medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20229203; this version posted May 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

The genetic risk for COVID-19 severity is associated with defective innate immune responses

Yunus Kuijpers[1,2,+], Xiaojing Chu[1,2,3,+], Martin Jaeger[4], Simone J.C.F.M. Moorlag[4], Valerie A.C.M. Koeken[1,2,4], Bowen Zhang[1,2,+], Aline de Nooijer[4], Inge Grondman[4], Nico Janssen[4], Vera P. Mourits[4], L. Charlotte J. de Bree[4], Quirijn de Mast[4], Frank L. van de Veerdonk[4], Leo A.B. Joosten[4,5], Yang Li[1,2,4], Mihai G. Netea[4,6,‡*] and Cheng-Jian Xu[1,2,4,7‡*]

[1]Centre for Individualised Infection Medicine, CiiM, a joint venture between the Hannover Medical School and the Helmholtz Centre for Infection Research, Hannover, Germany

[2]TWINCORE, Centre for Experimental and Clinical Infection Research, a joint venture between the Hannover Medical School and the Helmholtz Centre for Infection Research, Hannover, Germany

[3]Department of Genetics, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands

[4]Department of Internal Medicine and Radboud Institute for Center for Infectious Diseases, Radboud University Medical Center, Nijmegen, the Netherlands

[5]Núcleo de Pesquisa da Faculdade da Polícia Militar (FPM) do Estado de Goiás, Goiânia, Goiás, Brasil

[6]Department for Genomics & Immunoregulation, Life and Medical Sciences Institute (LIMES), University of Bonn, Bonn, Germany

[7]Department of Gastroenterology, Hepatology and Endocrinology, Hannover Medical School, Hannover, Germany

[†]shared first authors; these authors contribute equally.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20229203; this version posted May 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

‡shared senior authors: these authors contribute equally.

*Corresponding authors:

Mihai.Netea@radboudumc.nl; Xu.Chengjian@mh-hannover.de

One Sentence summary

In this study, we explore the physiological significance of the genetic variants associated with COVID-19 severity using detailed clinical, immunological and multi-omics data from large cohorts. Our findings allow a physiological understanding of genetic susceptibility to severe COVID-19, and indicate pathways that could be targeted for prevention and therapy.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20229203; this version posted May 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Abstract

Recent genome-wide association studies (GWASs) of COVID-19 patients of European ancestry have identified genetic loci significantly associated with disease severity *(1)*. Here, we employed the detailed clinical, immunological and multi-omics dataset of the Human Functional Genomics Projects (HFGP) to explore the physiological significance of the host genetic variants that influence susceptibility to severe COVID-19. A genomics investigation intersected with functional characterization of individuals with high genetic risk for severe COVID-19 susceptibility identified several major patterns: i. a large impact of genetically determined innate immune responses in COVID-19, with increased susceptibility for severe disease in individuals with defective monocyte-derived cytokine production; ii. genetic susceptibility related to ABO blood groups is probably mediated through the von Willebrand factor (VWF) and endothelial dysfunction. We further validated these identified associations at transcript and protein levels by using independent disease cohorts. These insights allow a physiological understanding of genetic susceptibility to severe COVID-19, and indicate pathways that could be targeted for prevention and therapy.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20229203; this version posted May 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Introduction

The novel coronavirus disease 2019 (COVID-19), caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) *(2)(3)*, firstly emerged in late December 2019 and has been spreading worldwide very quickly. The COVID-19 pandemic creates a severe disruption to the healthcare system and endangers the economy. While much has been learned about the pathophysiology of the disease, and immune-based treatment proven to be effective has included dexamethasone *(4)* and anti-IL-6/IL-6R monoclonal antibodies*(5)(6)*, the prognosis is still poor in many patients. Therefore, there is an urgent need to better understand the exact host-pathogen interactions leading to increased severity and mortality, in order to design additional prophylactic and therapeutic strategies in future *(7)(8)*.

