/sim.3034 PMID:17886233

8. Palmos AB, Millischer V, Menon DK, Nicholson TR, Taams LS, Michael B, Sunderland G, Griffiths MJ, Hübel C, Breen G, and COVID Clinical Neuroscience Study Consortium. Proteome-wide Mendelian randomization identifies causal links between blood proteins and severe COVID-19. PLoS Genet. 2022; 18:e1010042. https://doi.org/10.1371/journal.pgen.1010042 PMID:35239653

9. Gaziano L, Giambartolomei C, Pereira AC, Gaulton A, Posner DC, Swanson SA, Ho YL, Iyengar SK, Kosik NM, Vujkovic M, Gagnon DR, Bento AP, Barrio-Hernandez I, et al, and VA Million Veteran Program COVID-19 Science Initiative. Actionable druggable genome-wide Mendelian randomization identifies repurposing opportunities for COVID-19. Nat Med. 2021; 27:668–76. https://doi.org/10.1038/s41591-021-01310-z PMID:33837377

10. Perrot N, Pelletier W, Bourgault J, Couture C, Li Z, Mitchell PL, Ghodsian N, Bossé Y, Thériault S, Mathieu P, Arsenault BJ. A trans-omic Mendelian randomization study of parental lifespan uncovers novel aging biology and therapeutic candidates for chronic diseases. Aging Cell. 2021; 20:e13497. https://doi.org/10.1111/acel.13497 PMID:34704651

11. Sun BB, Chiou J, Traylor M, Benner C, Hsu YH, Richardson TG, Surendran P, Mahajan A, Robins C, Vasquez-Grinnell SG, Hou L, Kvikstad EM, Burren OS, et al. Genetic regulation of the human plasma proteome in 54,306 UK Biobank participants. bioRxiv. 2022. https://doi.org/10.1101/2022.06.17.496443

12. The COVID-19 Host Genetics Initiative. https://www.covid19hg.org/.

13. Zenin A, Tsepilov Y, Sharapov S, Getmantsev E, Menshikov LI, Fedichev PO, Aulchenko Y. Identification of 12 genetic loci associated with human healthspan. Commun Biol. 2019; 2:41.

https://doi.org/10.1038/s42003-019-0290-0
PMID:30729179

14. Pilling LC, Kuo CL, Sicinski K, Tamosauskaite J, Kuchel GA, Harries LW, Herd P, Wallace R, Ferrucci L, Melzer D. Human longevity: 25 genetic loci associated in 389,166 UK biobank participants. Aging (Albany NY). 2017; 9:2504–20.
https://doi.org/10.18632/aging.101334
PMID:29227965

15. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. Genet Epidemiol. 2016; 40:304–14.
https://doi.org/10.1002/gepi.21965
PMID:27061298

16. Hartwig FP, Davey Smith G, Bowden J. Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. Int J Epidemiol. 2017; 46:1985–98.
https://doi.org/10.1093/ije/dyx102
PMID:29040600

17. Yao M, Guo Z, Liu Z. Robust Mendelian Randomization Analysis by Automatically Selecting Valid Genetic Instruments for Inferring Causal Relationships between Complex Traits and Diseases. medRxiv. 2023.
https://doi.org/10.1101/2023.02.20.23286200

18. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P, Jensen LJ, Mering CV. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res. 2019; 47:D607–13.
https://doi.org/10.1093/nar/gky1131
PMID:30476243

19. Ortega FJ, Pueyo N, Moreno-Navarrete JM, Sabater M, Rodriguez-Hermosa JI, Ricart W, Tinahones FJ, Fernández-Real JM. The lung innate immune gene surfactant protein-D is expressed in adipose tissue and linked to obesity status. Int J Obes (Lond). 2013; 37:1532–8.
https://doi.org/10.1038/ijo.2013.23
PMID:23478426

20. Chen L, Zheng S. Understand variability of COVID-19 through population and tissue variations in expression of SARS-CoV-2 host genes. Inform Med Unlocked. 2020; 21:100443.
https://doi.org/10.1016/j.imu.2020.100443
PMID:33072849

21. Zhu J, Wu C, Wu L. Associations Between Genetically Predicted Protein Levels and COVID-19 Severity. J Infect Dis. 2021; 223:19–22.

https://doi.org/10.1093/infdis/jiaa660
PMID:33083826

22. Zheng J, Zhang Y, Zhao H, Liu Y, Baird D, Karim MA, Ghoussaini M, Schwartzentruber J, Dunham I, Elsworth B, Roberts K, Compton H, Miller-Molloy F, et al. Multi-ancestry Mendelian randomization of omics traits revealing drug targets of COVID-19 severity. EBioMedicine. 2022; 81:104112.
https://doi.org/10.1016/j.ebiom.2022.104112
PMID:35772218

23. Velavan TP, Pallerla SR, Rüter J, Augustin Y, Kremsner PG, Krishna S, Meyer CG. Host genetic factors determining COVID-19 susceptibility and severity. EBioMedicine. 2021; 72:103629.
https://doi.org/10.1016/j.ebiom.2021.103629
PMID:34655949

24. Amraei R, Yin W, Napoleon MA, Suder EL, Berrigan J, Zhao Q, Olejnik J, Chandler KB, Xia C, Feldman J, Hauser BM, Caradonna TM, Schmidt AG, et al. CD209L/L-SIGN and CD209/DC-SIGN Act as Receptors for SARS-CoV-2. ACS Cent Sci. 2021; 7:1156–65.
https://doi.org/10.1021/acscentsci.0c01537
PMID:34341769

25. van Eeden C, Khan L, Osman MS, Cohen Tervaert JW. Natural Killer Cell Dysfunction and Its Role in COVID-19. Int J Mol Sci. 2020; 21:6351.
https://doi.org/10.3390/ijms21176351
PMID:32883007

26. Lucas C, Wong P, Klein J, Castro TBR, Silva J, Sundaram M, Ellingson MK, Mao T, Oh JE, Israelow B, Takahashi T, Tokuyama M, Lu P, et al, and Yale IMPACT Team. Longitudinal analyses reveal immunological misfiring in severe COVID-19. Nature. 2020; 584:463–9.
https://doi.org/10.1038/s41586-020-2588-y
PMID:32717743

27. Mohan T, Deng L, Wang BZ. CCL28 chemokine: An anchoring point bridging innate and adaptive immunity. Int Immunopharmacol. 2017; 51:165–70.
https://doi.org/10.1016/j.intimp.2017.08.012
PMID:28843907

28. Yan Y, Jiang X, Wang X, Liu B, Ding H, Jiang M, Yang Z, Dai Y, Ding D, Yu H, Zhang S, Liu J, Sha M, et al. CCL28 mucosal expression in SARS-CoV-2-infected patients with diarrhea in relation to disease severity. J Infect. 2021; 82:e19–21.
https://doi.org/10.1016/j.jinf.2020.08.042
PMID:32871180

29. Gu L, Casserly D, Brady G, Carpenter S, Bracken AP, Fitzgerald KA, Unterholzner L, Bowie AG. Myeloid cell nuclear differentiation antigen controls the pathogen-stimulated type I interferon cascade in human monocytes by transcriptional regulation of IRF7. Nat Commun. 2022; 13:14.

https://doi.org/10.1038/s41467-021-27701-x
PMID:35013241

30. Beltrán-Camacho L, Eslava-Alcón S, Rojas-Torres M, Sánchez-Morillo D, Martinez-Nicolás MP, Martín-Bermejo V, de la Torre IG, Berrocoso E, Moreno JA, Moreno-Luna R, Durán-Ruiz MC. The serum of COVID-19 asymptomatic patients up-regulates proteins related to endothelial dysfunction and viral response in circulating angiogenic cells ex-vivo. Mol Med. 2022; 28:40.
https://doi.org/10.1186/s10020-022-00465-w
PMID:35397534

31. Trott DW, Lesniewski LA, Donato AJ. Selected life-extending interventions reduce arterial CXCL10 and macrophage colony-stimulating factor in aged mouse arteries. Cytokine. 2017; 96:102–6.
https://doi.org/10.1016/j.cyto.2017.03.008
PMID:28390264

32. de Almeida Chuffa LG, Freire PP, Dos Santos Souza J, de Mello MC, de Oliveira Neto M, Carvalho RF. Aging whole blood transcriptome reveals candidate genes for SARS-CoV-2-related vascular and immune alterations. J Mol Med (Berl). 2022; 100:285–301.
https://doi.org/10.1007/s00109-021-02161-4
PMID:34741638

33. Jonker MJ, Melis JP, Kuiper RV, van der Hoeven TV, Wackers PFK, Robinson J, van der Horst GT, Dollé ME, Vijg J, Breit TM, Hoeijmakers JH, van Steeg H. Life spanning murine gene expression profiles in relation to chronological and pathological aging in multiple organs. Aging Cell. 2013; 12:901–9.
https://doi.org/10.1111/acel.12118
PMID:23795901

