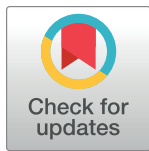


RESEARCH ARTICLE

Proteome-wide Mendelian randomization identifies causal links between blood proteins and severe COVID-19

Alish B. Palmos^{1,2‡}, Vincent Millischer^{3,4‡*}, David K. Menon⁵, Timothy R. Nicholson^{2,6}, Leonie S. Taams⁷, Benedict Michael⁸, Geraint Sunderland^{9,10}, Michael J. Griffiths^{9,11}, COVID Clinical Neuroscience Study Consortium, Christopher Hübel^{1,2,12‡}, Gerome Breen^{1,2‡}



1 Social, Genetic & Developmental Psychiatry Centre, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, United Kingdom, **2** UK National Institute for Health Research (NIHR) Biomedical Research Centre for Mental Health, South London and Maudsley Hospital, London, United Kingdom, **3** Department of Psychiatry and Psychotherapy, Medical University of Vienna, Vienna, Austria, **4** Department of Molecular Medicine and Surgery, Karolinska Institutet, Stockholm, Sweden, **5** Division of Anaesthesia, Department of Medicine, University of Cambridge, Cambridge, United Kingdom, **6** Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, United Kingdom, **7** Centre for Inflammation Biology & Cancer Immunology, Department of Inflammation Biology, School of Immunology & Microbial Sciences, King's College London, London, United Kingdom, **8** Institute of Infection and Global Health, University of Liverpool, Liverpool, United Kingdom, **9** Department of Clinical Infection, Microbiology and Immunology, Institute of Infection, Veterinary and Ecological Sciences, University of Liverpool, Liverpool, United Kingdom, **10** Department of Neurosurgery, Alder Hey Children's NHS Trust, Liverpool, United Kingdom, **11** Department of Neurology, Alder Hey Children's NHS Trust, Liverpool, United Kingdom, **12** National Centre for Register-based Research, Department of Economics and Business Economics, Aarhus University, Aarhus, Denmark

OPEN ACCESS

Citation: Palmos AB, Millischer V, Menon DK, Nicholson TR, Taams LS, Michael B, et al. (2022) Proteome-wide Mendelian randomization identifies causal links between blood proteins and severe COVID-19. *PLoS Genet* 18(3): e1010042. <https://doi.org/10.1371/journal.pgen.1010042>

Editor: Chris Cotsapas, Yale School of Medicine, UNITED STATES

Received: August 20, 2021

Accepted: January 18, 2022

Published: March 3, 2022

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pgen.1010042>

Copyright: © 2022 Palmos et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The analyses are based on open data. The links and origin of the summary statistics have been added as a column in [S1 Table](https://www.ebi.ac.uk/gwas/): <https://www.ebi.ac.uk/gwas/>

‡ ABP and VM indicates joint first authorship on this work. CH and GB are joint last authorship on this work.
* vincent.millischer@meduniwien.ac.at

Abstract

In November 2021, the COVID-19 pandemic death toll surpassed five million individuals. We applied Mendelian randomization including >3,000 blood proteins as exposures to identify potential biomarkers that may indicate risk for hospitalization or need for respiratory support or death due to COVID-19, respectively. After multiple testing correction, using genetic instruments and under the assumptions of Mendelian Randomization, our results were consistent with higher blood levels of five proteins GCNT4, CD207, RAB14, C1GALT1C1, and ABO being causally associated with an increased risk of hospitalization or respiratory support/death due to COVID-19 (ORs = 1.12–1.35). Higher levels of FAAH2 were solely associated with an increased risk of hospitalization (OR = 1.19). On the contrary, higher levels of SELL, SELE, and PECAM-1 decrease risk of hospitalization or need for respiratory support/death (ORs = 0.80–0.91). Higher levels of LCTL, SFTPD, KEL, and ATP2A3 were solely associated with a decreased risk of hospitalization (ORs = 0.86–0.93), whilst higher levels of ICAM-1 were solely associated with a decreased risk of respiratory support/death of COVID-19 (OR = 0.84). Our findings implicate blood group markers and binding proteins in both hospitalization and need for respiratory support/death. They, additionally, suggest that higher levels of endocannabinoid enzymes may increase the risk of hospitalization. Our research replicates findings of blood markers previously associated with COVID-19 and

publications/23696881 <https://www.ebi.ac.uk/gwas/publications/28240269> <https://www.ebi.ac.uk/gwas/publications/27989323> <https://www.ebi.ac.uk/gwas/publications/28369058> <https://www.ebi.ac.uk/gwas/publications/29875488> <https://www.ebi.ac.uk/gwas/publications/31217265> (Contacted corresponding author for full summary statistics). <https://datashare.ed.ac.uk/handle/10283/3649> <https://zenodo.org/record/2615265#.YaD2uJDMLQ> <https://www.ebi.ac.uk/gwas/publications/31727947> <https://www.ebi.ac.uk/gwas/publications/31320639> All results can be found in S1_Data.xlsx Analyses scripts are publicly available under https://github.com/tnggroup/PWMMR_Covid19.

Funding: BM and GB are supported to conduct COVID-19 neuroscience research (The Covid-19 Clinical Neuroscience Study (COVID-CNS)) by the Medical Research Council (UKRI/MRC; MR/V03605X/1); for additional neurological inflammation research due to viral infection BM is also supported by grants from the MRC/UKRI (MR/V007181/1), MRC (MR/T028750/1) and Wellcome (ISSF201902/3). CH acknowledges funding from Lundbeckfonden (R276-2018-4581). MJG is supported for neuroscience research internationally by MRC Newton Fund (MR/S019960/1), MRC Developmental Pathway Funding Scheme (MR/R015406/1), and National Institute for Health Research (NIHR; 153195 17/60/67, 126156 17/63/11, and 200907). DKM is also funded by the NIHR (through the Cambridge NIHR Biomedical Research Centre) and by the Addenbrooke's Charities Trust. This paper represents independent research partially funded by the National Institute for Health Research (NIHR) Biomedical Research Centre (BRC) at the South London and Maudsley NHS Foundation Trust and King's College London. The authors acknowledge use of the research computing facility at King's College London, Rosalind (<https://rosalind.kcl.ac.uk>), which is delivered in partnership with the National Institute for Health Research (NIHR) Biomedical Research Centres at South London & Maudsley and Guy's & St. Thomas' NHS Foundation Trusts, and part-funded by capital equipment grants from the Maudsley Charity (award 980) and Guy's & St. Thomas' Charity (TR130505). The funders had no role in study design, data collection, and analysis, decision to publish, or preparation of the manuscript.

Competing interests: I have read the journal's policy and the authors of this manuscript have the following competing interests: Gerome Breen has sat on pre-clinical advisory boards for Compass

prioritises additional blood markers for risk prediction of severe forms of COVID-19. Furthermore, we pinpoint druggable targets potentially implicated in disease pathology.

Author summary

As of November 2021, more than five million people have died due to COVID-19. Although vaccinations provide good protection, it is important to fully understand the biology behind the severe forms of COVID-19. Mendelian randomization facilitates the identification of blood proteins that may be involved in the pathophysiology of severe forms. Here, we investigated whether >3,000 blood proteins might play a role in hospitalization due to COVID-19 or the requirement of respiratory support or death due to COVID-19. Using genetic instruments and under the assumption of Mendelian randomization, our results are consistent with higher levels of five proteins being causally associated with an increased risk of both COVID-19 outcomes and higher levels of one protein associated with hospitalization. Our results are also consistent with higher levels of four proteins—mainly playing a role in cell adhesion—being causally associated with a decreased risk of hospitalization and respiratory support/death, and higher levels of four proteins being causally associated with a decreased risk of hospitalization. These proteins may represent new biomarkers useful in risk prediction of severity and may lead to new therapeutics by prioritizing druggable targets.

Introduction

The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) was identified in late 2019 in Wuhan, China, and is commonly referred to as coronavirus disease 2019 (COVID-19) that rapidly evolved into a global pandemic [1,2]. As of November 2021, more than 240 million cases have been confirmed worldwide with total deaths exceeding 5 million [3]. COVID-19 pathology encompasses a wide spectrum of clinical manifestations from asymptomatic, mild, moderate, to 15% being severe infections [1,2,4]. Severe COVID-19 commonly requires hospitalization and intensive care with assisted respiratory support, and respiratory failure is the most common reason for COVID-19 associated mortality [5].

A dysregulated pro- and anti-inflammatory immunomodulatory response is thought to drive much of the pathophysiology of COVID-19 and comprises alveolar damage, lung inflammation and pathology of an acute respiratory distress syndrome [1,2,6,7]. Given that the innate immune response has an individual-level genetic basis, genetic variants carried by an individual could play an important role in the individual-level immune response and, therefore, may influence progression and severity of COVID-19. This individual difference may also be key in our understanding of why some individuals require hospitalization due to the severity of their symptoms, whilst others are able to recover from COVID-19 without hospitalization [8]. In addition, once hospitalized, this individual difference may drive some people towards fatal outcomes or intensive care with respiratory support, whilst others are discharged from hospitals without respiratory complications.

Although immunomodulatory blood proteins can be studied in hospitalized and non-hospitalized patients with COVID-19, it is difficult to avoid potential confounding effects through factors, such as initial viral exposure/inoculum, smoking behavior, and high body mass index (BMI). These factors themselves are associated with high pro-inflammatory cytokine levels

Pathways Ltd and Otsuka Ltd. Michael J Griffiths has received research supported by Fast Track Diagnostics, Luxembourg and consultancy fees from Siemens Healthineers. There are no other conflicts to disclose.

and may represent independent risk factors for hospitalization, death, or respiratory failure as a result of COVID-19 [9,10].

Numerous genome-wide association studies (GWASs) in healthy populations associated genetic variants with immunomodulatory blood proteins [11–13]. In addition, the COVID-19 Host Genetics Initiative carried out GWASs on COVID-19 outcomes to understand the role of host genetic factors in susceptibility and severity of COVID-19 [8]: the first GWAS associated genetic variants with hospitalization due to COVID-19; and the second with need for respiratory support and death subsequently to a COVID-19 hospitalization. The findings suggest that once hospitalized another set of genetic variants may be responsible for a severe respiratory form of COVID-19, which may lead to the need for respiratory support or death.

These GWASs represent a powerful source of information to identify new biomarkers and therapeutic leads for drug development or repositioning. The method of Mendelian randomization can investigate the relationship between immunomodulatory blood proteins and a severe COVID-19 infection. Mendelian randomization exploits the fact that alleles are randomly inherited from parent to offspring in a manner analogous to a randomized-controlled trial, and allows estimation of putative causal effects of an exposure on a disease while avoiding confounding environmental effects, thus overcoming some of the limitations of observational studies. Recent advancements in Mendelian randomization methods allow use of GWAS summary statistics to identify genetic proxies (i.e., instrumental variables) of modifiable risk factors and test their association with disease outcomes [14,15]. We, therefore, conducted Mendelian randomization analyses between high levels of a large number of blood proteins and COVID-19, highlighting specific proteins associated with an increased risk of hospitalization due to COVID-19 and once hospitalized, an increased risk for need of respiratory support/death due to COVID-19. We identified putative causal associations that help us understand how innate differences in protein levels can affect the COVID-19 disease course and which proteins could be prioritized in clinical studies.

Methods

Blood protein GWAS data

In total, we amassed 5,305 sets of GWAS summary statistics for blood biomarkers [11–13,16–23]. A systematic search was performed based on the ontology lookup service (OLS; www.ebi.ac.uk/ols/index) using R and the packages ‘rols’ and ‘gwasrapidd’ between July 7th and July 27th, 2020. OLS is a repository for biomedical ontologies, such as gene ontology (GO) or the experimental factor ontology (EFO), including a systematic description of many experimental variables. First, all subnodes of the EFO ‘protein measurements’ (EFO:0004747) were determined using an iterative process based on the package ‘rols’. Overall, 628 unique EFO IDs were determined. Subsequently, all genetic associations reported in the GWAS Catalog [24] (www.ebi.ac.uk/gwas/) were identified and linked to the corresponding study using the package ‘gwasrapidd’. One hundred and seventy-eight unique GWAS catalog accession IDs with available summary statistics were curated manually before inclusion. In order to expand the dataset, studies published at a later date were included manually and the first and the last author of studies without publicly available summary statistics were contacted.