The severity of SARS-CoV-2 infection is highly variable, and ranges from asymptomatic to mild disease, and even to severe Acute Respiratory Distress Syndrome with a fatal outcome *(9)*. However, the causes for this broad variability in disease outcome between individuals are largely unknown. A recent study indicates that human host factors rather than viral genetic variation affect COVID-19 severity outcome *(10)*. Additionally, clinical and epidemiological data have shown that old age, male sex, and chronic comorbidity are associated with higher mortality *(11)(12)*. The first genome-wide association study in individuals with genetic European ancestry has identified several chemokine receptor genes, including *CCR9*, *CXCR6* and *XCR1* and the locus controlling the ABO blood type to be associated with severe symptoms of COVID-19 *(1)*. Nevertheless, little is known about the mechanisms through which these genetic variants influence COVID-19 severity. For example, several competing hypotheses may be envisaged for the involvement of immune genes in susceptibility to severe COVID-19: on the one hand, it may be hypothesized that genetic risk for severe COVID-19 is associated with defective innate immune responses that would allow viral multiplication with high viral

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20229203; this version posted May 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

loads; on the other hand, the opposite hypothesis may also be true, with an exaggerated genetically-mediated cytokine production being responsible for the late phase hyperinflammation and poor outcome. A purely genetic study cannot respond to this crucial question, that would have important consequences for the approach to prophylaxis and therapy.

By making use of resources from the Human Functional Genomics Project (HFGP) *(13, 14)*, we assessed the impact of COVID-19 associated genetic polymorphisms on variability of immune responses at the population level (Fig. 1), and we validated our findings using single cell transcriptomics and proteomics data from two independent COVID-19 cohorts. This study will help us to understand how genetic variability is related to disease susceptibility through the regulation of immune responses and endothelial function.

Results

COVID-19 loci are enriched for expression in immune organs, chemokine signaling pathways and enhancer region

To explore the functional impact of the identified COVID-19 loci, we firstly investigated if the identified independent genetic loci (p value $<1\times10^{-5}$) from the first COVID-19 GWAS study by Severe COVID-19 GWAS Group *(1)* are associated with any phenotypes available at the GWAS catalog (https://www.ebi.ac.uk/gwas/). We found that many of these loci are associated with traits such as blood proteins/biomarkers, LDL and VLDL concentrations (Table S1). We next performed functional annotation of significant loci and gene-mapping using Functional Mapping and Annotation (FUMA) *(15)*. The SNP2GENE function identified 32 independent SNPs located in 26 different loci which reached suggestive significance in this study (p-value $<1\times10^{-5}$, Table S2). Using multiple independent expression quantitative trait loci (eQTL)

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20229203; this version posted May 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

datasets, FUMA mapped 115 genes to these 26 genomic risk loci. Using RNA-seq data of 30 tissues from GTEx database (v8), we found significant enrichment of candidate genes in expression in immune organs such as spleen and blood (Fig. 2A), suggesting that they are important tissues contributing to the pathophysiology of COVID-19 *(16)(17)*. Moreover, we observed the enrichment of candidate genes to be mainly expressed in small intestine and lung (Fig. 2A), suggesting that COVID-19 represents a multisystem illness with involvement of different organs, consistent with the respiratory and intestinal symptoms of the disease *(18)*.

Pathway analysis using these 115 genes showed a strong enrichment in chemokine binding and chemokine receptors binding (Fig. S1), which is in line with the fact that chemokines can recruit immune cells to the site of infection and are critical for the function of the immune response *(19)*. In addition, chemokines have been reported as the most significantly elevated biomarkers in patients with severe COVID-19 on the intensive care unit *(17)*.