34. Tanaka T, Biancotto A, Moaddel R, Moore AZ, Gonzalez-Freire M, Aon MA, Candia J, Zhang P, Cheung F, Fantoni G, Semba RD, Ferrucci L, and CHI consortium. Plasma proteomic signature of age in healthy humans. Aging Cell. 2018; 17:e12799.
https://doi.org/10.1111/acel.12799
PMID:29992704

35. Ershler WB, Sun WH, Binkley N, Gravenstein S, Volk MJ, Kamoske G, Klopp RG, Roecker EB, Daynes RA, Weindruch R. Interleukin-6 and aging: blood levels and mononuclear cell production increase with advancing age and in vitro production is modifiable by dietary restriction. Lymphokine Cytokine Res. 1993; 12:225–30.
PMID:8218595

36. Willcox BJ, Tranah GJ, Chen R, Morris BJ, Masaki KH, He Q, Willcox DC, Allsopp RC, Moisyadi S, Poon LW, Rodriguez B, Newman AB, Harris TB, et al. The FoxO3 gene and cause-specific mortality. Aging Cell. 2016; 15:617–24.

https://doi.org/10.1111/acel.12452
PMID:27071935

37. Huang T, Liu D, Wang Y, Li P, Sun L, Xiong H, Dai Y, Zou M, Yuan X, Qiu H. FGFR2 Promotes Gastric Cancer Progression by Inhibiting the Expression of Thrombospondin4 via PI3K-Akt-Mtor Pathway. Cell Physiol Biochem. 2018; 50:1332–45.
https://doi.org/10.1159/000494590
PMID:30355943

38. Sonkar C, Doharey PK, Rathore AS, Singh V, Kashyap D, Sahoo AK, Mittal N, Sharma B, Jha HC. Repurposing of gastric cancer drugs against COVID-19. Comput Biol Med. 2021; 137:104826.
https://doi.org/10.1016/j.compbiomed.2021.104826
PMID:34537409

39. Selvaraj G, Kaliamurthi S, Peslherbe GH, Wei DQ. Identifying potential drug targets and candidate drugs for COVID-19: biological networks and structural modeling approaches. F1000Res. 2021; 10:127.
https://doi.org/10.12688/f1000research.50850.3
PMID:33968364

40. Lu WQ, Qiu JL, Huang ZL, Liu HY. Enhanced circulating transforming growth factor beta 1 is causally associated with an increased risk of hepatocellular carcinoma: a mendelian randomization meta-analysis. Oncotarget. 2016; 7:84695–704.
https://doi.org/10.18632/oncotarget.13218
PMID:27835897

41. Filbin MR, Mehta A, Schneider AM, Kays KR, Guess JR, Gentili M, Fenyves BG, Charland NC, Gonye ALK, Gushterova I, Khanna HK, LaSalle TJ, Lavin-Parsons KM, et al. Longitudinal proteomic analysis of severe COVID-19 reveals survival-associated signatures, tissue-specific cell death, and cell-cell interactions. Cell Rep Med. 2021; 2:100287.
https://doi.org/10.1016/j.xcrm.2021.100287
PMID:33969320

42. Cao F, Lee RT. PLA2G7, caloric restriction and cardiovascular aging. J Cardiovasc Aging. 2022; 2:19.
https://doi.org/10.20517/jca.2022.08
PMID:35497092

43. Spadaro O, Youm Y, Shchukina I, Ryu S, Sidorov S, Ravussin A, Nguyen K, Aladyeva E, Predeus AN, Smith SR, Ravussin E, Galban C, Artyomov MN, Dixit VD. Caloric restriction in humans reveals immunometabolic regulators of health span. Science. 2022; 375:671–7.
https://doi.org/10.1126/science.abg7292
PMID:35143297

44. Aschenbrenner AC, Mouktaroudi M, Krämer B, Oestreich M, Antonakos N, Nuesch-Germano M, Gkizeli K, Bonaguro L, Reusch N, Baßler K, Saridaki M, Knoll R, Pecht T, et al, and German COVID-19 Omics

Initiative (DeCOI). Disease severity-specific neutrophil signatures in blood transcriptomes stratify COVID-19 patients. Genome Med. 2021; 13:7.
https://doi.org/10.1186/s13073-020-00823-5
PMID:33441124

45. Bonilla C, Lewis SJ, Rowlands MA, Gaunt TR, Davey Smith G, Gunnell D, Palmer T, Donovan JL, Hamdy FC, Neal DE, Eeles R, Easton D, Kote-Jarai Z, et al, and PRACTICAL consortium. Assessing the role of insulin-like growth factors and binding proteins in prostate cancer using Mendelian randomization: Genetic variants as instruments for circulating levels. Int J Cancer. 2016; 139:1520–33.
https://doi.org/10.1002/ijc.30206
PMID:27225428

46. Larsson SC, Michaëlsson K, Burgess S. IGF-1 and cardiometabolic diseases: a Mendelian randomisation study. Diabetologia. 2020; 63:1775–82.
https://doi.org/10.1007/s00125-020-05190-9
PMID:32548700

47. Ilnytska O, Argyropoulos G. The role of the Agouti-Related Protein in energy balance regulation. Cell Mol Life Sci. 2008; 65:2721–31.
https://doi.org/10.1007/s00018-008-8104-4
PMID:18470724

48. Redmann SM Jr, Argyropoulos G. AgRP-deficiency could lead to increased lifespan. Biochem Biophys Res Commun. 2006; 351:860–4.
https://doi.org/10.1016/j.bbrc.2006.10.129
PMID:17097059

49. Davies NM, Holmes MV, Davey Smith G. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. BMJ. 2018; 362:k601.
https://doi.org/10.1136/bmj.k601
PMID:30002074

SUPPLEMENTARY MATERIALS

Supplementary Methods

The proteomics data come from the UK Biobank Pharma Proteomics Project (UKB-PPP), compromising 54,306 samples (46,673 UKB participants at baseline visit, 6,385 samples selected by the UKB-PPP consortium, and 1,268 samples who participated in the COVID-19 repeat imaging study). The selection of UKB-PPP samples was based on two temporally and algorithmically separated picking processes. The first process involved pre-selecting 5,500 samples from participants' baseline recruitment visits, supplemented by additional 44,502 samples selected from the UK Biobank cohort with stratification against age, sex, and recruitment centre. The second process included a total of 7,000 samples, with 1,020 pre-selected samples, 3,637 samples from participants in the COVID-19 case-control imaging study, and 2,343 baseline samples selected randomly as described for the first sample selection. The full description of the selection criteria can be found in the Supplementary Material in Sun, et al. [1]. Generally, the 46,673 UKB participants were highly representative of the overall UKB population across various demographic characteristics except for deprivation index, distribution of recruitment centres, and time since recruitment. The 6,385 UKB-PPP consortium-selected participants were on average 2.5 years older than the overall UKB participants, and there were also differences in sex, ethnic background, body mass index, and smoking prevalence. The 1,268 COVID-19 imaging participants were on average 6.3 years younger than the overall UKB participants, while the other demographic characteristics were comparable to the UKB population.

Supplementary References

1. Sun BB, Chiou J, Traylor M, Benner C, Hsu YH, Richardson TG, Surendran P, Mahajan A, Robins C, Vasquez-Grinnell SG, Hou L, Kvikstad EM, Burren OS, et al. Genetic regulation of the human plasma proteome in 54,306 UK Biobank participants. bioRxiv. 2022.
   https://doi.org/10.1101/2022.06.17.496443

Supplementary Tables

Please browse Full Text version to see the data of Supplementary Tables 1–5 and 15–18.

Supplementary Table 1. Genetic instruments for proteins used in the Mendelian randomization analysis.

Supplementary Table 2. Associations of genetically predicted proteins with severe COVID-19 using IVW method and cis+trans SNPs.

Supplementary Table 3. Associations of genetically predicted proteins with COVID-19 hospitalization using IVW method and cis+trans SNPs.

Supplementary Table 4. Associations of genetically predicted proteins with COVID-19 infection using IVW method and cis+trans SNPs.

Supplementary Table 5. Function of proteins affecting COVID-19 risk in MR analysis.

Supplementary Table 6. Sensitivity analysis of genetically predicted proteins with severe COVID-19 using the weighted mode and weighted median methods.