In total, we included ten publications for which summary data was readily available and processed those using standard GWAS summary statistics quality control metrics including removal of incomplete genetic variants, variants with information metrics of lower than 0.6 and allele frequencies more extreme than 0.005 or 0.995. Allele frequencies were estimated from raw genotypes of the European 1,000 Genomes Project dataset, where needed [25]. See [S1 Table](#) for a full list of studies included in these analyses. Links to the summary statistics are

also provided. Note that all protein measurements from all studies described above were included in our analyses, meaning that some proteins were analyzed more than once.

COVID-19 GWAS data

In order to capture increased risk of hospitalization as a result of severe COVID-19, we downloaded the COVID-19 Host Genome Initiative GWAS meta-analysis [8] of “Hospitalized covid vs. population”, European ancestry (B2_ALL_eur, release 5, January 2021; <https://www.covid19hg.org/results/r5/>). Cases were defined as SARS-CoV-2 infected individuals who required hospitalization due to COVID-19 related symptoms. Controls were defined as non-cases, i.e. the population. The European sample consisted of 9,986 cases and 1,877,672 controls. In our study, we refer to this GWAS as the hospitalization-COVID-19 GWAS.

In order to capture increased risk of very severe respiratory COVID-19, including respiratory support and death, we downloaded the COVID-19 Host Genome Initiative GWAS meta-analysis [8] of “very severe respiratory confirmed covid vs. population”, European ancestry (A2_ALL_eur, release 5, January 2021; www.covid19hg.org/results/r5). Cases were defined as SARS-CoV-2 infected individuals who were admitted to hospital, had COVID-19 as the primary reason for admission, and had died or needed respiratory support (i.e., intubation, continuous positive airway pressure, or bilevel positive airway pressure). Controls were defined as non-cases, i.e. the population [8]. The European sample consisted of 5,101 cases and 1,383,241 controls. In this study, we refer to this GWAS as the respiratory support/death-COVID-19 GWAS.

Mendelian randomization

To examine the influence of blood proteins on the risk of developing severe COVID-19, we selected genetic variants, single nucleotide polymorphisms (SNPs), that were strongly associated with actual blood protein levels in 5,504 genome-wide analyses of single proteins using robust methodologies (see [S1 Data](#), for more details on how the proteins were measured, and instruments for all significant proteins). Using these genetic loci as proxies for protein levels, we performed an analysis using Mendelian randomization, a method that enables tests of putative causal associations of these blood proteins with the development of severe COVID-19. We used the Generalized Summary data-based Mendelian randomization (GSMR) method as the base method [26]. GSMR tests for putative causal associations between a risk factor and a disease using multi-SNP effects from GWAS summary data. The HEIDI-outlier approach in GSMR removes SNP instruments with strong putative pleiotropic effects. In addition, GSMR accounts for linkage disequilibrium (LD) among SNPs not removed by clumping using a reference dataset for LD estimation. In this study, the European 1,000 Genomes dataset was used as the reference dataset [18].

For all GWASs, SNPs used as instrumental variables were selected by applying a suggestive genome-wide *p-value* threshold ($p < 5 \times 10^{-6}$), to identify enough SNPs (i.e., at least 5) in common between the exposure (e.g., blood marker) and outcome (e.g. COVID-19 hospitalization). Note that although reducing the *p-value* threshold may introduce potential false positive SNPs as instruments, SNPs with the strongest effect sizes are robust and reliable for conducting MR. The use of a lower *p-value* threshold in numerous MR studies is common [27–31], and we additionally calculated F-statistics and I-squared statistics to transparently present the strength of our instruments ([S2 Table](#)). We needed to take this analytical step as GWAS of blood proteins with more statistical power are not available at this time.

Where possible (due to SNPs in common between the exposure and the outcome), bidirectional analyses were performed. To account for multiple testing, we calculated false discovery rate (FDR) corrected Q values using the `p.adjust` function in R ($p_{\text{FDR}} = 0.05$) [32].

Sensitivity analyses

To test for robustness, we performed sensitivity analyses on our significant results from the GSMR analyses using additional Mendelian randomization methods, including the maximum likelihood, MR Egger, simple median, weighted median, inverse weighted median, inverse weighted median radial, inverse variance weighted (multiplicative random effects), inverse variance weighted (fixed effects), simple mode and weighted mode methods [33–35]. In order to pass our sensitivity analyses, at least nine of these ten methods must agree with the primary GSMR results.

Furthermore, when possible, we performed GSMR using only variants in the *cis* region of the gene encoding the blood marker (defined as variants either within a gene, up to 1 Mb proximal to the start of the gene, or up to 1 Mb distal to the end of the gene). Gene information was obtained from `ensembl` [36] using the `biomaRt` library [37], SNP information was obtained from NCBI dbSNP [38] using the `rsnps` library [39] (Gustavsen et al., “Get ‘SNP’ (‘Single-Nucleotide’ ‘Polymorphism’) Data on the Web [R Package `Rsnps` Version 0.4.0]” 2020).

With BMI being associated with both COVID-19 severity and the levels of many inflammatory proteins [9,40,41], we also performed GSMR with BMI both as exposure and outcome for all significant proteins.

Finally, given that many inflammatory and immunomodulatory proteins share genetic loci and may therefore be driving associations via genetically correlated SNPs in high linkage disequilibrium, we computed pairwise linkage disequilibrium for all SNPs used as instrumental variables of blood proteins that were significantly associated in our analyses. To calculate linkage disequilibrium, we used `LDlink` [42] and the CEU population panel (Utah residents from North and West Europe) as the reference.

Pathway analyses

KEGG pathway analysis was performed in R with the significant proteins for both outcomes separately, using the `clusterProfiler` library 4.0 [43]. To account for multiple testing, we calculated false discovery rate (FDR) on the pathways, the significance threshold was set at $p_{\text{FDR}} = 0.05$.

Results

We tested 3,890 associations with hospitalization-COVID-19 as the exposure and blood proteins as outcome (yielding 1 statistically significant association); and in reverse, 5,314 associations of blood proteins as the exposure and hospitalization-COVID-19 as the outcome (yielding 15 statistically significant associations). Additionally, we tested 2,687 associations with need for respiratory support/death-COVID-19 as the exposure (yielding 1 significant association); and in reverse, 3,273 associations with respiratory support/death-COVID-19 as the outcome (yielding 13 significant associations, [Table 1](#)). Our results show for some proteins, robust associations with the same proteins twice, as they were measured twice in independent GWASs, serving as direct replication. In order to easily identify these proteins, we added a suffix of the study name to the protein. In addition, note that units of protein measurement differed in studies, with some studies using standardized units (see studies in [S1 Table](#)). Thus, we will report our findings per standard deviation increase.

Table 1. Mendelian randomization results with COVID-19. Details of significant, false discovery rate-(FDR)-corrected Mendelian randomization results, using the Generalized Summary Data-based Mendelian randomization (GSMR) method. Using genetic instruments and under the assumptions of Mendelian randomization, the top section shows results consistent with six blood markers being significantly causally associated with an increased risk of hospitalization as a result of COVID-19 and the nine blood markers causally associated with a decreased risk of hospitalization, as well as one protein showing a decrease in risk of hospitalization. The bottom section shows results consistent with five blood markers being significantly causally associated with an increased risk for the need of respiratory support/death due to COVID-19 and eight blood markers causally associated with a decreased risk for the need of respiratory support/death due to COVID-19, as well as one protein decreasing risk for the need of respiratory support/death due to COVID-19. The table presents the log odds statistics (i.e., beta) and corresponding standard error as well as odds ratios, 95% confidence intervals, and the FDR-adjusted Q values ($p_{FDR} = 0.05$).

Outcome—Hospitalization as a Result of COVID-19								
Protein (exposure)	Beta	SE	p value	SNPs	OR	Lower 95% CI	Upper 95% CI	Q
FAAH2_Sun	0.17	0.03	1.54x10 ⁻⁰⁷	13	1.19	1.12	1.25	1.31x10 ⁻⁰⁴
GCNT4_Sun	0.15	0.03	8.15x10 ⁻⁰⁸	18	1.16	1.11	1.21	8.68x10 ⁻⁰⁵
CD207_Sun	0.11	0.02	1.95x10 ⁻⁰⁸	24	1.11	1.08	1.15	3.59x10 ⁻⁰⁵
RAB14_Sun	0.10	0.02	3.77x10 ⁻⁰⁸	24	1.11	1.07	1.14	5.29x10 ⁻⁰⁵
C1GALT1C1_Sun	0.10	0.02	4.06x10 ⁻⁰⁵	26	1.10	1.06	1.15	2.47x10 ⁻⁰²
ABO_Sun	0.07	0.01	1.36x10 ⁻⁰⁷	29	1.07	1.05	1.10	1.29x10 ⁻⁰⁴
LCTL_Sun	-0.08	0.02	1.51x10 ⁻⁰⁶	40	0.93	0.90	0.96	1.07x10 ⁻⁰³
SFTPD_Breth	-0.08	0.02	6.36x10 ⁻⁰⁵	16	0.92	0.88	0.96	3.61x10 ⁻⁰²
SELL_Sun	-0.09	0.02	3.65x10 ⁻⁰⁷	24	0.91	0.88	0.95	2.83x10 ⁻⁰⁴
SELE_Folk	-0.11	0.02	4.35x10 ⁻⁰⁸	16	0.90	0.86	0.94	5.29x10 ⁻⁰⁵
KEL_Sun	-0.11	0.03	9.05x10 ⁻⁰⁵	18	0.90	0.84	0.95	4.54x10 ⁻⁰²
SELE_Scal	-0.12	0.02	2.11x10 ⁻⁰⁸	50	0.88	0.84	0.93	3.59x10 ⁻⁰⁵
SELE_Breth	-0.13	0.03	8.49x10 ⁻⁰⁶	6	0.88	0.82	0.94	5.57x10 ⁻⁰³
ATP2A3_Sun	-0.16	0.04	8.99x10 ⁻⁰⁵	16	0.86	0.78	0.93	4.54x10 ⁻⁰²
PECAM1_Scal	-0.23	0.04	2.05x10 ⁻¹⁰	30	0.80	0.73	0.87	1.74x10 ⁻⁰⁶
Exposure—Hospitalization as a Result of COVID-19								
Protein (outcome)	Beta	SE	p value	SNPs				Q
MIP1b_Ahol	-0.16	0.03	7.76x10 ⁻⁰⁹	27				2.26x10 ⁻⁰⁵
Outcome—Respiratory support/death due to COVID-19								
Protein (exposure)	Beta	SE	p value	SNPs	OR	Lower 95% CI	Upper 95% CI	Q
GCNT4_Sun	0.30	0.05	3.36x10 ⁻¹¹	16	1.35	1.26	1.44	6.68x10 ⁻⁰⁸
RAB14_Sun	0.20	0.03	1.32x10 ⁻¹¹	27	1.22	1.16	1.28	3.92x10 ⁻⁰⁸
C1GALT1C1_Sun	0.19	0.04	3.19x10 ⁻⁰⁷	28	1.21	1.13	1.28	2.37x10 ⁻⁰⁴
CD207_Sun	0.16	0.03	2.97x10 ⁻⁰⁷	25	1.17	1.11	1.23	2.37x10 ⁻⁰⁴
ABO_Sun	0.11	0.02	8.35x10 ⁻⁰⁸	30	1.12	1.08	1.16	8.30x10 ⁻⁰⁵
SELE_Sliz	-0.11	0.02	5.87x10 ⁻⁰⁶	65	0.89	0.84	0.94	3.18x10 ⁻⁰³
SELL_Sun	-0.13	0.03	9.03x10 ⁻⁰⁶	24	0.88	0.83	0.94	4.49x10 ⁻⁰³
SELE_Scal	-0.17	0.04	4.35x10 ⁻⁰⁶	52	0.85	0.78	0.92	2.59x10 ⁻⁰³
sICAM1_Sliz	-0.17	0.04	2.98x10 ⁻⁰⁵	31	0.84	0.76	0.92	1.27x10 ⁻⁰²
SELE_Folk	-0.19	0.03	9.44x10 ⁻¹⁰	16	0.83	0.77	0.89	1.41x10 ⁻⁰⁶
SELE_Breth	-0.20	0.05	1.87x10 ⁻⁰⁵	6	0.82	0.73	0.91	8.57x10 ⁻⁰³
PECAM1_Folk	-0.26	0.05	1.49x10 ⁻⁰⁶	8	0.77	0.66	0.88	9.87x10 ⁻⁰⁴
PECAM1_Scal	-0.31	0.05	1.34x10 ⁻⁰⁹	31	0.73	0.63	0.83	1.60x10 ⁻⁰⁶
Exposure—Respiratory support/death due to COVID-19								
Protein (outcome)	Beta	SE	p value	SNPs				Q
NEP_Hill	-0.28	0.07	7.63x10 ⁻⁰⁵	24				3.03x10 ⁻⁰²