Considering that all SNPs in Linkage Disequilibrium (LD) with the 32 independent loci (p-value $<1\times10^{-5}$) identified by the COVID-19 GWAS, were significantly enriched in the non-coding intronic region (p value = 0.036, Fisher's exact test) (Table S3), we next examined whether the COVID-19 associated variants are enriched in regulatory DNA elements. We interrogated all significant SNPs (p-value$<1\times10^{-5}$ and stricter thresholds) with histone marks and chromatin states of 24 blood cell-types in the Roadmap Epigenome Project *(20)*. We found that the COVID-19 genetic loci were strongly enriched for enhancer markers and weakly enriched in promoter marker (Table S4). The strong enrichment of COVID-19 loci in enhancer marks indicates that the associated genetic variants are likely to be involved in the regulation of immunologically-related functions. This finding also suggests that epigenetic mechanisms/regulation may play an important role in the pathogenesis of COVID-19 infection.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20229203; this version posted May 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

3p21.31 loci are associated with lower production of monocyte-derived cytokines

Severe COVID-19 is characterized by complex immune dysregulation, combining immune defective features with hyperinflammatory innate immune traits *(21)(22)*. However, these analyses in patients could be performed only late during the disease, and whether genetic risk for severe COVID-19 is characterized by low or high innate immune responses in a non-infected person is not known. We therefore used the cytokine QTL data from the 500FG cohort *(14)* of the HFGP to test whether SNPs in 3p21.31 influence cytokine production upon stimulation. We checked all SNPs located within a 50 kilobase window of top variant rs11385942, and showed all nominal significant associations (p-value < 0.05, Fig. 2A). Interestingly, we observed that the risk alleles for a severe course of COVID-19 are consistently associated with lower production of monocyte-dependent cytokines (IL-6, IL-1β and TNF-α) upon various *in-vitro* stimulations (Fig. 2B). Of note, COVID-19 risk alleles also correspond to lower monocyte-derived cytokine production after influenza stimulation, a viral stimulus (Fig. 2B and C, Table S5). It is thus tempting to speculate that the people who carry risk alleles may not respond properly to an initial virus infection, leading to high viral loads, subsequent systemic inflammation and poor outcome. Next, we tested whether the COVID-19 risk SNPs are associated with the concentrations of circulating cytokines and levels of metabolites in blood (Table S6-7). Using the same cohort, we found that IL-18 and IL-18BP concentrations show a suggestive positive association with genetic risk of COVID-19 (Fig. S2, Table S6). Additionally, one of the 3p21.31 loci , rs2191031 is suggestively negative associated with high density lipoprotein cholesterol (HDL) (P value =0.004, Table S7-S8), which is consistent with previous findings that HDL levels were significantly lower in the severe COVID-19 disease group *(23)*.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20229203; this version posted May 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Von Willebrand Factor (VWF) and lymphocytes are strongly colocalized with ABO loci

It is known that ABO blood group influences the plasma levels of von Willebrand factor (VWF)*(24)* and elevated VWF levels are associated with severe COVID-19 *(25)*. We therefore tested the association of ABO locus with VWF circulating concentrations from the individuals in the 500FG cohort. Of note, we found the risk allele rs687621-G is significantly associated with elevated levels of VWF (p-value = 9.58×10^{-20}) (Fig. 3A and 3B). Recent studies have reported that the VWF level is highly related to COVID-19 severity *(26, 27)*. As VWF level in plasma is an indicator of inflammation, endothelial activation and damage *(28)*, our results suggest that the association of VWF and COVID-19 severity is very likely mediated through genetic regulation.

We next tested if this specific locus is associated with immune functions. Interestingly, we observed consistent negative correlation of VWF and T-cell derived cytokine production in response to various *ex-vivo* stimulations (Fig. 3C and Fig. S3). In addition, the ABO locus led by the variant rs687621 also showed statistically significant co-localization with several immune-mediated traits, including cell counts of lymphocytes (Coloc analysis H4: 0.98), monocytes (Coloc analysis H4: 1), neutrophils (Coloc analysis H4: 0.8), and whole blood cells (Coloc analysis H4: 1) (Fig. 3D).

Association between genetic risk score of COVID-19 severity with gender and BMI

Polygenic risk scores (PRS) combine multiple risk alleles and capture an individual's load of common genetic variants associated with a disease phenotype *(29)*. Using the summary statistics provided in the GWAS study *(1)*, we calculated the PRS for the samples from 500FG and 300BCG cohorts. Since higher mortality of COVID-19 has been reported to be associated

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20229203; this version posted May 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

with male sex and BMI *(11, 12)*, we investigated whether these host factors are associated with the PRS, a predictive measure of risk for development of severe COVID-19.