Protein	No. of SNPs	BETA (Weighted Mode)	SE (Weighted Mode)	*P*-val (Weighted Mode)	BETA (Weighted Median)	SE (Weighted Median)	*P*-val (Weighted Median)
ADGRG2	6	0.15	0.15	3.36E-01	0.18	0.16	2.60E-01
AMY2B	9	−0.08	0.04	3.09E-02	−0.09	0.04	2.36E-02
CCL15	4	0.47	0.08	1.23E-09	0.39	0.08	1.85E-07
CD109	6	−0.49	0.08	5.78E-10	−0.48	0.08	1.31E-10
CD209	8	0.11	0.02	3.56E-11	0.11	0.02	1.28E-11
CD34	3	−0.24	0.04	5.26E-10	−0.23	0.04	2.46E-09
CDH15	4	0.31	0.05	6.57E-11	0.30	0.05	1.62E-10
CKMT1A_CKMT1B	6	0.10	0.20	6.16E-01	0.25	0.14	8.03E-02
CX3CL1	7	0.02	0.49	9.61E-01	−0.55	0.17	1.45E-03
ERBB4	16	−0.14	0.08	7.33E-02	−0.15	0.07	3.54E-02
FGF19	5	0.37	0.06	5.58E-09	0.31	0.06	8.35E-07
GOLM2	9	0.29	0.07	1.88E-05	0.17	0.08	3.01E-02
ICAM5	6	−0.09	0.02	3.58E-05	−0.09	0.02	7.01E-05
ISLR2	8	−0.16	0.03	4.74E-10	−0.17	0.03	9.15E-11
KEL	16	−0.13	0.04	4.93E-04	−0.08	0.05	9.28E-02
KLK1	13	−0.10	0.02	2.16E-09	−0.10	0.02	4.65E-10
LAMP3	10	−0.09	0.09	2.94E-01	−0.19	0.10	5.00E-02
LEFTY2	4	0.04	0.35	9.07E-01	−0.57	0.31	6.21E-02
LGALS4	3	−0.19	0.03	3.31E-11	−0.19	0.03	2.59E-11
LGALS8	6	−0.22	0.04	3.48E-07	−0.20	0.04	3.83E-06
MNDA	1	NA	NA	NA	NA	NA	NA
MUC13	4	−0.03	0.10	7.50E-01	−0.18	0.09	3.98E-02
NRCAM	6	−0.11	0.17	5.02E-01	−0.25	0.15	8.32E-02
PECAM1	8	−0.19	0.03	6.47E-10	−0.19	0.03	5.32E-10
PODXL	6	−0.20	0.03	1.59E-09	−0.19	0.03	1.68E-08
PTPRM	8	−0.16	0.03	1.18E-09	−0.16	0.03	8.58E-10
REG1A	5	0.13	0.06	1.95E-02	0.17	0.06	5.56E-03
REG1B	7	0.09	0.05	7.77E-02	0.15	0.06	1.69E-02
SCG2	4	−0.16	0.25	5.22E-01	−0.85	0.19	1.17E-05
SCGN	5	0.66	0.27	1.55E-02	0.23	0.13	7.21E-02

Protein	No. of SNPs	BETA (Weighted Mode)	SE (Weighted Mode)	*P*-val (Weighted Mode)	BETA (Weighted Median)	SE (Weighted Median)	*P*-val (Weighted Median)
SEMA4C	4	−0.43	0.07	6.90E-10	−0.43	0.07	1.95E-10
SFTPD	11	−0.08	0.02	1.04E-05	−0.09	0.02	7.16E-07
TDGF1	1	NA	NA	NA	NA	NA	NA
VAMP5	2	NA	NA	NA	NA	NA	NA
VTCN1	2	NA	NA	NA	NA	NA	NA

The results for proteins with number of SNPs less than 3 are filled with NA values.

Supplementary Table 7. Sensitivity analysis of genetically predicted proteins with COVID-19 hospitalization using the weighted mode and weighted median methods.

Protein	No. of SNPs	BETA (Weighted Mode)	SE (Weighted Mode)	*P*-val (Weighted Mode)	BETA (Weighted Median)	SE (Weighted Median)	*P*-val (Weighted Median)
ADGRG1	9	−0.04	0.08	6.66E-01	−0.13	0.07	7.43E-02
ADGRG2	6	0.11	0.09	2.52E-01	0.17	0.13	1.80E-01
AMY2A	8	−0.04	0.03	2.65E-01	−0.05	0.04	2.11E-01
AMY2B	9	−0.04	0.02	1.31E-01	−0.05	0.03	8.59E-02
CA4	9	0.04	0.03	1.67E-01	0.05	0.03	1.16E-01
CCL15	4	0.28	0.05	9.09E-08	0.26	0.05	2.61E-07
CD109	6	−0.42	0.05	1.28E-14	−0.40	0.06	2.91E-12
CD209	8	0.11	0.01	4.03E-21	0.11	0.01	5.57E-20
CD34	3	−0.22	0.03	4.84E-16	−0.19	0.03	1.23E-13
CDH15	4	0.17	0.03	1.06E-07	0.17	0.03	8.76E-08
CLEC14A	8	−0.27	0.04	4.35E-14	−0.22	0.04	2.22E-07
CTSS	7	−0.06	0.08	4.95E-01	−0.09	0.08	2.66E-01
CX3CL1	7	−0.02	0.22	9.46E-01	−0.26	0.12	3.27E-02
EFNA1	4	0.13	0.03	1.84E-05	0.13	0.03	1.37E-05
ERBB4	16	−0.03	0.05	5.63E-01	−0.06	0.05	2.04E-01
FGF19	5	0.24	0.04	3.28E-08	0.23	0.04	2.98E-08
GOLM2	9	0.11	0.08	2.12E-01	0.12	0.05	1.54E-02
GP2	8	0.16	0.02	2.67E-15	0.15	0.02	6.80E-10
ICAM2	10	−0.05	0.01	1.51E-06	−0.02	0.03	4.19E-01
ICAM5	7	−0.04	0.02	1.12E-02	−0.07	0.02	5.37E-05
ISLR2	8	−0.14	0.02	1.38E-15	−0.13	0.02	3.87E-13
KLK1	13	−0.06	0.01	8.04E-08	−0.06	0.01	5.47E-08
LEFTY2	4	−0.10	0.15	4.79E-01	−0.33	0.19	8.40E-02
LGALS4	3	−0.19	0.02	3.97E-22	−0.18	0.02	8.61E-22
MET	20	−0.26	0.04	8.78E-10	−0.21	0.05	1.49E-05
NELL2	14	−0.44	0.07	2.36E-09	−0.28	0.08	7.18E-04
NRCAM	7	0.00	0.08	9.67E-01	−0.04	0.09	6.38E-01
PECAM1	8	−0.16	0.02	1.35E-15	−0.15	0.02	1.37E-13
PODXL	6	−0.17	0.02	3.06E-14	−0.14	0.02	1.63E-09
PRSS27	18	−0.08	0.02	2.14E-07	−0.07	0.02	1.08E-05
PTPRM	8	−0.14	0.02	5.37E-16	−0.14	0.02	3.28E-16
REG1A	5	0.12	0.04	7.72E-04	0.17	0.04	2.68E-05
REG1B	7	0.09	0.03	1.15E-02	0.08	0.04	3.96E-02
SCG2	4	0.01	0.15	9.33E-01	−0.48	0.12	1.20E-04
SCGN	5	0.09	0.16	5.68E-01	0.13	0.07	7.15E-02
SELE	9	−0.10	0.01	3.55E-16	−0.10	0.01	5.32E-15
SEMA3F	5	−0.12	0.10	2.31E-01	−0.15	0.10	1.36E-01
SEMA4C	5	−0.36	0.05	1.03E-14	−0.34	0.05	1.18E-13
SFTPA2	6	0.07	0.02	1.61E-05	0.07	0.02	1.75E-05

Protein			
SFTPD	11	−0.07	0.01	6.42E-09	−0.07	0.01	1.70E-08
TDGF1	1	NA	NA	NA	NA	NA	NA
TGFBR2	8	−0.57	0.08	8.99E-12	−0.30	0.11	5.09E-03
VAMP5	2	NA	NA	NA	NA	NA	NA

The results for proteins with number of SNPs less than 3 are filled with NA values.

Supplementary Table 8. Sensitivity analysis of genetically predicted proteins with COVID-19 infection using the weighted mode and weighted median methods.