Note: SE = standard error, SNPs = number of single nucleotide polymorphisms in common between the exposure and outcome; OR = odds ratio, CI = confidence interval, Q = false discovery rate-adjusted p value; ABO = ABO system transferase; ATP2A3 = ATPase Sarcoplasmic/Endoplasmic Reticulum Ca²⁺ Transporting 3; C1GALT1C1 = C1GALT1 specific chaperone 1; CD207 = langerin; FAAH2 = Fatty Acid Amide Hydrolase 2; GCNT4 = glucosaminyl (N-Acetyl) transferase 4; KEL = Kell Metallo-Endopeptidase (Kell Blood Group); LCTL = Lactase-like protein; MIP1b = macrophage inflammatory protein; NEP = neprilysin; PECAM1 = platelet endothelial cell adhesion molecule; RAB14 = ras-related protein rab-14; SELE = E-selectin; SELL = L-selectin; SFTPD = Surfactant Protein D; sICAM1 = Soluble intercellular adhesion molecule-1. Names after the underscore are abbreviations for the study the protein was measured in.

<https://doi.org/10.1371/journal.pgen.1010042.t001>

Proteins associated with an elevated risk of hospitalization as a result of COVID-19

After multiple testing correction ($p_{FDR} = 0.05$), using genetic instruments and under the assumptions of Mendelian randomization, our results were consistent with six blood markers being significantly causally associated with an elevated risk of hospitalization as a result of COVID-19 (Figs 1 and S1); the reverse associations with risk of hospitalization as exposure and these six blood markers as outcome revealed no significant associations (S3 Table). Per

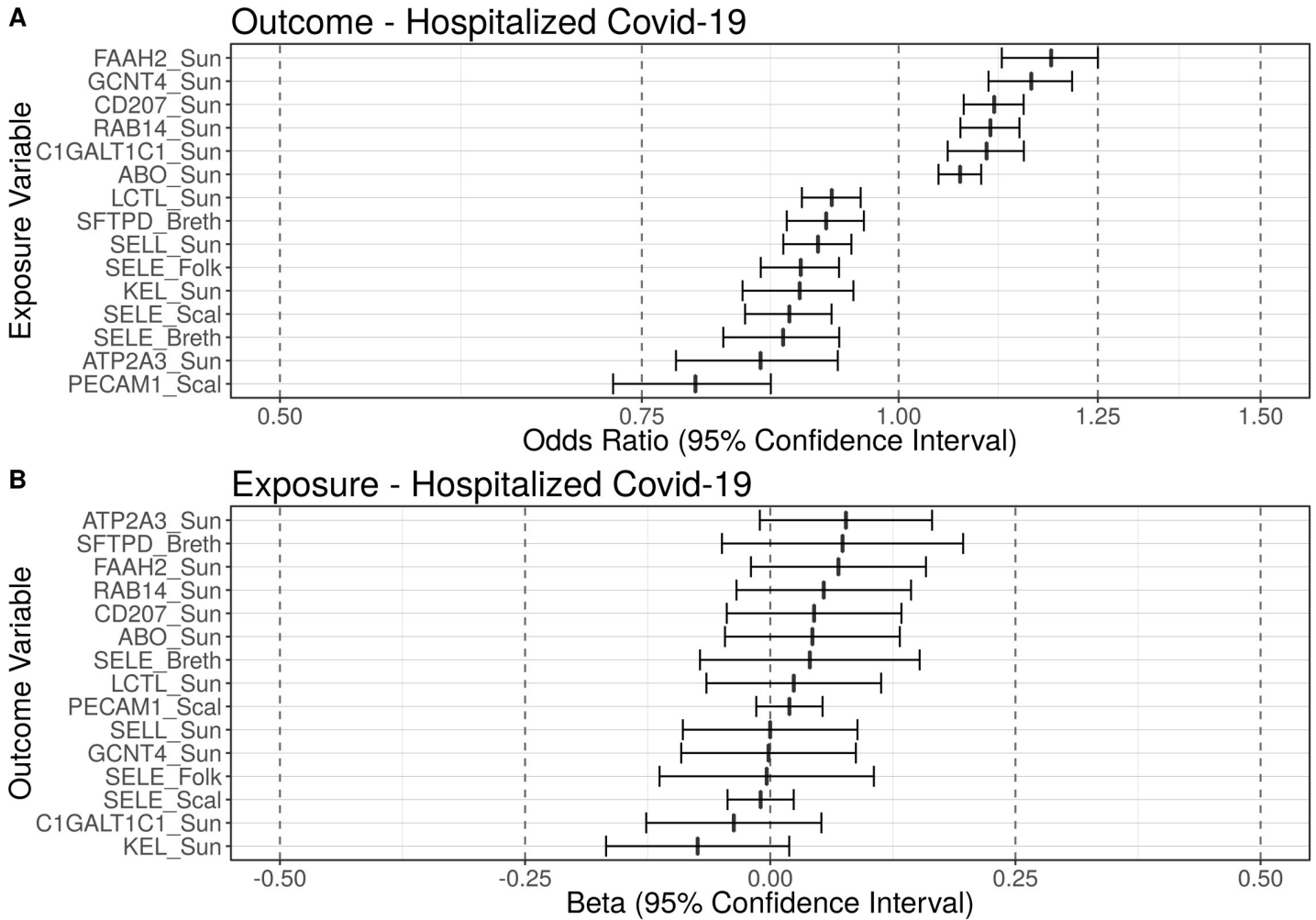


Fig 1. Blood markers putatively causally associated with hospitalized COVID-19. Summary figure of the false discovery rate-corrected ($p_{FDR} = 0.05$) Mendelian randomization results using the Generalized Summary data-based Mendelian randomization (GSMR) method. Using genetic instruments and under the assumptions of Mendelian randomization, this figure displays: (A) Summary figure when hospitalization-COVID-19 is the outcome of interest; (B) Summary figure when hospitalization-COVID-19 is the exposure of interest. Odds ratios (ORs) of the blood markers causally associated with hospitalized-Covid-19 are displayed on the x-axis (with 95% confidence intervals). The blood markers are displayed on the y-axis. The dashed line at one represents an odds ratio of one (i.e., no effect). Using genetic instruments and under the assumptions of Mendelian randomization, six blood markers were causally associated with a significantly increased risk for hospitalization COVID-19 and nine blood markers were causally associated with a significantly decreased risk for hospitalization ($q_{FDR} \leq 0.05$). ABO = ABO system transferase; ATP2A3 = ATPase sarcoplasmic/endoplasmic reticulum Ca²⁺ transporting 3; C1GALT1C1 = C1GALT1 specific chaperone 1; CD207 = langerin; FAAH2 = fatty acid amide hydrolase 2; GCNT4 = glucosaminyl (N-Acetyl) transferase 4; KEL = Kell metallo-endopeptidase (Kell Blood Group); LCTL = lactase-like protein; PECAM1 = platelet endothelial cell adhesion molecule; RAB14 = ras-related protein rab-14; SELE = E-selectin; SELL = L-selectin; SFTPD = surfactant protein D.

<https://doi.org/10.1371/journal.pgen.1010042.g001>

standard deviation (SD) increase in the respective blood marker our results were consistent with an increase in odds for hospitalization ranging from 7 to 19%, with fatty acid amide hydrolase 2 (FAAH2) showing the strongest effect: odds ratio (OR) = 1.19 (95% CI: 1.12, 1.25, $q \leq 0.01$; [Table 1](#)).

Proteins associated with a decreased risk of hospitalization as a result of COVID-19

After multiple testing correction ($p_{\text{FDR}} = 0.05$), using genetic instruments and under the assumptions of Mendelian randomization, our results were consistent with nine blood markers being significantly causally associated with a decreased risk of hospitalization as a result of COVID-19 ([Figs 1 and S2](#)); the reverse associations with these nine blood markers were non-significant ([S3 Table](#)). Per SD increase in the respective blood marker the decreases in odds for hospitalization ranged from 7 to 20%, with the platelet endothelial cell adhesion molecule (PECAM-1) showing the strongest effect: OR = 0.80 (95% CI: 0.73, 0.87, $q \leq 0.01$; [Table 1](#)).

Hospitalization as a result of COVID-19 associated with protein levels

In addition, using genetic instruments and under the assumptions of MR, our results were consistent with hospitalization being significantly causally associated with decreased levels of macrophage inflammatory protein (MIP1b): beta = -0.16 (SE = 0.03), $q \leq 0.01$; [Table 1](#)).

Proteins associated with an elevated risk of need for respiratory support/death due to COVID-19

After multiple testing correction ($p_{\text{FDR}} = 0.05$), using genetic instruments and under the assumptions of Mendelian randomization, our results were consistent with five blood markers being causally associated with need for respiratory support/death due to COVID-19 ([Figs 2 and S3](#)); the reverse associations with these blood markers as outcomes were nonsignificant ([S3 Table](#)). Per standard deviation (SD) increase in these respective blood marker the increase in odds for respiratory support/death ranged from 12 to 35%, with glucosaminyl (N-Acetyl) transferase 4 (GCNT4) showing the strongest effect: OR = 1.35 (95% CI: 1.26, 1.44, $q \leq 0.01$; [Table 1](#)).

Proteins associated with a decreased risk of need for respiratory support/death due to COVID-19

After multiple testing correction ($p_{\text{FDR}} = 0.05$), using genetic instruments and under the assumptions of Mendelian randomization, our results were consistent with eight blood markers being causally associated with a statistically significantly decreased risk of need for respiratory support/death due to COVID-19 ([Figs 2 and S4](#)); the reverse associations with these blood markers were nonsignificant ([S3 Table](#)). Per standard deviation (SD) increase in the respective blood marker the increases in odds for respiratory support/death ranged from 11 to 27%, with the platelet endothelial cell adhesion molecule (PECAM-1) showing the strongest effect size: OR = 0.73 (95% CI: 0.63, 0.83, $q \leq 0.01$; [Table 1](#)).