We firstly assessed if males have a higher genetic risk compared to females in 500FG. Hereby, we defined people with top 10% PRS as a high-risk group and those with bottom 10% PRS as a low-risk group. As shown in Fig. 4A, male tend to have higher severe COVID-19 risk than female (odd ratio: 1.47, 95% CI: 0.98-2.22, p-value = 0.045 (Fisher's exact test)) (Table S9). We next used different percentile cut-offs (15%, 20%, 25% and 30%) to re-define low and high-risk groups. Interestingly, we observed a consistent pattern that males have higher genetic risk (PRS) than females at different percentile cut-offs. These results can be replicated in a similar, but independent, cohort (300BCG cohort, Fig. 4B). Meanwhile, the genetic risk difference between male and female can be attenuated when a loose cut-off has been defined. The meta-analysis of two cohorts showed a significant p-value at various percentile cut-offs (10%, 15%, 20%, 25%) and marginal significant p-value of 0.051 at the percentile cut-off of 30%. Furthermore, this result persisted when PRS was computed using summary statistics from the GWAS model after age and sex correction, reported in the original GWAS study (Table S10). However, when excluding variants from the X and Y chromosomes in the PRS calculation, the enrichment of men in the higher PRS group was not significant anymore (Fig. 4C and 4D). This suggested that higher genetic severity risk at least partially originates from the SNPs in the sex chromosomes.

As obesity or overweight has been reported as a risk factor for serious illness or death from COVID-19, we tested if the PRS is associated with BMI (Fig. 4E and 4F). We did not observe any significant correlation between PRS and BMI.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20229203; this version posted May 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Association between severity risk factors and cytokine response was replicated in a second GWAS study

To investigate whether our findings are only specific to one cohort, we implemented the same functional analyses in the recent published GenOMICC GWAS study*(30)*. We firstly run FUMA analyses with the same parameter settings, and the 135 mapped genes show significant enrichment in expression in spleen, blood, lung and small intestine (Fig. S4), which is consistent with our interpretation of the function of the genetic loci on immune organs. Next, we sought to test whether COVID-19 GWAS SNPs identified in GenOMICC studies influence cytokine production upon stimulation. We again observed risk alleles from 3p21.31 loci are consistently associated with lower production of monocyte-dependent cytokines (IL-6, IL-1β and TNF-α) upon various *in-vitro* stimulations (Fig. 5A and Table S11). Table S11 listed 816 normally significantly stimulation pairs SNP- cytokine (p value <0.05). Notably, 777 out of 816 SNP- cytokine stimulation pairs are from 3p21.31 loci. As 3p21.31 loci is the major genetic risk factor explaining severity outcome(https://www.covid19hg.org), our result may indicate a potential mechanism of major genetic factor impacting the severity of COVID-19 through an innate immune response.

Validation in single-cell transcriptomes and proteomics data of COVID-19 cohorts

To investigate in which cells those COVID-19 GWAS genes expressed, we utilized single-cell transcriptomics data from a COVID-19 German Berlin cohort*(31)* to illustrate the cell type-specific expression of COVID-19 risk genes (Fig. 5B). Among 37 COVID-19 risk genes, 10 were significantly differential expressed between severe and mild COVID-19. In general, the people with mild symptoms show high expression of those risk genes that those with severe symptoms (p-value $_{\chi 2test}$ $=6.9\times10^{-13}$) (Fig.S5). Interestingly, *OAS1* (2'-5'-Oligoadenylate Synthetase 1) shows higher expression in mild symptoms patients, which agrees with the recent