Protein	No. of SNPs	BETA (Weighted Mode)	SE (Weighted Mode)	*P*-val (Weighted Mode)	BETA (Weighted Median)	SE (Weighted Median)	*P*-val (Weighted Median)
ADAM15	8	0.04	0.01	1.50E-09	0.04	0.01	1.58E-09
ADGRG1	10	0.02	0.03	5.94E-01	−0.01	0.03	7.22E-01
ADGRG2	6	0.04	0.04	3.23E-01	0.07	0.05	1.30E-01
AMY2A	8	0.00	0.01	8.32E-01	−0.02	0.02	3.48E-01
AMY2B	9	−0.01	0.01	3.75E-01	−0.03	0.01	4.90E-02
BST2	11	−0.01	0.03	8.44E-01	0.00	0.03	9.41E-01
CA4	11	0.00	0.01	8.09E-01	0.00	0.01	7.70E-01
CCL15	4	0.19	0.03	1.32E-12	0.16	0.03	6.19E-09
CCL28	18	0.00	0.02	9.62E-01	−0.01	0.02	5.38E-01
CD109	6	−0.36	0.03	1.20E-37	−0.29	0.04	1.84E-12
CD209	8	0.10	0.01	5.45E-54	0.10	0.01	1.10E-48
CD34	3	−0.18	0.01	3.99E-42	−0.16	0.01	2.50E-38
CD58	12	−0.19	0.01	1.21E-40	−0.13	0.03	5.94E-06
CDH17	12	0.09	0.01	1.19E-14	0.02	0.03	5.69E-01
CKMT1A_CKMT1B	8	0.01	0.03	7.50E-01	0.02	0.03	5.08E-01
CLEC14A	9	−0.24	0.02	1.05E-37	−0.21	0.02	2.65E-18
CTSS	7	−0.06	0.04	8.33E-02	−0.07	0.04	7.34E-02
CX3CL1	8	−0.05	0.06	3.36E-01	−0.16	0.07	3.30E-02
DPP10	5	−0.01	0.02	4.96E-01	−0.02	0.02	2.59E-01
DRAXIN	17	−0.05	0.02	6.29E-03	−0.05	0.02	5.51E-03
EFNA1	4	0.10	0.01	8.92E-12	0.10	0.01	1.81E-11
F2R	8	−0.03	0.03	3.60E-01	−0.03	0.03	1.87E-01
FCGR2B	4	−0.01	0.01	4.36E-01	−0.01	0.02	5.55E-01
FGF19	5	0.16	0.02	4.81E-14	0.15	0.02	6.08E-15
FGFR2	8	0.00	0.04	9.14E-01	0.01	0.03	7.81E-01
FLT4	12	0.00	0.01	7.98E-01	0.00	0.01	5.96E-01
FOLR1	8	−0.06	0.05	2.88E-01	−0.13	0.05	5.44E-03
GKN1	4	0.10	0.11	3.50E-01	0.27	0.07	5.46E-05
GOLM2	8	0.03	0.02	1.24E-01	0.06	0.02	1.71E-02
ICAM2	10	−0.01	0.01	3.96E-01	0.00	0.02	9.41E-01
ICAM5	7	−0.01	0.01	1.37E-01	−0.01	0.01	1.77E-01
IDS	3	−0.01	0.02	7.23E-01	−0.02	0.02	3.63E-01
ISLR2	8	−0.13	0.01	6.41E-39	−0.11	0.01	2.62E-26
ITGA6	7	−0.03	0.02	1.43E-01	−0.03	0.02	8.71E-02
ITGB1	7	0.01	0.03	7.85E-01	0.00	0.03	9.52E-01
KEL	17	−0.01	0.01	6.63E-01	−0.01	0.01	4.39E-01
KLK1	13	−0.04	0.01	1.32E-14	−0.04	0.01	8.85E-14
LEFTY2	3	−0.02	0.06	7.72E-01	−0.24	0.09	5.42E-03
MNDA	1	NA	NA	NA	NA	NA	NA

Protein	No. of SNPs	BETA	SE	P-val	BETA	SE	P-val
MUC13	4	−0.03	0.02	2.23E-01	−0.15	0.05	1.88E-03
NELL2	15	0.04	0.06	4.54E-01	−0.20	0.05	3.72E-05
NME3	14	−0.02	0.02	2.30E-01	−0.02	0.02	1.39E-01
NRCAM	7	0.00	0.03	9.40E-01	−0.04	0.04	4.03E-01
PECAM1	8	−0.14	0.01	1.53E-43	−0.14	0.01	9.26E-37
PLAT	4	−0.05	0.05	3.13E-01	−0.11	0.05	4.46E-02
PODXL	6	−0.15	0.01	5.01E-29	−0.11	0.01	1.81E-16
PTPRM	8	−0.12	0.01	1.26E-44	−0.12	0.01	7.06E-46
REG1A	6	0.02	0.02	1.76E-01	0.04	0.02	1.13E-01
REG1B	7	0.03	0.02	3.16E-02	0.02	0.02	4.75E-01
S100A16	3	0.02	0.04	6.24E-01	0.03	0.05	5.49E-01
SCARF2	8	−0.03	0.02	6.28E-02	−0.03	0.02	3.68E-02
SCG2	4	−0.13	0.06	2.44E-02	−0.33	0.06	1.93E-08
SEMA3F	6	−0.01	0.06	8.04E-01	−0.03	0.05	5.54E-01
SEMA4C	6	−0.32	0.02	3.69E-38	−0.26	0.03	1.49E-21
SFTPD	12	−0.03	0.01	1.27E-06	−0.03	0.01	1.88E-05
SLITRK2	11	0.18	0.04	3.55E-06	0.09	0.03	4.29E-03
SPINK5	7	−0.02	0.01	2.06E-01	−0.02	0.02	2.14E-01
STC1	6	0.04	0.03	2.45E-01	0.01	0.04	6.90E-01
TDGF1	2	NA	NA	NA	NA	NA	NA
TGFBR2	8	−0.04	0.07	5.73E-01	−0.19	0.07	8.11E-03
ULBP2	8	0.01	0.01	5.83E-01	0.01	0.01	5.51E-01
VAMP5	2	NA	NA	NA	NA	NA	NA
VTCN1	2	NA	NA	NA	NA	NA	NA

The results for proteins with number of SNPs less than 3 are filled with NA values.

Supplementary Table 9. Sensitivity analysis of genetically predicted proteins with severe COVID-19 using MR-SPI.

Protein	No. of SNPs	No. of valid SNPs	BETA (MR-SPI)	SE (MR-SPI)	*P*-val (MR-SPI)
ADGRG2	6	5	0.25	0.13	5.34E-02
AMY2B	9	7	−0.06	0.04	7.58E-02
CCL15	4	4	0.34	0.07	1.74E-06
CD109	6	6	−0.35	0.08	4.11E-06
CD209	8	8	0.11	0.02	1.47E-10
CD34	3	3	−0.21	0.04	3.12E-07
CDH15	4	4	0.28	0.05	2.19E-09
CKMT1A_CKMT1B	6	5	0.10	0.11	3.74E-01
CX3CL1	7	6	−0.46	0.11	3.84E-05
ERBB4	16	15	−0.21	0.06	3.10E-04
FGF19	5	5	0.30	0.05	2.53E-08
GOLM2	9	8	0.22	0.05	4.70E-06
ICAM5	6	6	−0.09	0.02	1.31E-05
ISLR2	8	8	−0.13	0.03	5.72E-07
KEL	16	15	−0.10	0.03	3.83E-04
KLK1	13	11	−0.10	0.02	2.04E-10
LAMP3	10	6	−0.11	0.08	1.64E-01
LEFTY2	4	4	−0.16	0.21	4.40E-01
LGALS4	3	3	−0.19	0.03	5.43E-11
LGALS8	6	6	−0.22	0.03	5.08E-10
MUC13	4	4	−0.35	0.07	1.25E-07

NRCAM	6	5	−0.12	0.12	2.99E-01
PECAM1	8	8	−0.16	0.03	5.00E-08
PODXL	6	6	−0.13	0.03	2.34E-06
PTPRM	8	7	−0.15	0.03	5.94E-09
REG1A	5	5	0.53	0.06	2.30E-17
REG1B	7	7	0.09	0.04	2.55E-02
SCG2	4	4	−0.06	0.18	7.59E-01
SCGN	5	5	0.11	0.10	2.42E-01
SEMA4C	4	4	−0.37	0.07	7.94E-08
SFTPD	11	8	−0.09	0.02	5.49E-07

The results for proteins with number of SNPs less than 3 are omitted.

Supplementary Table 10. Sensitivity analysis of genetically predicted proteins with COVID-19 hospitalization using MR-SPI.

Protein	No. of SNPs	No. of valid SNPs	BETA (MR-SPI)	SE (MR-SPI)	*P*-val (MR-SPI)
ADGRG1	9	8	−0.04	0.05	4.80E-01
ADGRG2	6	4	0.16	0.08	5.14E-02
AMY2A	8	6	−0.03	0.03	2.51E-01
AMY2B	9	7	−0.03	0.02	1.45E-01
CA4	9	8	0.03	0.03	3.23E-01
CCL15	4	4	0.24	0.05	6.20E-07
CD109	6	6	−0.32	0.05	5.44E-11
CD209	8	8	0.10	0.01	6.80E-19
CD34	3	3	−0.17	0.03	3.61E-11
CDH15	4	4	0.16	0.03	1.98E-07
CLEC14A	8	8	0.09	0.06	1.83E-01
CTSS	7	6	−0.07	0.06	2.56E-01
CX3CL1	7	5	0.00	0.09	9.63E-01
EFNA1	4	4	0.13	0.03	4.10E-06
ERBB4	16	14	−0.04	0.03	2.56E-01
FGF19	5	5	0.18	0.04	2.13E-07
GOLM2	9	7	0.06	0.04	9.16E-02
GP2	8	7	0.17	0.02	1.91E-21
ICAM2	10	9	−0.10	0.01	1.85E-12
ICAM5	7	7	−0.20	0.03	3.39E-13
ISLR2	8	7	0.02	0.04	6.94E-01
KLK1	13	13	−0.05	0.01	1.20E-06
LEFTY2	4	4	−0.15	0.12	2.11E-01
LGALS4	3	3	−0.18	0.02	6.56E-21
MET	20	18	0.01	0.03	6.78E-01
NELL2	14	11	−0.03	0.06	6.35E-01
NRCAM	7	6	−0.03	0.07	6.39E-01
PECAM1	8	8	−0.14	0.02	1.26E-12
PODXL	6	6	−0.12	0.02	9.56E-11
PRSS27	18	17	−0.08	0.02	9.75E-08
PTPRM	8	7	−0.13	0.02	9.40E-15
REG1A	5	4	0.14	0.03	3.03E-07
REG1B	7	5	0.06	0.03	4.02E-02
SCG2	4	4	0.15	0.12	2.32E-01
SCGN	5	5	0.20	0.05	4.09E-05