Need for respiratory support/death due to COVID-19 associated with protein levels

In addition, using genetic instruments and under the assumptions of Mendelian randomization, our results were consistent with respiratory support/death due to COVID-19 being

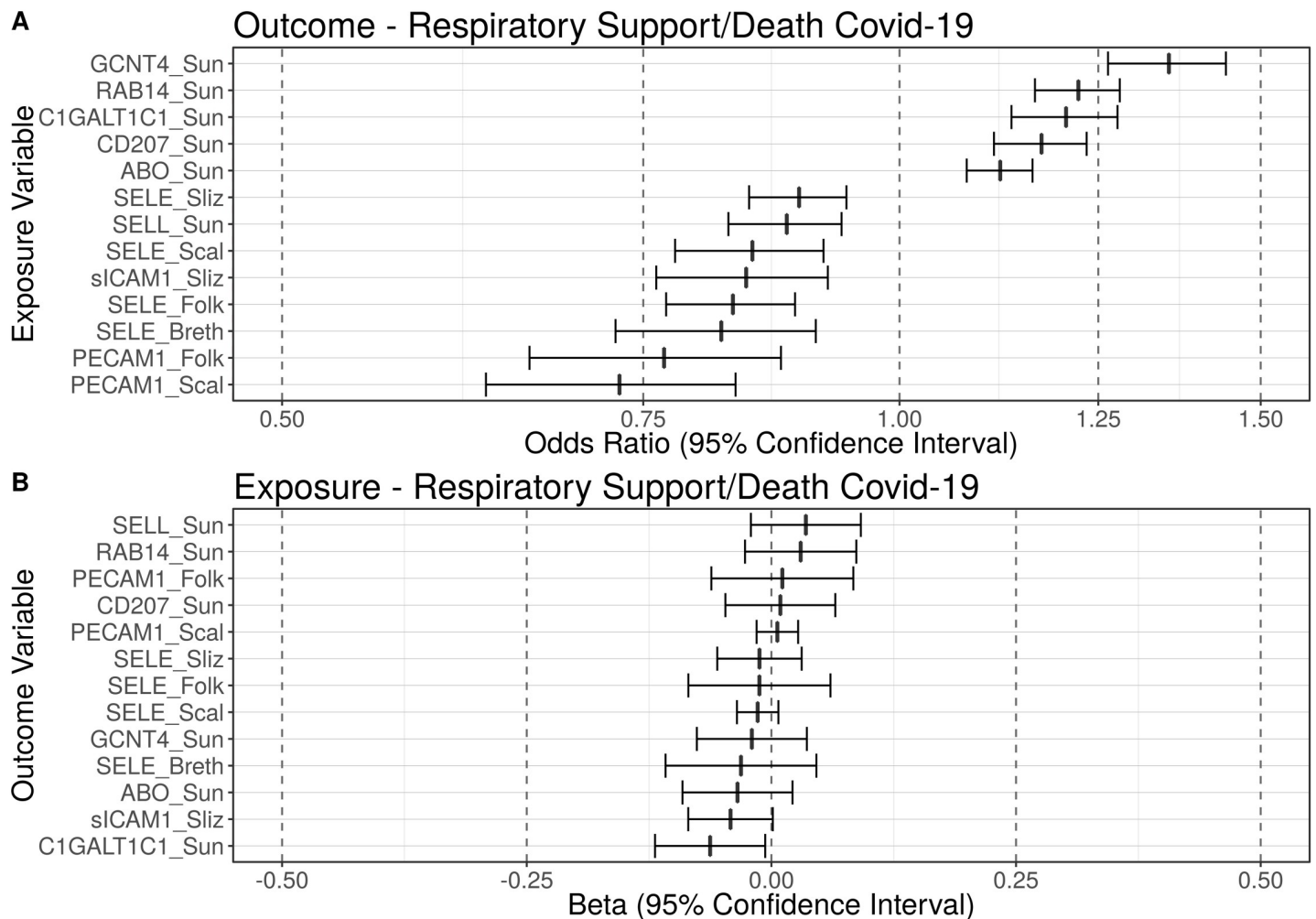


Fig 2. Blood markers putatively causally associated with need for respiratory support/death due to COVID-19. Summary figure of the false discovery rate-corrected ($p_{FDR} = 0.05$) Mendelian randomization results using the Generalised Summary data-based Mendelian randomization (GSMR) method. Using genetic instruments and under the assumptions of Mendelian randomization, this figure displays: (A) Summary figure when respiratory support/death-COVID-19 is the outcome of interest; (B) Summary figure when respiratory support/death-COVID-19 is the exposure of interest. Odds ratios (ORs) of blood markers causally associated with the need for respiratory support/death due to COVID-19 are displayed on the x-axis (with 95% confidence intervals). The blood markers are displayed on the y-axis. The dashed line at one represents an odds ratio of one (i.e., no effect). Using genetic instruments and under the assumptions of Mendelian randomization, five blood markers were causally associated with a significantly increased risk for need for respiratory support/death due to COVID-19 and eight blood markers were causally associated with a significantly decreased risk for respiratory support/death ($q_{FDR} \leq 0.05$). ABO = ABO system transferase; C1GALT1C1 = C1GALT1 specific chaperone 1; CD207 = langerin; GCNT4 = glucosaminyl (N-Acetyl) transferase 4; LCTL = lactase-like protein; PECAM1 = platelet endothelial cell adhesion molecule; RAB14 = ras-related protein rab-14; SELE = E-selectin; SELL = L-selectin; sICAM1 = soluble intercellular adhesion molecule-1.

<https://doi.org/10.1371/journal.pgen.1010042.g002>

significantly causally associated with decreased levels of neprilysin (NEP): beta = -0.28 (SE = 0.07, $q \leq 0.01$; [Table 1](#)).

Sensitivity analyses

Additional Mendelian randomization analyses. To further increase confidence in the findings, we additionally filtered our results for those concurrent with results calculated by additional Mendelian randomization methods (Maximum likelihood, MR Egger, Simple median, Weighted median, Inverse variance weighted (IVW), IVW radial, IVW multiplicative random effects, IVW fixed effects, Simple mode, Weighted mode).

Our sensitivity analyses confirmed the robustness of the association between 14 out of 15 blood markers and hospitalized-COVID-19, with ATP2A3 not satisfying the sensitivity analysis criteria (see [S4 and S5 Tables](#), for a full breakdown of these sensitivity analyses).

Our sensitivity analyses also confirmed the association between GCNT4_Sun, RAB14_Sun, CD207_Sun, SELL_Sun, sICAM1_Sliz, SELE_Folk, PECAM1_Scal and severe-COVID-19. C1GALT1C1_Sun, ABO_Sun, sEselectin_Sliz, SELE_Scal, SELE_Breth, PECAM1_Folk did not survive our sensitivity analyses (see [S4 and S5 Tables](#) for a full breakdown of these sensitivity analyses).

Strength of genetic instruments. Given that we are using a reduced p-value threshold to identify SNPs as our genetic instruments, we calculated F-statistic and I-squared on all our significantly associated proteins. We found that with the exception of ATP2A3, all F-statistics were larger than 20 and all I-squared statistics were larger than 0.9, suggesting strong genetic instruments for use in analyses [[44–46](#)]. See [S2 Table](#) for full results of these analyses.

Linkage disequilibrium analyses. Given that many inflammatory and immunomodulatory proteins share genetic loci which may therefore be driving associations via SNPs in high linkage disequilibrium, we computed pairwise linkage disequilibrium statistics for all SNPs used as instrumental variables for blood markers significantly associated with our outcomes.

In our analysis of hospitalization as a result of COVID-19, we used 296 unique SNPs as instrument variables for blood proteins, in 30 cases, one SNP was used as an instrument for two or more different proteins, 27 of which were located on chromosome 9. In the assessment of pairwise LD, 40 SNP pairs (based on 55 unique SNPs) showed high LD ($r^2 > 0.6$). Most pairs ($n = 37$) were located on chromosome 9 carrying the ABO gene ([S6 Table and S5 Fig](#)).

In our analysis of the need for respiratory support/death due to COVID-19, we used 305 unique SNPs as instrument variables for blood proteins; however, in 31 cases, one SNP was used twice as an instrument for two or more different proteins, 28 of which were located on chromosome 9. Furthermore, assessing pairwise LD between all SNPs, we found 50 SNP pairs (based on 65 unique SNPs) in high LD ($r^2 > 0.6$). Most pairs ($n = 44$) were located on chromosome 9 carrying the ABO gene ([S7 Table and S6 Fig](#)) [[47–49](#)].

cis-SNP effects from significantly associated proteins. In order to establish whether the significant associations identified in our study were being driven by cis-regulatory variants, we identified cis-SNPs from all significant blood markers and performed Mendelian randomization analyses using only these SNPs with the respective COVID-19 GWASs. Out of the 28 exposure-outcome pairs, only seven (25%) were based on at least one cis-SNP and could therefore be analysed. The results show that ABO_Sun and SFTPD_Breth cis-SNPs are significantly associated with hospitalization, and ABO_Sun and sICAM-1_Sliz cis-SNPs are significantly associated with need for respiratory support/death. All other significant associations are deemed to be driven by trans-SNPs ([S8 Table](#)).

Mendelian randomization analyses with body mass index. High BMI has been robustly associated with both inflammatory cytokine levels in blood and an increased risk of severe COVID-19 [[9,40,41](#)]. After validating this relationship by performing Mendelian randomization analysis between BMI and the two COVID-19 outcomes ([S9 Table](#)), we performed bidirectional Mendelian randomization analyses between all significant blood markers and BMI, using the largest publicly available BMI GWAS [[50](#)]. Our results indicate that genetic susceptibility for higher BMI is associated with higher levels of SELE_Scal, C1GALT1C1_Sun, SELE_Folk, KEL_Sun, SELL_Sun, RAB14_Sun and SFTPD_Breth. In addition, the genetic susceptibility for higher levels of LCTL_Sun, SELE_Sliz, SFTPD_Breth, PECAM-1_Scal and RAB14_Sun is associated with higher BMI (see [S10 Table](#) for full results). Note that some protein GWASs had controlled for BMI, whereas others, the majority of which show a significant association with BMI, did not (see [S1 Table](#)).

Table 2. Groupings of statistically significantly associated proteins by their biological processes.

Protein	Association
Blood group proteins	
ABO	Risk of hospitalization & need of respiratory support or death due to COVID
KEL	Protection against hospitalization
Antigen recognition	
CD207	Risk of hospitalization & need of respiratory support or death due to COVID
SFTPD	Protection against hospitalization
Adhesion molecules	
SELL	Protection against hospitalization & need of respiratory support or death due to COVID
SELE	Protection against hospitalization & need of respiratory support or death due to COVID
PECAM-1	Protection against hospitalization & need of respiratory support or death due to COVID
sICAM-1	Protection against need of respiratory support or death due to COVID
Transporters	
RAB14	Risk of hospitalization & need of respiratory support or death due to COVID
ATP2A3	Protection against hospitalization
Enzymes	
GCNT4	Risk of hospitalization & need of respiratory support or death due to COVID
C1GALT1C1	Risk of hospitalization & need of respiratory support or death due to COVID
FAAH2	Risk of hospitalization
LCTL	Protection against hospitalization

Note: ABO = ABO system transferase; ATP2A3 = ATPase sarcoplasmic/endoplasmic reticulum Ca²⁺ transporting 3; C1GALT1C1 = C1GALT1 specific chaperone 1; CD207 = langerin; FAAH2 = fatty acid amide hydrolase 2; GCNT4 = glucosaminyl (N-Acetyl) transferase 4; KEL = Kell metallo-endopeptidase (Kell Blood Group); LCTL = lactase-like protein; PECAM1 = platelet endothelial cell adhesion molecule; RAB14 = ras-related protein rab-14; SELE = E-selectin; SELL = L-selectin; SFTPD = surfactant protein D; sICAM1 = soluble intercellular adhesion molecule-1.

<https://doi.org/10.1371/journal.pgen.1010042.t002>

Pathway analysis. Proteins significantly associated with respiratory support/death as a result of COVID-19 were significantly enriched in three KEGG pathways: “Cell adhesion molecules” (hsa04514), “Mucin type O-glycan biosynthesis” (hsa00512) and “Malaria” (hsa05144). No enrichment was found at the defined significance threshold for hospitalization as a result of COVID-19, however “cell adhesion molecules” and “Mucin type O-glycan biosynthesis” (hsa00512) showed the strongest signal ($q_{FDR} = 0.07$)

Discussion

Using genetic instruments and under the assumptions of Mendelian randomization, our proteome-wide analyses are consistent with higher levels of certain blood proteins being causally associated with risk of being hospitalized due to COVID-19, and subsequently experiencing the most severe form including respiratory support or ending lethal (i.e., respiratory support/death in the following). All these proteins have detectable blood plasma or serum levels. For our discussion, we grouped the proteins by function in [Table 2](#) and provided more detail in [S11 Table](#).

It is important to note that, in our analyses, we did not identify typical canonical immune proteins, such as interleukin 6 or C-reactive protein [51,52]. This suggests that with a larger database of proteins we can pinpoint non-canonical immunomodulatory proteins relevant to disease pathophysiology. We did, however, estimate associations for some proteins twice, as they were measured separately in independent GWASs. Results from these analyses displayed

the same direction of effects, rendering them a direct replication and increasing the validity of our findings.

Blood group proteins

In the blood group protein group, using genetic instruments and under the assumptions of Mendelian randomization, our findings were consistent with ABO being causally associated with both an increased risk of hospitalization as well as the requirement of respiratory support or death by COVID-19 (i.e., respiratory support/death). ABO is an enzyme with glycosyltransferase activity that determines the ABO blood group of an individual [53]. However, the precise blood group associated with the increased risk for hospitalization as a result of COVID-19 cannot be determined from our results, as the probe for the blood marker measures both the A and B isoform of the protein while not showing a signal for O. Given the underlying British population of the original GWAS, A should be the more prevalent blood group (24%) in the sample compared to B (8%) [54]. Nevertheless, it is more likely that A, B, or the combination of A and B is associated with higher risk for hospitalization. Our findings confirm previous reports of the ABO blood group system being an important risk factor for a severe COVID-19 infection. For example, the proportion of group A is higher in COVID-19 positive individuals than in controls [55–60], and group A has been associated with higher mortality [61]. All evidence taken together suggests that blood group A is the more likely candidate for follow-up studies. Additionally, KEL, which is part of the complex Kell blood group system that contains many highly immunogenic antigens [62], was associated with a decreased risk of hospitalization as a result of COVID-19. This supports the notion that Kell negative individuals may be more susceptible to COVID-19 [63].

Antigen recognition

In our study, CD207, also known as langerin, was associated with hospitalization as well as the requirement of respiratory support or death by COVID-19. This protein is exclusively expressed in Langerhans cells (LC)—the first dendritic cells to encounter pathogens entering the body via the mucosa or skin [64]. Langerin binds COVID-19 glycoprotein glycans; however, it does not mediate transfection of COVID-19 pseudovirions in a T-lymphocyte cell line [65], rendering its role in COVID-19 infections inconclusive [66]. Our findings showed evidence consistent with high levels of SFTPD potentially protecting against COVID-19 hospitalization. SFTPD is strongly expressed in lung, brain, and adipose tissue, and contributes to the lung's defense against microorganisms, antigens, and toxins [67]. SFTPD also interacts with COVID-19 spike proteins [68]. In COVID-19, expression findings are mixed: some studies show that SFTPD is highly expressed in the lungs of COVID-19 patients [68], whereas others evidence a decreased expression [69]. Moreover, another study which investigated gene expression patterns in COVID-19-affected lung tissue and SARS-CoV-2 infected cell-lines, report a downregulation of SFTPD along with several regulatory partners [70]. Given its role in immunomodulation and air exchange in the lung, this supports our finding that higher levels of SFTPD may be causally associated with COVID-19 immunity [71,72]. Although more research is needed; however, ours and others' findings imply that SFTPD may protect against severe forms of COVID-19.

Adhesion molecules

Using genetic instruments and under the assumptions of Mendelian randomization, our analysis was consistent with the adhesion molecules SELE, SELL, and PECAM-1 being causally associated with a decreased risk of both hospitalization and the requirement of respiratory

support/death by COVID-19, while ICAM-1 was only protective against respiratory support/death. This is in keeping with results from out pathway analyses which suggest a significant enrichment in the KEGG pathway “cell adhesion molecules”. Studies have suggested that late stage COVID-19 is an endothelial disease [73]. The vascular endothelium is the crucial interface between blood and other tissues, regulating vascular structure, permeability, vasomotion, inflammation, and oxidative stress [73]. SELL and SELE are members of the selectin class of leukocyte adhesion molecules, which facilitate slow rolling of blood leukocytes along the endothelium [73]. Specifically, SELL promotes initial tethering and rolling of leukocytes to the endothelium [74,75] and SELE is responsible for the accumulation of blood leukocytes at sites of inflammation by mediating the adhesion of cells to the vascular lining [76,77]. The firm binding of leukocytes to the endothelial surface depends upon other molecules, such as PECAM-1, which is a cell adhesion molecule required for leukocyte transendothelial migration under most inflammatory conditions [78,79], and our results were consistent with it being protective against hospitalization. Once tightly bound, chemoattractant cytokines can signal to the bound leukocytes to traverse the endothelial monolayer and enter tissues where they can combat pathogenic invaders and initiate tissue repair [80]. This may be one of the biological explanations why we saw elevated levels of SELL, SELE, and PECAM-1 as being protective against hospitalization. ICAM-1—our results being consistent with it being protective against hospitalization—mediates cell-cell adhesion and is involved in inflammation [81]. Contrary to our findings, higher ICAM-1 levels have been associated with COVID-19 severity [82,83], requiring follow-up investigations. In summary, molecules that mediate the interaction between immune cells and blood vessels may be important in late stage COVID-19 and moderate severity.

Transporter molecules

In the protein transporter/trafficking group, using genetic instruments and under the assumptions of Mendelian randomization, our results were consistent with RAB14 being causally associated with an increased risk of hospitalization and respiratory support/death, whereas ATP2A3 may protect against hospitalization. Rab proteins are central regulators of phagosome maturation; RAB14 particularly regulates the interaction of phagosomes with early endocytic compartments [84]. One study identified RAB14 GTPases as a critical COVID-19 host factor: coronaviruses hijack Rab GTPase in host cells to replicate [85]. Additionally, whole genome analysis of COVID-19 lung tissue identified RAB14 polymorphisms that alter its binding to some miRNAs [86]. Therefore, Rab GTPases could be therapeutic targets. ATP2A3 is a magnesium-dependent ATP hydrolase, transports calcium from the cytosol into the sarcoplasmic/endoplasmic reticulum involved in muscular excitation/contraction and contributes to calcium sequestration [87,88]. Note that ATP2A3 has previously been genetically associated with severe COVID-19, but its exact role in infection remains unclear [89]. Cardiac failure in severe COVID-19 has been reported [90,91], hence, ATP2A3 may be involved as it regulates cardiomyocyte contraction [92]. However, note that ATP2A3 did not survive our sensitivity analyses, so this finding would need further validation.

Enzymes

In the enzyme group, which consists predominantly of glycosylases and hydrolases, using genetic instruments and under the assumptions of Mendelian randomization, GCNT3, a member of the GCNT family, was consistent with increasing the risk for hospitalization and respiratory support/death. GCTN3 proteins mediate mucin synthesis, branching, and oligomerization [93], a pathway showing a significant enrichment in our analyses. Glycosylation of COVID-19 viral surface antigens may help the virus evade the host immune system by

shielding its protein surface and, therefore, may prevent the development of an effective immune response [94]. In addition, as part of our innate immunity, the epithelial barrier made up of mucins acts as the first line of defense [95]. Our analysis also showed evidence consistent with C1GALT1C1 being risk increasing. C1GALT1C1 is a molecular chaperone required for the expression of active T-synthase, the only enzyme that glycosylates the Tn antigen [96]. Anti-Tn antibodies are lower in COVID-19 patients than noninfected individuals and individuals with blood group A [97]. Anti-Tn antibodies may protect against COVID-19 [97]. The glycosylation of Tn by C1GALT1C1 may suppress these antibodies. The glycobiology of COVID-19 includes glycans on viral proteins and host glycosaminoglycans that are critical in infections [98]. FDA-approved drugs, such as glycans for vaccines, anti-glycan antibodies, recombinant lectins, lectin inhibitors, glycosidase inhibitors, polysaccharides, and numerous glycosides may be repurposing targets for COVID-19 [98]. Our analysis also showed evidence consistent with FAAH2 being risk increasing for hospitalization as a result of COVID-19. FAAH2, is a fatty acid hydrolase involved in endocannabinoid uptake and inactivation [99]. Cannabinoids may reduce pulmonary inflammation through immunomodulation, decrease polymorphonuclear leukocytes infiltration, reduce fibrosis, decrease viral replication, and modulate the 'cytokine storm' in COVID-19 [100–103]. Cannabinoids have been suggested as anti-inflammatory treatment in COVID-19 [102,103]. Our results are also consistent with LCTL, being protective against hospitalization. LCTL is a glycosidase which hydrolyses glycosidic bonds. Little is known about this protein, and nothing in the context of COVID-19.

Using genetic instruments and under the assumptions of Mendelian randomization, our analysis was also consistent with hospitalization due to COVID-19 decreasing levels MIP1b. MIP1b is a major proinflammatory factor, acting as a chemoattractant for natural killer cells [104]. This association has been reported previously [105], with studies suggesting that MIP1b is a key mediator in the immune response against COVID-19 [106]. Also, in line with our findings of respiratory support/death due to COVID-19 decreasing levels of NEP, NEP protects against pulmonary inflammation and fibrosis [107]. Other studies suggest repurposing of roflumilast, a treatment for chronic obstructive pulmonary disorder, that increases NEP activity and, hence, increases anti-inflammatory activity [107].

Our study has several limitations. First, although we required confirmation of our findings by several Mendelian randomization methods, we set the p -value threshold for selecting genetic variants as our instruments at genome-wide suggestive significance ($p < 5 \times 10^{-6}$) to identify enough SNP instruments for each protein to run Mendelian randomization analyses. Some genetic variants may therefore be false positive associations with protein levels. This procedure is common in Mendelian randomization analyses [27–31]. Additionally, we report F-statistics and I-squared statistics per SNP instrument (S2 Table) so that the quality of our instruments is transparent. This step was necessary as the available GWASs of the blood proteins are still of limited sample size, and therefore statistical power. However, we identified several blood markers robustly associated with COVID-19, including ABO, suggesting that our SNP instruments pick up true associations and potential underlying biology of COVID-19. In addition, we note that for most associations (such as between FAAH2 and hospitalized COVID-19, or sICAM1 and severe COVID-19) we observe a clear linear relationship (S1–S4 Figs). However, some associations, such as between ABO and hospitalized COVID-19, or GCNT4 and severe COVID-19, display SNP effects between the exposure and outcome with very large standard errors. While these SNPs have less statistical power, the joint effect over all SNPs remains significant. Second, some blood marker GWASs were excluded from our analyses due to unavailability, therefore, we may have missed associations with these markers. Third, although the COVID-19 GWASs used in our analyses were carefully chosen to

represent two different phenotypes—hospitalization and respiratory support/death due to COVID-19—other GWAS may be better powered for identifying severe COVID-19 phenotypes. However, those were not publicly available when we conducted our study. Fourth, some of our SNP instruments for blood proteins are either the same or are in high LD, potentially tagging the same causal variant. Although this may indicate pleiotropy across blood markers, these instruments only represent a minority of our instruments (28% for hospitalization and 31% for respiratory support/death, respectively). Furthermore, a significant proportion of these SNPs are either used as instruments for the same protein from separate studies (e.g., in the case of SELE or PECAM), or are used as pleiotropic instruments for the different adhesion molecules (the family of SELE, PECAM-1, ICAM and SELL). Therefore, they should not overly influence our results. Fifth, the proteins in the original GWAS were measured in blood, thus not necessarily reflecting their intracellular concentrations. Therefore, it is difficult to draw conclusions on intracellular concentrations based on our results and further cellular research is required. Sixth, we used several robust MR methods with varying abilities to detect heterogeneity and pleiotropy; however, residual heterogeneity or pleiotropy may still be present. However, this is common to most Mendelian randomization analyses. Sixth, our sensitivity analyses demonstrate that a genetic predisposition to high BMI is significantly associated with some of our blood markers. High BMI is also associated with severe COVID-19 ([41], [S9 Table](#)), suggesting that BMI may drive risk of severe forms of COVID-19 and the change in blood protein levels, potentially confounding our results. However, our Mendelian randomization analyses showed that higher BMI is causal for higher levels of the COVID-19 protective proteins SELE, KEL, SELL, and causal for lower levels of the COVID-19 risk increasing proteins C1GALT and RAB14. Our sensitivity analysis therefore shows that BMI influences these blood markers in opposite directions than would be concurrent with confounding effects. This indicates the effects of these proteins are independent of BMI. Also, some blood protein GWASs controlled for BMI which should in the first place reduce the potential confounding ([S1 Table](#)). Finally, our findings are based on measurements in ancestral European populations due to data availability; therefore, future endeavors should include participants from more diverse ancestry.