SELE	9	8	−0.09	0.01	3.79E-14
SEMA3F	5	4	−0.07	0.08	3.57E-01
SEMA4C	5	5	−0.25	0.04	2.42E-08
SFTPA2	6	6	0.07	0.02	1.83E-05
SFTPD	11	9	−0.07	0.01	3.26E-09
TGFBR2	8	7	−0.05	0.07	4.95E-01

The results for proteins with number of SNPs less than 3 are omitted.

Supplementary Table 11. Sensitivity analysis of genetically predicted proteins with COVID-19 infection using MR-SPI.

Protein	No. of SNPs	No. of valid SNPs	BETA (MR-SPI)	SE (MR-SPI)	*P*-val (MR-SPI)
ADAM15	8	7	0.03	0.01	1.57E-09
ADGRG1	10	9	−0.02	0.02	3.00E-01
ADGRG2	6	4	0.04	0.04	2.76E-01
AMY2A	8	6	0.00	0.01	9.70E-01
AMY2B	9	7	0.00	0.01	8.96E-01
BST2	11	9	−0.02	0.02	3.13E-01
CA4	11	10	0.00	0.01	7.67E-01
CCL15	4	3	−0.10	0.05	4.49E-02
CCL28	18	17	−0.01	0.02	3.55E-01
CD109	6	5	−0.01	0.04	8.30E-01
CD209	8	8	0.11	0.01	8.13E-76
CD34	3	2	0.02	0.04	5.63E-01
CD58	12	11	−0.02	0.01	2.66E-01
CDH17	12	11	0.08	0.01	4.59E-15
CKMT1A_CKMT1B	8	6	0.02	0.03	5.17E-01
CLEC14A	9	7	−0.01	0.03	6.32E-01
CTSS	7	6	−0.11	0.03	1.73E-04
CX3CL1	8	6	0.03	0.04	3.79E-01
DPP10	5	4	−0.01	0.02	5.21E-01
DRAXIN	17	16	−0.04	0.01	2.21E-03
EFNA1	4	4	0.09	0.01	1.01E-10
F2R	8	7	−0.02	0.02	3.01E-01
FCGR2B	4	3	−0.01	0.01	6.02E-01
FGF19	5	5	0.13	0.02	1.22E-15
FGFR2	8	7	0.01	0.03	7.29E-01
FLT4	12	11	0.00	0.01	6.78E-01
FOLR1	8	5	−0.05	0.03	1.18E-01
GKN1	4	4	0.47	0.06	6.15E-16
GOLM2	8	6	0.02	0.02	2.22E-01
ICAM2	10	8	0.00	0.01	6.79E-01
ICAM5	7	6	−0.01	0.01	8.60E-02
IDS	3	2	−0.01	0.02	5.18E-01
ISLR2	8	6	−0.02	0.02	2.58E-01
ITGA6	7	6	−0.03	0.02	5.95E-02
ITGB1	7	6	0.00	0.02	8.30E-01
KEL	17	14	0.00	0.01	6.26E-01
KLK1	13	12	0.01	0.02	5.11E-01
LEFTY2	3	2	−0.03	0.06	6.71E-01
MUC13	4	3	−0.02	0.02	2.47E-01

NELL2	15	13	−0.03	0.02	1.85E-01
NME3	14	12	−0.01	0.01	5.66E-01
NRCAM	7	6	−0.04	0.03	1.44E-01
PECAM1	8	8	−0.01	0.02	8.32E-01
PLAT	4	4	−0.24	0.05	1.62E-06
PODXL	6	5	−0.02	0.01	1.59E-01
PTPRM	8	7	−0.11	0.01	4.97E-43
REG1A	6	4	0.03	0.01	4.30E-02
REG1B	7	5	0.01	0.01	2.62E-01
S100A16	3	2	0.02	0.04	6.09E-01
SCARF2	8	7	−0.03	0.01	3.07E-02
SCG2	4	2	−0.10	0.06	8.04E-02
SEMA3F	6	5	−0.01	0.04	8.91E-01
SEMA4C	6	5	−0.02	0.04	6.01E-01
SFTPD	12	10	−0.03	0.01	5.31E-06
SLITRK2	11	11	0.09	0.02	1.23E-05
SPINK5	7	6	−0.02	0.01	1.34E-01
STC1	6	5	0.03	0.03	3.91E-01
TGFBR2	8	7	−0.02	0.03	4.77E-01
ULBP2	8	7	0.00	0.01	8.05E-01

The results for proteins with number of SNPs less than 3 are omitted.

Supplementary Table 12. Sensitivity analysis of genetically predicted proteins with severe COVID-19 using the cis-SNPs only.

Protein	BETA (cis SNPs)	SE (cis SNPs)	P-val (cis SNPs)
ADGRG2	NA	NA	NA
AMY2B	−0.08	0.04	2.32E-02
CCL15	NA	NA	NA
CD109	NA	NA	NA
CD209	0.08	0.04	6.45E-02
CD34	NA	NA	NA
CDH15	0.31	0.05	4.55E-11
CKMT1A_CKMT1B	NA	NA	NA
CX3CL1	NA	NA	NA
ERBB4	−0.15	0.08	7.59E-02
FGF19	NA	NA	NA
GOLM2	NA	NA	NA
ICAM5	−0.09	0.02	1.80E-05
ISLR2	−0.08	0.07	2.71E-01
KEL	−0.04	0.06	4.55E-01
KLK1	−0.10	0.02	2.42E-10
LAMP3	−0.09	0.10	3.63E-01
LEFTY2	NA	NA	NA
LGALS4	−0.31	0.22	1.66E-01
LGALS8	−0.01	0.05	8.01E-01
MNDA	NA	NA	NA
MUC13	−0.02	0.07	7.64E-01
NRCAM	NA	NA	NA
PECAM1	NA	NA	NA
PODXL	−0.01	0.06	8.97E-01

PTPRM	−0.37	0.22	8.93E-02
REG1A	0.10	0.06	8.88E-02
REG1B	0.08	0.04	8.88E-02
SCG2	NA	NA	NA
SCGN	0.18	0.17	2.85E-01
SEMA4C	NA	NA	NA
SFTPD	−0.09	0.02	8.14E-07
TDGF1	NA	NA	NA
VAMP5	0.16	0.20	4.24E-01
VTCN1	−0.02	0.28	9.56E-01

The results for proteins without cis-SNPs are filled with NA values.

Supplementary Table 13. Sensitivity analysis of genetically predicted proteins with COVID-19 hospitalization using the cis-SNPs only.

Protein	BETA (cis SNPs)	SE (cis SNPs)	*P*-val (cis SNPs)
ADGRG1	−0.31	0.15	3.57E-02
ADGRG2	NA	NA	NA
AMY2A	−0.04	0.03	1.95E-01
AMY2B	−0.04	0.02	8.00E-02
CA4	0.04	0.03	1.68E-01
CCL15	NA	NA	NA
CD109	NA	NA	NA
CD209	0.05	0.03	1.33E-01
CD34	NA	NA	NA
CDH15	0.17	0.03	4.43E-08
CLEC14A	NA	NA	NA
CTSS	NA	NA	NA
CX3CL1	NA	NA	NA
EFNA1	0.13	0.03	1.55E-05
ERBB4	−0.05	0.06	3.52E-01
FGF19	NA	NA	NA
GOLM2	NA	NA	NA
GP2	0.00	0.04	9.30E-01
ICAM2	NA	NA	NA
ICAM5	−0.04	0.01	2.52E-03
ISLR2	−0.05	0.05	2.78E-01
KLK1	−0.06	0.01	2.65E-08
LEFTY2	NA	NA	NA
LGALS4	−0.05	0.12	6.60E-01
MET	−0.02	0.08	8.05E-01
NELL2	NA	NA	NA
NRCAM	NA	NA	NA
PECAM1	NA	NA	NA
PODXL	−0.05	0.04	2.45E-01
PRSS27	−0.03	0.06	6.77E-01
PTPRM	−0.25	0.15	1.09E-01
REG1A	0.08	0.04	5.68E-02
REG1B	0.06	0.03	5.68E-02
SCG2	NA	NA	NA
SCGN	0.11	0.12	3.38E-01

Protein			
SELE	−0.03	0.12	7.85E-01
SEMA3F	NA	NA	NA
SEMA4C	NA	NA	NA
SFTPA2	0.07	0.02	1.30E-05
SFTPD	−0.07	0.01	1.40E-07
TDGF1	NA	NA	NA
TGFBR2	NA	NA	NA
VAMP5	0.14	0.14	3.37E-01

The results for proteins without cis-SNPs are filled with NA values.