Our results highlight the utility of applying large scale Mendelian randomization analyses to identify blood markers that may be causal for severe COVID-19. Using genetic instruments and under the assumptions of Mendelian randomization, our findings are consistent with higher levels of GCNT4, RAB14, C1GALT1C1, CD207 and ABO causally increasing the risk of both hospitalization and need of respiratory support or death due to COVID-19, and higher levels of FAAH2 increasing the risk of hospitalization. Our results were also consistent with higher levels of a number of adhesion molecules, including SELE, SELL, PECAM-1 and ICAM-1, as being protective against both hospitalization and a need of respiratory support or death. This adds to a growing body of evidence for the involvement of adhesion and endothelial dysfunction in severe COVID-19. Moreover, our results were consistent with higher levels of LCTL, SFTPD and KEL being protective against hospitalization alone. Together, our findings support previous findings and identify novel blood markers associated with a severe COVID-19 phenotype, indicating possible avenues to develop prognostic biomarkers and therapeutics for COVID-19.

Supporting information

S1 Table. A breakdown of all studies from which inflammatory marker genome-wide association study (GWAS) data originated.

(DOCX)

S2 Table. Validation of our genetic instruments.

(DOCX)

S3 Table. Reverse effect results from Mendelian randomization analyses.

(DOCX)

S4 Table. Results from sensitivity analyses for all markers and risk hospitalization as a result of COVID-19.

(DOCX)

S5 Table. Results from sensitivity analyses for all markers and respiratory support/death as a result of COVID-19.

(DOCX)

S6 Table. Table indicating the heterogeneous SNP used as instruments for at least 2 biomarkers identified in the hospitalization from COVID-19 GWAS.

(DOCX)

S7 Table. Table indicating the heterogeneous SNP used as instruments for at least 2 biomarkers identified in the respiratory support/death as a result of COVID-19 GWAS.

(DOCX)

S8 Table. cis-SNP effects from the significantly associated proteins.

(DOCX)

S9 Table. COVID-19 associations with BMI.

(DOCX)

S10 Table. Blood marker associations with BMI.

(DOCX)

S11 Table. Details on the tissue, function, and Covid-19 relevance of each significant blood biomarker.

(DOCX)

S1 Fig. Effect plots of blood markers associated with an increase in hospitalized COVID-19 risk.

(DOCX)

S2 Fig. Effect plots of blood markers associated with a decrease in hospitalized COVID-19 risk.

(DOCX)

S3 Fig. Effect plots of blood markers associated with an increase in risk of respiratory support/death as a result of COVID-19.

(DOCX)

S4 Fig. Effect plots of blood markers associated with a decrease in risk of respiratory support/death as a result of COVID-19.

(DOCX)

S5 Fig. Plot indicating the number of heterogeneous SNP instruments on chromosome 9 from biomarkers identified in the hospitalization from COVID-19 GWAS.

(DOCX)

S6 Fig. Plot indicating the number of heterogeneous SNP instruments on chromosome 9 from biomarkers identified in the respiratory support/death as a result of COVID-19

GWAS.
(DOCX)

S1 Data. The sheets in this spreadsheet displays all of our results, including a summary table of our main findings, all sensitivity analyses (different MR methods, sensitivity analyses relating to cis-SNP effects of all blood markers, sensitivity analyses relating to the association between significant blood markers and BMI, sensitivity analyses relating to the genetic instruments used in our analyses), the lists of the associations between both COVID phenotypes and all blood markers, in both directions, as well the full details of all SNP instruments used for the association with significant blood markers.
(XLSX)

Acknowledgments

The views expressed are those of the author(s) and not necessarily those of the National Health Service (NHS), the National Institute of Health Research (NIHR), King's College London, or the Department of Health and Social Care.

Author Contributions

Conceptualization: Alish B. Palmos, Vincent Millischer, David K. Menon, Timothy R. Nicholson, Leonie S. Taams, Benedict Michael, Christopher Hübel, Gerome Breen.

Data curation: Alish B. Palmos, Vincent Millischer, Christopher Hübel.

Formal analysis: Alish B. Palmos, Vincent Millischer, Christopher Hübel.

Funding acquisition: Benedict Michael, Gerome Breen.

Investigation: Alish B. Palmos, Vincent Millischer, Christopher Hübel.

Methodology: Alish B. Palmos, Vincent Millischer, Geraint Sunderland, Christopher Hübel.

Project administration: Alish B. Palmos, Vincent Millischer, Christopher Hübel.

Resources: Gerome Breen.

Software: Alish B. Palmos, Vincent Millischer.

Supervision: Michael J. Griffiths, Christopher Hübel, Gerome Breen.

Validation: Alish B. Palmos, Vincent Millischer, Christopher Hübel.

Visualization: Alish B. Palmos, Vincent Millischer.

Writing – original draft: Alish B. Palmos, Vincent Millischer, Christopher Hübel, Gerome Breen.

Writing – review & editing: Alish B. Palmos, Vincent Millischer, David K. Menon, Timothy R. Nicholson, Leonie S. Taams, Benedict Michael, Geraint Sunderland, Michael J. Griffiths, Christopher Hübel, Gerome Breen.

References

1. Blanco-Melo D, Nilsson-Payant BE, Liu W-C, Uhl S, Hoagland D, Møller R, et al. Imbalanced Host Response to SARS-CoV-2 Drives Development of COVID-19. *Cell*. 2020; 181: 1036–1045.e9. <https://doi.org/10.1016/j.cell.2020.04.026> PMID: 32416070
2. García LF. Immune Response, Inflammation, and the Clinical Spectrum of COVID-19. *Front Immunol*. 2020; 11: 1441. <https://doi.org/10.3389/fimmu.2020.01441> PMID: 32612615

3. Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect Dis.* 2020; 20: 533–534. [https://doi.org/10.1016/S1473-3099\(20\)30120-1](https://doi.org/10.1016/S1473-3099(20)30120-1) PMID: 32087114
4. Docherty AB, Harrison EM, Green CA, Hardwick HE, Pius R, Norman L, et al. Features of 20 133 UK patients in hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: prospective observational cohort study. *BMJ.* 2020; 369: m1985. <https://doi.org/10.1136/bmj.m1985> PMID: 32444460
5. Berlin DA, Gulick RM, Martinez FJ. Severe Covid-19. *N Engl J Med.* 2020; 383: 2451–2460. <https://doi.org/10.1056/NEJMcp2009575> PMID: 32412710
6. Kadkhoda K. COVID-19: an Immunopathological View. *mSphere.* 2020; 5. <https://doi.org/10.1128/mSphere.00344-20> PMID: 32321823
7. Ovsyannikova IG, Haralambieva IH, Crooke SN, Poland GA, Kennedy RB. The role of host genetics in the immune response to SARS-CoV-2 and COVID-19 susceptibility and severity. *Immunol Rev.* 2020; 296: 205–219. <https://doi.org/10.1111/immr.12897> PMID: 32658335
8. COVID-19 Host Genetics Initiative. Mapping the human genetic architecture of COVID-19. *Nature.* 2021; 600: 472–477. <https://doi.org/10.1038/s41586-021-03767-x> PMID: 34237774
9. Palmos AB, Chung R, Frissa S, Goodwin L, Hotopf M, Hatch SL, et al. Reconsidering the reasons for heightened inflammation in major depressive disorder. *J Affect Disord.* 2020. <https://doi.org/10.1016/j.jad.2020.12.109> PMID: 33422819
10. Monteiro AC, Suri R, Emeruwa IO, Stretch RJ, Cortes-Lopez RY, Sherman A, et al. Obesity and smoking as risk factors for invasive mechanical ventilation in COVID-19: A retrospective, observational cohort study. *PLoS One.* 2020; 15: e0238552. <https://doi.org/10.1371/journal.pone.0238552> PMID: 33351817
11. Sliz E, Kalaoja M, Ahola-Olli A, Raitakari O, Perola M, Salomaa V, et al. Genome-wide association study identifies seven novel loci associating with circulating cytokines and cell adhesion molecules in Finns. *J Med Genet.* 2019; 56: 607–616. <https://doi.org/10.1136/jmedgenet-2018-105965> PMID: 31217265
12. Ahola-Olli AV, Würtz P, Havulinna AS, Aalto K, Pitkänen N, Lehtimäki T, et al. Genome-wide association study identifies 27 loci influencing concentrations of circulating cytokines and growth factors. *Am J Hum Genet.* 2017; 100: 40–50. <https://doi.org/10.1016/j.ajhg.2016.11.007> PMID: 27989323
13. Höglund J, Rafati N, Rask-Andersen M, Enroth S, Karlsson T, Ek WE, et al. Improved power and precision with whole genome sequencing data in genome-wide association studies of inflammatory biomarkers. *Sci Rep.* 2019; 9: 16844. <https://doi.org/10.1038/s41598-019-53111-7> PMID: 31727947
14. Burgess S, Foley CN, Allara E, Staley JR, Howson JMM. A robust and efficient method for Mendelian randomization with hundreds of genetic variants. *Nat Commun.* 2020; 11: 376. <https://doi.org/10.1038/s41467-019-14156-4> PMID: 31953392
15. Burgess S, Davey Smith G, Davies NM, Dudbridge F, Gill D, Glymour MM, et al. Guidelines for performing Mendelian randomization investigations. *Wellcome Open Res.* 2020; 4: 186. <https://doi.org/10.12688/wellcomeopenres.15555.2> PMID: 32760811
16. Sun BB, Maranville JC, Peters JE, Stacey D, Staley JR, Blackshaw J, et al. Genomic atlas of the human plasma proteome. *Nature.* 2018; 558: 73–79. <https://doi.org/10.1038/s41586-018-0175-2> PMID: 29875488
17. Suhre K, Arnold M, Bhagwat AM, Cotton RJ, Engelke R, Raffler J, et al. Connecting genetic risk to disease end points through the human blood plasma proteome. *Nat Commun.* 2017; 8: 14357. <https://doi.org/10.1038/ncomms14357> PMID: 28240269
18. Wood AR, Perry JRB, Tanaka T, Hernandez DG, Zheng H-F, Melzer D, et al. Imputation of variants from the 1000 Genomes Project modestly improves known associations and can identify low-frequency variant-phenotype associations undetected by HapMap based imputation. *PLoS One.* 2013; 8: e64343. <https://doi.org/10.1371/journal.pone.0064343> PMID: 23696881
19. Folkersen L, Gustafsson S, Wang Q, Hansen DH, Hedman ÅK, Schork A, et al. Genomic and drug target evaluation of 90 cardiovascular proteins in 30,931 individuals. *Nat Metab.* 2020; 2: 1135–1148. <https://doi.org/10.1038/s42255-020-00287-2> PMID: 33067605
20. Folkersen L, Fauman E, Sabater-Lleal M, Strawbridge RJ, Frånberg M, Sennblad B, et al. Mapping of 79 loci for 83 plasma protein biomarkers in cardiovascular disease. *PLoS Genet.* 2017; 13: e1006706. <https://doi.org/10.1371/journal.pgen.1006706> PMID: 28369058
21. Bretherick AD, Canela-Xandri O, Joshi PK, Clark DW, Rawlik K, Boutin TS, et al. Linking protein to phenotype with Mendelian Randomization detects 38 proteins with causal roles in human diseases and traits. *PLoS Genet.* 2020; 16: e1008785. <https://doi.org/10.1371/journal.pgen.1008785> PMID: 32628676