Supplementary Table 14. Sensitivity analysis of genetically predicted proteins with COVID-19 infection using the cis-SNPs only.

Protein	BETA (cis SNPs)	SE (cis SNPs)	*P*-val (cis SNPs)
ADAM15	0.03	0.01	4.44E-08
ADGRG1	−0.03	0.07	6.93E-01
ADGRG2	NA	NA	NA
AMY2A	−0.01	0.01	6.83E-01
AMY2B	−0.01	0.01	3.94E-01
BST2	0.00	0.03	9.13E-01
CA4	0.00	0.01	8.81E-01
CCL15	NA	NA	NA
CCL28	NA	NA	NA
CD109	NA	NA	NA
CD209	0.02	0.01	8.28E-02
CD34	NA	NA	NA
CD58	−0.02	0.03	4.40E-01
CDH17	NA	NA	NA
CKMT1A_CKMT1B	0.02	0.05	6.82E-01
CLEC14A	NA	NA	NA
CTSS	NA	NA	NA
CX3CL1	NA	NA	NA
DPP10	−0.02	0.02	3.90E-01
DRAXIN	−0.05	0.02	9.96E-03
EFNA1	0.10	0.01	6.46E-12
F2R	−0.01	0.04	8.88E-01
FCGR2B	−0.01	0.01	4.67E-01
FGF19	NA	NA	NA
FGFR2	0.01	0.04	7.20E-01
FLT4	−0.02	0.02	2.91E-01
FOLR1	−0.12	0.08	1.27E-01
GKN1	NA	NA	NA
GOLM2	NA	NA	NA
ICAM2	NA	NA	NA
ICAM5	−0.01	0.01	1.39E-01
IDS	NA	NA	NA
ISLR2	−0.04	0.02	1.05E-01
ITGA6	−0.03	0.02	1.29E-01
ITGB1	NA	NA	NA
KEL	0.00	0.02	8.64E-01
KLK1	−0.04	0.01	1.21E-14

LEFTY2	NA	NA	NA
MNDA	NA	NA	NA
MUC13	−0.03	0.02	1.11E-01
NELL2	NA	NA	NA
NME3	−0.03	0.03	3.89E-01
NRCAM	NA	NA	NA
PECAM1	NA	NA	NA
PLAT	−0.03	0.06	6.39E-01
PODXL	−0.04	0.02	3.86E-02
PTPRM	−0.03	0.07	6.20E-01
REG1A	0.01	0.02	4.95E-01
REG1B	0.01	0.01	4.95E-01
S100A16	0.02	0.05	7.45E-01
SCARF2	−0.03	0.02	4.90E-02
SCG2	NA	NA	NA
SEMA3F	NA	NA	NA
SEMA4C	NA	NA	NA
SFTPD	−0.03	0.01	5.36E-07
SLITRK2	NA	NA	NA
SPINK5	−0.02	0.01	1.93E-01
STC1	−0.07	0.08	3.52E-01
TDGF1	NA	NA	NA
TGFBR2	NA	NA	NA
ULBP2	0.01	0.01	6.55E-01
VAMP5	0.15	0.08	5.59E-02
VTCN1	−0.11	0.06	9.55E-02

The results for proteins without cis-SNPs are filled with NA values.

Supplementary Table 15. Associations of genetically predicted proteins with healthspan using IVW method and cis+trans SNPs.

Supplementary Table 16. Associations of genetically predicted proteins with father's attained age using IVW method and cis+trans SNPs.

Supplementary Table 17. Associations of genetically predicted proteins with mother's attained age using IVW method and cis+trans SNPs.

Supplementary Table 18. Function of proteins affecting healthspan or lifespan in MR analysis.

Supplementary Table 19. Sensitivity analysis of genetically predicted proteins with healthspan using the weighted mode and weighted median methods.

Protein	No. of SNPs	BETA (Weighted Mode)	SE (Weighted Mode)	*P*-val (Weighted Mode)	BETA (Weighted Median)	SE (Weighted Median)	*P*-val (Weighted Median)
FOXO3	1	NA	NA	NA	NA	NA	NA
GPNMB	6	0.43	0.14	2.59E-03	0.44	0.14	1.96E-03
HLA-DRA	3	−0.41	0.09	4.89E-06	−0.42	0.09	3.67E-06
PLA2G7	8	−0.30	0.67	6.51E-01	−1.26	0.40	1.80E-03

The results for proteins with number of SNPs less than 3 are filled with NA values.

Supplementary Table 20. Sensitivity analysis of genetically predicted proteins with father's attained age using the weighted mode and weighted median methods.

Protein	No. of SNPs	BETA (Weighted Mode)	SE (Weighted Mode)	*P*-val (Weighted Mode)	BETA (Weighted Median)	SE (Weighted Median)	*P*-val (Weighted Median)
AGRP	9	−0.74	0.41	7.16E-02	−0.83	0.35	1.83E-02
CA11	2	NA	NA	NA	NA	NA	NA
CCN1	11	−0.34	0.38	3.59E-01	−0.40	0.31	2.01E-01
CD27	7	−0.46	0.18	1.09E-02	−0.54	0.18	1.94E-03
CD74	6	−0.29	0.23	1.98E-01	−0.32	0.23	1.54E-01
CDH1	6	0.34	0.09	7.10E-05	0.35	0.09	6.90E-05
CEACAM21	2	NA	NA	NA	NA	NA	NA
CPE	6	1.32	0.42	1.54E-03	1.35	0.37	2.83E-04
CXCL13	7	−0.40	0.36	2.73E-01	−0.75	0.40	6.28E-02
CXCL9	5	−0.46	0.32	1.52E-01	−1.09	0.36	2.13E-03
F3	8	0.47	0.12	1.01E-04	0.49	0.12	8.17E-05
FASLG	18	−0.20	0.15	1.87E-01	−0.25	0.15	7.99E-02
FES	1	NA	NA	NA	NA	NA	NA
FURIN	3	−1.32	0.24	7.34E-08	−1.03	0.22	3.57E-06
GCNT1	4	−2.15	0.33	1.27E-10	−1.81	0.39	2.73E-06
GP1BA	14	−0.03	0.30	9.16E-01	−0.12	0.24	6.10E-01
GRN	10	−0.41	0.07	2.41E-08	−0.38	0.07	2.35E-07
GZMB	8	−0.56	0.29	5.12E-02	−0.58	0.27	3.36E-02
IGFBP1	2	NA	NA	NA	NA	NA	NA
KIR2DL3	2	NA	NA	NA	NA	NA	NA
LAIR1	3	−0.72	0.71	3.13E-01	−1.63	0.82	4.61E-02
LDLR	15	−1.00	0.26	9.71E-05	−0.96	0.22	1.46E-05
LEFTY2	3	1.74	0.58	2.96E-03	2.21	0.73	2.39E-03
LGALS9	3	−1.43	1.03	1.67E-01	−1.77	0.56	1.66E-03
LILRB4	2	NA	NA	NA	NA	NA	NA
PAG1	6	−0.56	0.43	1.95E-01	−0.63	0.40	1.19E-01
PCSK9	8	−0.48	0.14	6.45E-04	−0.49	0.14	4.77E-04
POLR2F	1	NA	NA	NA	NA	NA	NA
RP2	2	NA	NA	NA	NA	NA	NA
SIT1	6	−0.30	0.64	6.39E-01	−0.75	0.40	5.84E-02
VCAM1	10	−0.42	0.82	6.08E-01	−0.76	0.30	1.22E-02
VSTM2L	2	NA	NA	NA	NA	NA	NA

The results for proteins with number of SNPs less than 3 are filled with NA values.

Supplementary Table 21. Sensitivity analysis of genetically predicted proteins with mother's attained age using the weighted mode and weighted median methods.

Protein	No. of SNPs	BETA (Weighted Mode)	SE (Weighted Mode)	*P*-val (Weighted Mode)	BETA (Weighted Median)	SE (Weighted Median)	*P*-val (Weighted Median)
CDH1	6	0.50	0.09	1.35E-08	0.50	0.09	2.03E-08
CDH17	10	0.36	0.08	1.05E-05	0.31	0.10	2.07E-03
CDHR2	6	−0.17	0.49	7.36E-01	0.05	0.48	9.22E-01
CPE	6	0.54	0.42	1.98E-01	0.67	0.38	8.17E-02
CXADR	5	1.06	0.23	4.07E-06	1.10	0.23	1.41E-06
CXCL9	5	−0.32	0.45	4.75E-01	−0.90	0.36	1.20E-02

F3	8	0.69	0.12	1.71E-08	0.71	0.13	2.46E-08
FOXO3	1	NA	NA	NA	NA	NA	NA
GFAP	7	−0.06	0.38	8.85E-01	−0.21	0.32	5.16E-01
IL19	3	−0.41	0.11	1.11E-04	−0.40	0.10	3.29E-05
ITGB6	9	0.52	0.16	1.47E-03	0.53	0.16	6.94E-04
LAIR1	3	−1.21	1.06	2.53E-01	−2.01	0.88	2.32E-02
LEFTY2	3	2.09	0.77	6.45E-03	2.54	0.57	1.03E-05
LGALS9	3	−2.64	0.51	1.99E-07	−2.46	0.49	4.59E-07
LILRB4	2	NA	NA	NA	NA	NA	NA
POLR2F	1	NA	NA	NA	NA	NA	NA
RP2	2	NA	NA	NA	NA	NA	NA
STC2	4	−0.92	0.65	1.55E-01	−1.13	0.56	4.34E-02
TNFRSF8	7	−0.50	0.22	2.03E-02	−0.56	0.22	9.55E-03

The results for proteins with number of SNPs less than 3 are filled with NA values.

Supplementary Table 22. Sensitivity analysis of genetically predicted proteins with healthspan using MR-SPI.

Protein	No. of SNPs	No. of valid SNPs	BETA (MR-SPI)	SE (MR-SPI)	*P*-val (MR-SPI)
GPNMB	6	5	0.31	0.13	2.22E-02
HLA-DRA	3	3	−0.42	0.09	3.16E-06
PLA2G7	8	8	−1.20	0.27	8.80E-06

The results for proteins with number of SNPs less than 3 are omitted.

Supplementary Table 23. Sensitivity analysis of genetically predicted proteins with father's attained age using MR-SPI.

Protein	No. of SNPs	No. of valid SNPs	BETA (MR-SPI)	SE (MR-SPI)	*P*-val (MR-SPI)
AGRP	9	7	−0.93	0.26	2.94E-04
CCN1	11	8	−0.49	0.22	2.49E-02
CD27	7	6	−0.49	0.17	2.92E-03
CD74	6	5	−0.18	0.19	3.51E-01
CDH1	6	6	0.37	0.08	8.32E-06
CPE	6	5	0.91	0.28	1.33E-03
CXCL13	7	6	−0.94	0.28	8.23E-04
CXCL9	5	4	−0.67	0.26	9.59E-03
F3	8	8	0.43	0.11	3.54E-05
FASLG	18	17	−0.17	0.11	1.31E-01
FURIN	3	3	−1.05	0.21	3.14E-07
GCNT1	4	4	−2.19	0.32	5.14E-12
GP1BA	14	13	−0.24	0.17	1.63E-01
GRN	10	10	−0.36	0.07	3.06E-08
GZMB	8	7	−0.56	0.22	1.02E-02
LAIR1	3	2	−0.71	0.64	2.69E-01
LDLR	15	13	−1.21	0.16	2.68E-14
LEFTY2	3	3	4.15	0.65	1.64E-10
LGALS9	3	3	−2.82	0.42	2.88E-11
PAG1	6	5	−0.45	0.29	1.16E-01
PCSK9	8	7	−0.54	0.13	3.34E-05
SIT1	6	5	−0.10	0.29	7.17E-01

VCAM1	10	8	−0.05	0.19	8.14E-01

The results for proteins with number of SNPs less than 3 are omitted.

Supplementary Table 24. Sensitivity analysis of genetically predicted proteins with mother's attained age using MR-SPI.

Protein	No. of SNPs	No. of valid SNPs	BETA (MR-SPI)	SE (MR-SPI)	P-val (MR-SPI)
CDH1	6	6	0.53	0.08	3.84E-10
CDH17	10	10	0.37	0.07	3.96E-07
CDHR2	6	5	−0.05	0.36	8.89E-01
CPE	6	6	1.23	0.26	2.27E-06
CXADR	5	5	0.79	0.18	9.95E-06
CXCL9	5	5	−1.65	0.29	8.56E-09
F3	8	8	0.64	0.11	1.60E-09
GFAP	7	6	−0.23	0.25	3.43E-01
IL19	3	3	−0.39	0.09	2.37E-05
ITGB6	9	8	0.51	0.14	2.86E-04
LAIR1	3	3	−2.28	0.56	4.09E-05
LEFTY2	3	3	2.65	0.46	7.68E-09
LGALS9	3	2	−2.65	0.42	2.34E-10
STC2	4	4	−1.66	0.42	6.53E-05
TNFRSF8	7	7	−0.33	0.18	6.78E-02

The results for proteins with number of SNPs less than 3 are omitted.

Supplementary Table 25. Sensitivity analysis of genetically predicted proteins with healthspan using the cis-SNPs only.

Protein	BETA (cis SNPs)	SE (cis SNPs)	P-val (cis SNPs)
FOXO3	NA	NA	NA
GPNMB	0.47	0.15	1.28E-03
HLA-DRA	−0.40	0.09	9.04E-06
PLA2G7	NA	NA	NA

The results for proteins without cis-SNPs are filled with NA values.

Supplementary Table 26. Sensitivity analysis of genetically predicted proteins with father's attained age using the cis-SNPs only.

Protein	BETA (cis SNPs)	SE (cis SNPs)	P-val (cis SNPs)
AGRP	−0.07	0.60	9.01E-01
CA11	NA	NA	NA
CCN1	−0.16	0.42	7.08E-01
CD27	−0.52	0.18	3.53E-03
CD74	−0.23	0.81	7.80E-01
CDH1	NA	NA	NA
CEACAM21	NA	NA	NA
CPE	NA	NA	NA
CXCL13	−0.80	0.86	3.53E-01
CXCL9	−0.50	0.35	1.55E-01
F3	0.28	0.24	2.59E-01

Protein	BETA (cis SNPs)	SE (cis SNPs)	P-val (cis SNPs)
FASLG	−0.99	0.60	9.87E-02
FES	1.78	0.33	4.80E-08
FURIN	−1.32	0.24	2.69E-08
GCNT1	NA	NA	NA
GP1BA	NA	NA	NA
GRN	−0.19	0.19	3.19E-01
GZMB	NA	NA	NA
IGFBP1	NA	NA	NA
KIR2DL3	NA	NA	NA
LAIR1	NA	NA	NA
LDLR	NA	NA	NA
LEFTY2	NA	NA	NA
LGALS9	NA	NA	NA
LILRB4	NA	NA	NA
PAG1	−0.64	0.61	2.91E-01
PCSK9	−0.44	0.14	2.25E-03
POLR2F	NA	NA	NA
RP2	NA	NA	NA
SIT1	NA	NA	NA
VCAM1	NA	NA	NA
VSTM2L	−0.04	0.39	9.23E-01

The results for proteins without cis-SNPs are filled with NA values.

Supplementary Table 27. Sensitivity analysis of genetically predicted proteins with mother's attained age using the cis-SNPs only.

Protein	BETA (cis SNPs)	SE (cis SNPs)	*P*-val (cis SNPs)
CDH1	NA	NA	NA
CDH17	NA	NA	NA
CDHR2	NA	NA	NA
CPE	NA	NA	NA
CXADR	0.85	0.38	2.54E-02
CXCL9	−0.07	0.35	8.32E-01
F3	0.30	0.24	2.22E-01
FOXO3	NA	NA	NA
GFAP	−0.16	0.54	7.67E-01
IL19	−0.18	0.15	2.31E-01
ITGB6	0.49	0.16	3.17E-03
LAIR1	NA	NA	NA
LEFTY2	NA	NA	NA
LGALS9	NA	NA	NA
LILRB4	NA	NA	NA
POLR2F	NA	NA	NA
RP2	NA	NA	NA
STC2	NA	NA	NA
TNFRSF8	−0.36	0.21	9.29E-02

The results for proteins without cis-SNPs are filled with NA values.

Supplementary Table 28. Heterogeneity test of protein on severe COVID.