22. Hillary RF, McCartney DL, Harris SE, Stevenson AJ, Seeboth A, Zhang Q, et al. Genome and epigenome wide studies of neurological protein biomarkers in the Lothian Birth Cohort 1936. *Nat Commun.* 2019; 10: 3160. <https://doi.org/10.1038/s41467-019-11177-x> PMID: 31320639
23. Enroth S, Maturi V, Berggrund M, Enroth SB, Moustakas A, Johansson A, et al. Systemic and specific effects of antihypertensive and lipid-lowering medication on plasma protein biomarkers for cardiovascular diseases. *Sci Rep.* 2018; 8. <https://doi.org/10.1038/s41598-018-23860-y> PMID: 29615742
24. MacArthur J, Bowler E, Cerezo M, Gil L, Hall P, Hastings E, et al. The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). *Nucleic Acids Res.* 2017; 45: D896–D901. <https://doi.org/10.1093/nar/gkw1133> PMID: 27899670
25. 1000 Genomes Project Consortium, The 1000 Genomes Project, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, et al. A global reference for human genetic variation. *Nature.* 2015; 526: 68–74. <https://doi.org/10.1038/nature15393> PMID: 26432245
26. Zhu Z, Zheng Z, Zhang F, Wu Y, Trzaskowski M, Maier R, et al. Causal associations between risk factors and common diseases inferred from GWAS summary data. *Nat Commun.* 2018; 9: 224. <https://doi.org/10.1038/s41467-017-02317-2> PMID: 29335400
27. Hübel C, Gaspar HA, Coleman JRI, Hanscombe KB, Purves K, Prokopenko I, et al. Genetic correlations of psychiatric traits with body composition and glycaemic traits are sex- and age-dependent. *Nat Commun.* 2019; 10: 5765. <https://doi.org/10.1038/s41467-019-13544-0> PMID: 31852892
28. Holmes MV, Asselbergs FW, Palmer TM, Drenos F, Lanktree MB, Nelson CP, et al. Mendelian randomization of blood lipids for coronary heart disease. *Eur Heart J.* 2015; 36: 539–550. <https://doi.org/10.1093/eurheartj/ehv571> PMID: 24474739
29. Censin JC, Nowak C, Cooper N, Bergsten P, Todd JA, Fall T. Childhood adiposity and risk of type 1 diabetes: A Mendelian randomization study. *PLoS Med.* 2017; 14: e1002362. <https://doi.org/10.1371/journal.pmed.1002362> PMID: 28763444
30. Wootton RE, Greenstone HSR, Abdellaoui A, Denys D, Verweij KJH, Munafò MR, et al. Bidirectional effects between loneliness, smoking and alcohol use: evidence from a Mendelian randomization study. *Addiction.* 2021; 116: 400–406. <https://doi.org/10.1111/add.15142> PMID: 32542815
31. Garipey G, Nitka D, Schmitz N. The association between obesity and anxiety disorders in the population: a systematic review and meta-analysis. *Int J Obes (Lond).* 2010; 34: 407–419. <https://doi.org/10.1038/ijo.2009.252> PMID: 19997072
32. Benjamini Y, Hochberg Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J R Stat Soc Series B Stat Methodol.* 1995; 57: 289–300.
33. Hemani G, Bowden J, Davey Smith G. Evaluating the potential role of pleiotropy in Mendelian randomization studies. *Hum Mol Genet.* 2018; 27: R195–R208. <https://doi.org/10.1093/hmg/ddy163> PMID: 29771313
34. Bowden J, Del Greco M F, Minelli C, Davey Smith G, Sheehan N, Thompson J. A framework for the investigation of pleiotropy in two-sample summary data Mendelian randomization. *Stat Med.* 2017; 36: 1783–1802. <https://doi.org/10.1002/sim.7221> PMID: 28114746
35. Teumer A. Common methods for performing Mendelian randomization. *Front Cardiovasc Med.* 2018; 5: 51. <https://doi.org/10.3389/fcvm.2018.00051> PMID: 29892602
36. Howe KL, Achuthan P, Allen J, Allen J, Alvarez-Jarreta J, Amode MR, et al. Ensembl 2021. *Nucleic Acids Res.* 2021; 49: D884–D891. <https://doi.org/10.1093/nar/gkaa942> PMID: 33137190
37. Durinck S, Spellman PT, Birney E, Huber W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat Protoc.* 2009; 4: 1184–1191. <https://doi.org/10.1038/nprot.2009.97> PMID: 19617889
38. Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, et al. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res.* 2001; 29: 308–311. <https://doi.org/10.1093/nar/29.1.308> PMID: 11125122
39. Get “SNP” (“Single-Nucleotide” “Polymorphism”) Data on the Web [R package rsnp version 0.4.0]. 2020 [cited 2 Nov 2021]. Available: <https://CRAN.R-project.org/package=rsnp>
40. Greenberg AS, Obin MS. Obesity and the role of adipose tissue in inflammation and metabolism. *Am J Clin Nutr.* 2006; 83: 461S–465S. <https://doi.org/10.1093/ajcn/83.2.461S> PMID: 16470013
41. Soeroto AY, Soetedjo NN, Purwiga A, Santoso P, Kulsum ID, Suryadinata H, et al. Effect of increased BMI and obesity on the outcome of COVID-19 adult patients: A systematic review and meta-analysis. *Diabetes Metab Syndr.* 2020; 14: 1897–1904. <https://doi.org/10.1016/j.dsx.2020.09.029> PMID: 33007661
42. Machiela MJ, Chanock SJ. LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. *Bioinformatics.* 2015; 31: 3555–3557. <https://doi.org/10.1093/bioinformatics/btv402> PMID: 26139635

43. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, et al. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *The Innovation*. 2021; 2: 100141. <https://doi.org/10.1016/j.xinn.2021.100141> PMID: 34557778
44. Bowden J, Del Greco M F, Minelli C, Davey Smith G, Sheehan NA, Thompson JR. Assessing the suitability of summary data for two-sample Mendelian randomization analyses using MR-Egger regression: the role of the I^2 statistic. *Int J Epidemiol*. 2016; 45: 1961–1974. <https://doi.org/10.1093/ije/dyw220> PMID: 27616674
45. Zhao Q, Wang J, Hemani G, Bowden J, Small DS. Statistical inference in two-sample summary-data Mendelian randomization using robust adjusted profile score. *Ann Stat*. 2020; 48: 1742–1769.
46. Bowden J, Del Greco M F, Minelli C, Zhao Q, Lawlor DA, Sheehan NA, et al. Improving the accuracy of two-sample summary-data Mendelian randomization: moving beyond the NOME assumption. *Int J Epidemiol*. 2019; 48: 728–742. <https://doi.org/10.1093/ije/dyy258> PMID: 30561657
47. Sumitran-Holgersson S. Beyond ABO and human histocompatibility antigen: other histocompatibility antigens with a role in transplantation. *Curr Opin Organ Transplant*. 2008; 13: 425–429. <https://doi.org/10.1097/MOT.0b013e328307ebd7> PMID: 18685340
48. Howell WM, Carter V, Clark B. The HLA system: immunobiology, HLA typing, antibody screening and crossmatching techniques. *J Clin Pathol*. 2010; 63: 387–390. <https://doi.org/10.1136/jcp.2009.072371> PMID: 20418230
49. Klein J, Sato A. The HLA system. First of two parts. *N Engl J Med*. 2000; 343: 702–709. <https://doi.org/10.1056/NEJM200009073431006> PMID: 10974135
50. Yengo L, Sidorenko J, Kemper KE, Zheng Z, Wood AR, Weedon MN, et al. Meta-analysis of genome-wide association studies for height and body mass index in ~ 700000 individuals of European ancestry. *Hum Mol Genet*. 2018; 27: 3641–3649. <https://doi.org/10.1093/hmg/ddy271> PMID: 30124842
51. Georgakis MK, Gill D, Rannikmäe K, Traylor M, Anderson CD, Lee J-M, et al. Genetically determined levels of circulating cytokines and risk of stroke. *Circulation*. 2019. pp. 256–268. <https://doi.org/10.1161/CIRCULATIONAHA.118.035905> PMID: 30586705
52. McGowan LM, Davey Smith G, Gaunt TR, Richardson TG. Integrating Mendelian randomization and multiple-trait colocalization to uncover cell-specific inflammatory drivers of autoimmune and atopic disease. *Hum Mol Genet*. 2019; 28: 3293–3300. <https://doi.org/10.1093/hmg/ddz155> PMID: 31276585
53. Yamamoto F. Molecular genetics of ABO. *Vox Sang*. 2000; 78 Suppl 2: 91–103. PMID: 10938936
54. Groot HE, Villegas Sierra LE, Said MA, Lipsic E, Karper JC, van der Harst P. Genetically determined ABO blood group and its associations with health and disease. *Arterioscler Thromb Vasc Biol*. 2020; 40: 830–838. <https://doi.org/10.1161/ATVBAHA.119.313658> PMID: 31969017
55. Mahmud R, Rassel MA, Monayem FB, Sayeed SKJB, Islam MS, Islam MM, et al. Association of ABO blood groups with presentation and outcomes of confirmed SARS CoV-2 infection: a prospective study in the largest COVID-19 dedicated hospital in Bangladesh. *PLoS One*. 2021; 16: e0249252. <https://doi.org/10.1371/journal.pone.0249252> PMID: 33826648
56. Zhao J, Yang Y, Huang H, Li D, Gu D, Lu X, et al. Relationship between the ABO blood group and the COVID-19 susceptibility. *Clin Infect Dis*. 2020; 73: 328–331.
57. Wu Y, Feng Z, Li P, Yu Q. Relationship between ABO blood group distribution and clinical characteristics in patients with COVID-19. *Clin Chim Acta*. 2020; 509: 220–223. <https://doi.org/10.1016/j.cca.2020.06.026> PMID: 32562665
58. Fan Q, Zhang W, Li B, Li D-J, Zhang J, Zhao F. Association between ABO blood group system and COVID-19 susceptibility in Wuhan. *Front Cell Infect Microbiol*. 2020; 10: 404. <https://doi.org/10.3389/fcimb.2020.00404> PMID: 32793517
59. Wu B-B, Gu D-Z, Yu J-N, Yang J, Shen W-Q. Association between ABO blood groups and COVID-19 infection, severity and demise: A systematic review and meta-analysis. *Infect Genet Evol*. 2020; 84: 104485. <https://doi.org/10.1016/j.meegid.2020.104485> PMID: 32739464
60. Liu N, Zhang T, Ma L, Zhang H, Wang H, Wei W, et al. The impact of ABO blood group on COVID-19 infection risk and mortality: A systematic review and meta-analysis. *Blood Rev*. 2020; 100785. <https://doi.org/10.1016/j.blre.2020.100785> PMID: 33309392
61. Muñiz-Diaz E, Llopis J, Parra R, Roig I, Ferrer G, Grifols J, et al. Relationship between the ABO blood group and COVID-19 susceptibility, severity and mortality in two cohorts of patients. *Blood Transfus*. 2021; 19: 54–63. <https://doi.org/10.2450/2020.0256-20> PMID: 33196417
62. Dean L. *Blood Groups and Red Cell Antigens*. 1st ed. Bethesda (MD): National Center for Biotechnology Information (US); 2005.
63. Bhandari S, Shaktawat A Singh, Tak A, Shukla J, Gupta J, Patel B, et al. Relationship between ABO blood group phenotypes and nCOVID-19 susceptibility: A retrospective observational study. *Scr medica*. 2020; 51: 217–239.