Protein	No. of SNPs	Q statistic	P-val of Q statistic
ADGRG2	6	42.14	5.51E-08
AMY2B	9	72.35	1.67E-12
CCL15	4	11.65	8.70E-03
CD109	6	6.06	3.01E-01
CD209	8	11.84	1.06E-01
CD34	3	3.65	1.61E-01
CDH15	4	2.48	4.79E-01
CKMT1A_CKMT1B	6	26.02	8.86E-05
CX3CL1	7	43.16	1.09E-07
ERBB4	16	44.94	7.82E-05
FGF19	5	11.67	1.99E-02
GOLM2	9	62.23	1.70E-10
ICAM5	6	16.57	5.38E-03
ISLR2	8	23.48	1.40E-03
KEL	16	72.11	1.88E-09
KLK1	13	42.97	2.28E-05
LAMP3	10	144.56	1.18E-26
LEFTY2	4	22.69	4.69E-05
LGALS4	3	5.58	6.14E-02
LGALS8	6	15.10	9.96E-03
MNDA	1	NA	NA
MUC13	4	24.43	2.03E-05
NRCAM	6	27.86	3.88E-05
PECAM1	8	13.05	7.08E-02
PODXL	6	14.95	1.06E-02
PTPRM	8	30.44	7.87E-05
REG1A	5	35.36	3.91E-07
REG1B	7	49.90	4.93E-09
SCG2	4	31.72	6.01E-07
SCGN	5	16.97	1.96E-03
SEMA4C	4	3.36	3.40E-01
SFTPD	11	98.80	9.47E-17
TDGF1	1	NA	NA
VAMP5	2	NA	NA
VTCN1	2	NA	NA

Supplementary Table 29. Heterogeneity test of protein on COVID hospitalization.

Protein	No. of SNPs	Q statistic	P-val of Q statistic
ADGRG1	9	57.47	1.46E-09
ADGRG2	6	92.04	2.50E-18
AMY2A	8	95.88	7.64E-18
AMY2B	9	99.13	6.44E-18
CA4	9	77.39	1.64E-13
CCL15	4	4.34	2.27E-01
CD109	6	23.11	3.22E-04
CD209	8	23.60	1.34E-03
CD34	3	11.84	2.68E-03
CDH15	4	6.17	1.04E-01

Protein	No. of SNPs	Q statistic	P-val of Q statistic
CLEC14A	8	73.72	2.60E-13
CTSS	7	84.18	4.89E-16
CX3CL1	7	47.96	1.20E-08
EFNA1	4	2.57	4.63E-01
ERBB4	16	79.76	7.72E-11
FGF19	5	13.19	1.04E-02
GOLM2	9	89.74	5.25E-16
GP2	8	86.98	5.15E-16
ICAM2	10	70.56	1.18E-11
ICAM5	7	42.99	1.17E-07
ISLR2	8	41.26	7.22E-07
KLK1	13	23.58	2.32E-02
LEFTY2	4	28.70	2.59E-06
LGALS4	3	7.15	2.81E-02
MET	20	77.14	5.75E-09
NELL2	14	89.19	2.00E-13
NRCAM	7	53.77	8.22E-10
PECAM1	8	26.75	3.69E-04
PODXL	6	19.27	1.71E-03
PRSS27	18	100.56	6.99E-14
PTPRM	8	44.44	1.75E-07
REG1A	5	69.14	3.46E-14
REG1B	7	85.00	3.30E-16
SCG2	4	55.74	4.78E-12
SCGN	5	13.73	8.21E-03
SELE	9	52.18	1.55E-08
SEMA3F	5	47.19	1.39E-09
SEMA4C	5	18.87	8.34E-04
SFTPA2	6	5.05	4.10E-01
SFTPD	11	55.32	2.75E-08
TDGF1	1	NA	NA
TGFBR2	8	33.02	2.63E-05
VAMP5	2	NA	NA

Supplementary Table 30. Heterogeneity test of protein on SARS-CoV-2 infection.

Protein	No. of SNPs	Q statistic	P-val of Q statistic
ADAM15	8	17.87	1.26E-02
ADGRG1	10	180.92	3.25E-34
ADGRG2	6	260.26	3.46E-54
AMY2A	8	331.37	1.19E-67
AMY2B	9	342.70	3.28E-69
BST2	11	351.39	2.02E-69
CA4	11	306.13	7.87E-60
CCL15	4	30.43	1.12E-06
CCL28	18	192.87	7.72E-32
CD109	6	70.21	9.24E-14
CD209	8	40.82	8.77E-07
CD34	3	28.96	5.16E-07
CD58	12	86.67	7.46E-14
CDH17	12	391.15	4.63E-77

CKMT1A_CKMT1B	8	161.54	1.52E-31
CLEC14A	9	126.33	1.63E-23
CTSS	7	275.64	1.34E-56
CX3CL1	8	158.83	5.66E-31
DPP10	5	190.95	3.31E-40
DRAXIN	17	68.97	1.51E-08
EFNA1	4	7.70	5.27E-02
F2R	8	185.14	1.60E-36
FCGR2B	4	310.89	4.36E-67
FGF19	5	12.76	1.25E-02
FGFR2	8	328.93	3.98E-67
FLT4	12	203.99	1.10E-37
FOLR1	8	109.09	1.42E-20
GKN1	4	49.76	8.98E-11
GOLM2	8	238.79	6.71E-48
ICAM2	10	166.16	3.89E-31
ICAM5	7	161.01	3.62E-32
IDS	3	333.48	3.86E-73
ISLR2	8	59.19	2.19E-10
ITGA6	7	176.19	2.18E-35
ITGB1	7	184.40	3.93E-37
KEL	17	223.16	1.59E-38
KLK1	13	67.09	1.12E-09
LEFTY2	3	109.90	1.37E-24
MNDA	1	NA	NA
MUC13	4	109.02	1.78E-23
NELL2	15	178.72	1.18E-30
NME3	14	227.26	3.30E-41
NRCAM	7	165.94	3.27E-33
PECAM1	8	57.85	4.04E-10
PLAT	4	16.33	9.69E-04
PODXL	6	68.83	1.80E-13
PTPRM	8	72.39	4.86E-13
REG1A	6	291.32	7.34E-61
REG1B	7	326.00	2.18E-67
S100A16	3	38.70	3.96E-09
SCARF2	8	180.12	1.84E-35
SCG2	4	117.26	3.00E-25
SEMA3F	6	168.95	1.22E-34
SEMA4C	6	45.21	1.31E-08
SFTPD	12	65.80	7.62E-10
SLITRK2	11	22.54	1.26E-02
SPINK5	7	40.67	3.37E-07
STC1	6	197.29	1.08E-40
TDGF1	2	NA	NA
TGFBR2	8	108.75	1.67E-20
ULBP2	8	327.41	8.38E-67
VAMP5	2	NA	NA
VTCN1	2	NA	NA

Supplementary Table 31. Heterogeneity test of protein on healthspan.

Protein	No. of SNPs	Q statistic	P-val of Q statistic
FOXO3	1	NA	NA
GPNMB	6	113.95	5.99E-23
HLA-DRA	3	1.29	5.24E-01
PLA2G7	8	15.97	2.54E-02

Supplementary Table 32. Heterogeneity test of protein on father's attained age.

Protein	No. of SNPs	Q statistic	P-val of Q statistic
AGRP	9	68.91	8.08E-12
CA11	2	NA	NA
CCN1	11	115.51	4.10E-20
CD27	7	62.53	1.38E-11
CD74	6	36.35	8.08E-07
CDH1	6	11.08	4.98E-02
CEACAM21	2	NA	NA
CPE	6	30.68	1.08E-05
CXCL13	7	54.56	5.70E-10
CXCL9	5	25.75	3.55E-05
F3	8	4.36	7.37E-01
FASLG	18	53.29	1.29E-05
FES	1	NA	NA
FURIN	3	4.72	9.44E-02
GCNT1	4	48.76	1.46E-10
GP1BA	14	41.85	6.94E-05
GRN	10	15.78	7.16E-02
GZMB	8	39.66	1.46E-06
IGFBP1	2	NA	NA
KIR2DL3	2	NA	NA
LAIR1	3	25.64	2.71E-06
LDLR	15	53.37	1.65E-06
LEFTY2	3	14.54	6.97E-04
LGALS9	3	35.69	1.78E-08
LILRB4	2	NA	NA
PAG1	6	36.71	6.83E-07
PCSK9	8	27.93	2.26E-04
POLR2F	1	NA	NA
RP2	2	NA	NA
SIT1	6	43.76	2.59E-08
VCAM1	10	59.77	1.48E-09
VSTM2L	2	NA	NA

Supplementary Table 33. Heterogeneity test of protein on mother's attained age.

Protein	No. of SNPs	Q statistic	P-val of Q statistic
CDH1	6	3.37	6.43E-01
CDH17	10	15.02	9.05E-02
CDHR2	6	283.66	3.26E-59
CPE	6	17.55	3.56E-03

CXADR	5	7.74	1.01E-01
CXCL9	5	17.38	1.63E-03
F3	8	9.46	2.21E-01
FOXO3	1	NA	NA
GFAP	7	252.84	1.01E-51
IL19	3	5.17	7.52E-02
ITGB6	9	24.14	2.17E-03
LAIR1	3	10.44	5.42E-03
LEFTY2	3	3.51	1.73E-01
LGALS9	3	17.25	1.79E-04
LILRB4	2	NA	NA
POLR2F	1	NA	NA
RP2	2	NA	NA
STC2	4	11.75	8.29E-03
TNFRSF8	7	38.97	7.25E-07