64. van der Vlist M, Geijtenbeek TBH. Langerin functions as an antiviral receptor on Langerhans cells. *Immunol Cell Biol.* 2010; 88: 410–415. <https://doi.org/10.1038/icb.2010.32> PMID: 20309013
65. Thépaut M, Luczkowiak J, Vivès C, Labiod N, Bally I, Lasala F, et al. DC/L-SIGN recognition of spike glycoprotein promotes SARS-CoV-2 trans-infection and can be inhibited by a glycomimetic antagonist. *PLoS Pathog.* 2021; 17: e1009576. <https://doi.org/10.1371/journal.ppat.1009576> PMID: 34015061
66. Trbojević-Akmačić I, Petrović T, Lauc G. SARS-CoV-2 S glycoprotein binding to multiple host receptors enables cell entry and infection. *Glycoconj J.* 2021; 38: 611–623. <https://doi.org/10.1007/s10719-021-10021-z> PMID: 34542788
67. Ortega FJ, Pueyo N, Moreno-Navarrete JM, Sabater M, Rodriguez-Hermosa JI, Ricart W, et al. The lung innate immune gene surfactant protein-D is expressed in adipose tissue and linked to obesity status. *Int J Obes.* 2013; 37: 1532–1538. <https://doi.org/10.1038/ijo.2013.23> PMID: 23478426
68. 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
69. He B, Garmire L. Prediction of repurposed drugs for treating lung injury in COVID-19. *F1000Res.* 2020; 9: 609. <https://doi.org/10.12688/f1000research.23996.2> PMID: 32934806
70. Islam ABMMK, Khan MA-A-K. Lung transcriptome of a COVID-19 patient and systems biology predictions suggest impaired surfactant production which may be druggable by surfactant therapy. *Sci Rep.* 2020; 10: 19395. <https://doi.org/10.1038/s41598-020-76404-8> PMID: 33173052
71. Mulugeta S, Beers MF. Surfactant protein C: its unique properties and emerging immunomodulatory role in the lung. *Microbes Infect.* 2006; 8: 2317–2323. <https://doi.org/10.1016/j.micinf.2006.04.009> PMID: 16782390
72. Nayak A, Dodagatta-Marri E, Tsolaki AG, Kishore U. An insight into the diverse roles of surfactant proteins, SP-A and SP-D in innate and adaptive immunity. *Front Immunol.* 2012; 3: 131. <https://doi.org/10.3389/fimmu.2012.00131> PMID: 22701116
73. Libby P, Luscher T. COVID-19 is, in the end, an endothelial disease. *Eur Heart J.* 2020; 41: 3038–3044. <https://doi.org/10.1093/eurheartj/ehaa623> PMID: 32882706
74. Bernimoulin MP, Zeng X-L, Abbal C, Giraud S, Martinez M, Michielin O, et al. Molecular basis of leukocyte rolling on PSGL-1. Predominant role of core-2 O-glycans and of tyrosine sulfate residue 51. *J Biol Chem.* 2003; 278: 37–47. <https://doi.org/10.1074/jbc.M204360200> PMID: 12403782
75. Mehta-D'souza P, Klopocki AG, Oganessian V, Terzyan S, Mather T, Li Z, et al. Glycan bound to the selectin low affinity state engages Glu-88 to stabilize the high affinity state under force. *J Biol Chem.* 2017; 292: 2510–2518. <https://doi.org/10.1074/jbc.M116.767186> PMID: 28011641
76. Koch AE, Halloran MM, Haskell CJ, Shah MR, Polverini PJ. Angiogenesis mediated by soluble forms of E-selectin and vascular cell adhesion molecule-1. *Nature.* 1995; 376: 517–519. <https://doi.org/10.1038/376517a0> PMID: 7543654
77. Nimrichter L, Burdick MM, Aoki K, Laroy W, Fierro MA, Hudson SA, et al. E-selectin receptors on human leukocytes. *Blood.* 2008; 112: 3744–3752. <https://doi.org/10.1182/blood-2008-04-149641> PMID: 18579791
78. Dasgupta B, Dufour E, Mamdouh Z, Muller WA. A novel and critical role for tyrosine 663 in platelet endothelial cell adhesion molecule-1 trafficking and transendothelial migration. *J Immunol.* 2009; 182: 5041–5051. <https://doi.org/10.4049/jimmunol.0803192> PMID: 19342684
79. Sachs UJH, Andrei-Selmer CL, Maniar A, Weiss T, Paddock C, Orlova VV, et al. The neutrophil-specific antigen CD177 is a counter-receptor for platelet endothelial cell adhesion molecule-1 (CD31). *J Biol Chem.* 2007; 282: 23603–23612. <https://doi.org/10.1074/jbc.M701120200> PMID: 17580308
80. Noels H, Weber C, Koenen RR. Chemokines as therapeutic targets in cardiovascular disease. *Arterioscler Thromb Vasc Biol.* 2019; 39: 583–592. <https://doi.org/10.1161/ATVBAHA.118.312037> PMID: 30760014
81. Sasaki M, Namioka Y, Ito T, Izumiyama N, Fukui S, Watanabe A, et al. Role of ICAM-1 in the aggregation and adhesion of human alveolar macrophages in response to TNF-alpha and INF-gamma. *Mediators Inflamm.* 2001; 10: 309–313. <https://doi.org/10.1080/09629350120102325> PMID: 11817671
82. Tong M, Jiang Y, Xia D, Xiong Y, Zheng Q, Chen F, et al. Elevated expression of serum endothelial cell adhesion molecules in COVID-19 patients. *J Infect Dis.* 2020; 222: 894–898. <https://doi.org/10.1093/infdis/jiaa349> PMID: 32582936
83. Nagashima S, Mendes MC, Camargo Martins AP, Borges NH, Godoy TM, Miggiolaro AFR dos S, et al. Endothelial dysfunction and thrombosis in patients with COVID-19—brief report. *Arterioscler Thromb Vasc Biol.* 2020; 40: 2404–2407. <https://doi.org/10.1161/ATVBAHA.120.314860> PMID: 32762443

84. Okai B, Lyall N, Gow NAR, Bain JM, Erwig L-P. Rab14 regulates maturation of macrophage phagosomes containing the fungal pathogen *Candida albicans* and outcome of the host-pathogen interaction. *Infect Immun*. 2015; 83: 1523–1535. <https://doi.org/10.1128/IAI.02917-14> PMID: 25644001
85. Hoffmann H-H, Sánchez-Rivera FJ, Schneider WM, Luna JM, Soto-Feliciano YM, Ashbrook AW, et al. Functional interrogation of a SARS-CoV-2 host protein interactome identifies unique and shared coronavirus host factors. *Cell Host Microbe*. 2021; 29: 267–280.e5. <https://doi.org/10.1016/j.chom.2020.12.009> PMID: 33357464
86. Cotroneo CE, Mangano N, Dragani TA, Colombo F. Lung expression of genes putatively involved in SARS-CoV-2 infection is modulated in cis by germline variants. *Eur J Hum Genet*. 2021; 29: 1019–1026. <https://doi.org/10.1038/s41431-021-00831-y> PMID: 33649539
87. Martin V, Bredoux R, Corvazier E, Van Gorp R, Kovacs T, Gelebart P, et al. Three novel sarco/endoplasmic reticulum Ca²⁺-ATPase (SERCA) 3 isoforms. Expression, regulation, and function of the membranes of the SERCA3 family. *J Biol Chem*. 2002; 277: 24442–24452. <https://doi.org/10.1074/jbc.M202011200> PMID: 11956212
88. Bobe R, Bredoux R, Corvazier E, Andersen JP, Clausen JD, Dode L, et al. Identification, expression, function, and localization of a novel (sixth) isoform of the human sarco/endoplasmic reticulum Ca²⁺-ATPase 3 gene. *J Biol Chem*. 2004; 279: 24297–24306. <https://doi.org/10.1074/jbc.M314286200> PMID: 15028735
89. 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
90. Akhmerov A, Marbán E. COVID-19 and the Heart. *Circ Res*. 2020; 126: 1443–1455. <https://doi.org/10.1161/CIRCRESAHA.120.317055> PMID: 32252591
91. Shi S, Qin M, Shen B, Cai Y, Liu T, Yang F, et al. Association of Cardiac Injury With Mortality in Hospitalized Patients With COVID-19 in Wuhan, China. *JAMA Cardiol*. 2020; 5: 802–810. <https://doi.org/10.1001/jamacardio.2020.0950> PMID: 32211816
92. Dally S, Corvazier E, Bredoux R, Bobe R, Enouf J. Multiple and diverse coexpression, location, and regulation of additional SERCA2 and SERCA3 isoforms in nonfailing and failing human heart. *J Mol Cell Cardiol*. 2010; 48: 633–644. <https://doi.org/10.1016/j.yjmcc.2009.11.012> PMID: 19962989
93. Radhakrishnan P, Cheng PW. GCNT3 (glucosaminyl (N-acetyl) transferase 3, mucin type). *Atlas Genet Cytogenet Oncol Haematol*. 2011. <https://doi.org/10.4267/2042/38541>
94. Grant OC, Montgomery D, Ito K, Woods RJ. Analysis of the SARS-CoV-2 spike protein glycan shield reveals implications for immune recognition. *Sci Rep*. 2020; 10: 14991. <https://doi.org/10.1038/s41598-020-71748-7> PMID: 32929138
95. Bose M, Mitra B, Mukherjee P. Mucin signature as a potential tool to predict susceptibility to COVID-19. *Physiol Rep*. 2021; 9: e14701. <https://doi.org/10.14814/phy2.14701> PMID: 33373502
96. Ju T, Cummings RD. A unique molecular chaperone Cosmc required for activity of the mammalian core 1 beta 3-galactosyltransferase. *Proc Natl Acad Sci U S A*. 2002; 99: 16613–16618. <https://doi.org/10.1073/pnas.262438199> PMID: 12464682
97. Breiman A, Ruvoën-Clouet N, Deleers M, Beauvais T, Jouand N, Rocher J, et al. Low levels of natural anti- α -N-acetylgalactosamine (Tn) antibodies are associated with COVID-19. *Front Microbiol*. 2021; 12: 641460. <https://doi.org/10.3389/fmicb.2021.641460> PMID: 33643275
98. Lardone RD, Garay YC, Parodi P, de la Fuente S, Angeloni G, Bravo EO, et al. How glycobiology can help us treat and beat the COVID-19 pandemic. *J Biol Chem*. 2021; 296: 100375. <https://doi.org/10.1016/j.jbc.2021.100375> PMID: 33548227
99. Kaczocha M. [Thesis]. “Role of Fatty Acid Binding Proteins and FAAH-2 in Endocannabinoid Uptake and Inactivation.” Edited by Dale G. Deutsch. Ann Arbor, United States: State University of New York at Stony Brook. Available: <https://www.proquest.com/dissertations-theses/role-fatty-acid-binding-proteins-faah-2/docview/305091515/se-2>.
100. Lucaciu O, Aghiorghiesei O, Petrescu NB, Mirica IC, Benea HRC, Apostu D. In quest of a new therapeutic approach in COVID-19: the endocannabinoid system. *Drug Metab Rev*. 2021; 1–13. <https://doi.org/10.1080/03602532.2021.1895204> PMID: 33683968
101. Rizzo MD, Henriquez JE, Blevins LK, Bach A, Crawford RB, Kaminski NE. Targeting Cannabinoid Receptor 2 on Peripheral Leukocytes to Attenuate Inflammatory Mechanisms Implicated in HIV-Associated Neurocognitive Disorder. *J Neuroimmune Pharmacol*. 2020; 15: 780–793. <https://doi.org/10.1007/s11481-020-09918-7> PMID: 32409991
102. Costiniuk CT, Jenabian M-A. Acute inflammation and pathogenesis of SARS-CoV-2 infection: cannabidiol as a potential anti-inflammatory treatment? *Cytokine Growth Factor Rev*. 2020; 53: 63–65. <https://doi.org/10.1016/j.cytogfr.2020.05.008> PMID: 32467020

103. Rossi F, Tortora C, Argenziano M, Di Paola A, Punzo F. Cannabinoid Receptor Type 2: a Possible Target in SARS-CoV-2 (CoV-19) Infection? *Int J Mol Sci.* 2020; 21. <https://doi.org/10.3390/ijms21113809> PMID: 32471272
104. Menten P, Wuyts A, Van Damme J. Macrophage inflammatory protein-1. *Cytokine Growth Factor Rev.* 2002; 13: 455–481. [https://doi.org/10.1016/s1359-6101\(02\)00045-x](https://doi.org/10.1016/s1359-6101(02)00045-x) PMID: 12401480
105. Li M, Yeung CHC, Schooling CM. Circulating cytokines and Coronavirus disease: a bi-directional Mendelian randomization study. *Front Genet.* 2021; 12: 680646. <https://doi.org/10.3389/fgene.2021.680646> PMID: 34163532
106. Tan Y, Tang F. SARS-CoV-2-mediated immune system activation and potential application in immunotherapy. *Med Res Rev.* 2021; 41: 1167–1194. <https://doi.org/10.1002/med.21756> PMID: 33185926
107. Mohammed El Tabaa M, Mohammed El Tabaa M. Targeting Neprilysin (NEP) pathways: A potential new hope to defeat COVID-19 ghost. *Biochem Pharmacol.* 2020; 178: 114057. <https://doi.org/10.1016/j.bcp.2020.114057> PMID: 32470547