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20229203; this version posted May 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

findings that higher plasma OAS1 levels were associated with reduced COVID-19 severity*(32)* . We next tested the expression of monocyte-derived and T cell-derived cytokine genes in single-cell transcriptomics data. We found that monocyte-derived cytokine genes are expressed in patients, while T cell-derived cytokine genes are not, which indicates that innate immune response is of importance in disease development (Fig.5C). Therefore, we further tested the longitudinal expression change of *IL-1β* and *TNF-α* using the data from the same cohort (Fig. S6). Notedly, we found that patients with mild symptoms usually show a decreasing pattern of expression in cytokines (*IL-1β*, *TNF-α*) during the disease course (from high to low), while the severe symptom patients show an increasing pattern of expression in cytokine genes (*IL-1β*, *TNF-α*) with time (from low to high). Such a pattern may suggest that mild patients have a sufficiently good response in the first phase of the infection, which inhibits efficiently the viral replication, subsequently resulting in decreased inflammation. On the contrary, severe patients show defective activation of antiviral innate immunity in the beginning, leading to prolongation of the disease process with inefficient and deleterious inflammatory response in the end. This observation is in line with the recent findings of immunosuppression in severely affected patients*(33)*.

We next validate whether two inflammatory proteins CCL25 and CXCL16, which are ligands for Chromosome 3 GWAS gene *CCR9* and *CXCR6*, respectively, are associated with disease severity. We measured CCL25 and CXCL16 protein levels in two different patient cohorts by using Olink platform technology and ELISA, respectively (Methods). We observed a significantly lower expression of CCL25 in ICU patients (P-value = 0.0255) (Fig. 5D), while a significantly higher expression of CXCL16 in ICU patient (P-value <0.0001) (Fig. 5E and Fig. S7). These results are in line with the reported protective effect of *CXCR6* and the risk effect of *CCR9* in transcriptomic regulation *(34)*. Interestingly, the longitudinal data from

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20229203; this version posted May 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

COVID-19 patients showed that a clinical improvement is often associated with increased CCL25 concentration (Fig. S8).

Additionally, we measured VWF levels in 159 COVID-19 patients (Methods) and found that VWF levels were significantly higher in the ICU patients (Fig.5F and Fig.S9). This further supported our findings that the ABO genetic risk is associated with VWF levels.

Discussion

Understanding the pathophysiology of COVID-19 is urgently needed for designing novel preventive and therapeutic approaches against the disease. One important tool for identifying the most important mechanisms mediating severity of a disease process is genomics: genetic variants that influence susceptibility or severity to a disease are usually located in genetic loci that impact important mechanisms for that particular disease. Using the information of a recently published GWAS assessing the severity of COVID-19 *(1)*, and the rich datasets available in the HFGP, we interrogated the mechanisms through which genetic variants associated with severe COVID-19 exert their effects.

Among the genetic loci associated with severe COVID-19, the 3p21.31 gene cluster has been well replicated by independent studies from the COVID-19 Host Genetics Initiative (https://www.covid19hg.org), and it was reported to be inherited by Neanderthals *(35)*. This locus is currently regarded as a marker of COVID-19 severity, but crucial information is missing: are the risk alleles in this locus (that encode several cytokines and chemokines) associated with a lower or higher cytokine production? The answer to this question is crucial for understanding COVID-19: a genetic risk associated with low cytokine production would imply that severe COVID-19 is the consequence of a relative immunodeficiency, while a high

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20229203; this version posted May 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

cytokine production associated with genetic risk would mean that severe COVID-19 is a genetic hyperinflammatory disease. In our study, the 3p21.31 genetic polymorphisms associated with a high risk of severe COVID-19 were associated with lower production of monocyte-derived cytokines, especially to viral (influenza) stimuli. This important discovery has significant prophylactic and therapeutic consequences. On the one hand, it implies that improvement of innate immune responses in healthy individuals would decrease the probability that they undergo a severe form of COVID-19: this supports the rationale of clinical trials that improve innate immune responses through induction of trained innate immunity *(7)*. On the other hand, this also implies that the dysregulated immune responses that have been described at late time points in patients with severe COVID-19 *(31)(36)* are likely the consequence of accelerated viral multiplication due to defective innate immune responses, and subsequent systemic inflammation due to high viral loads.

Several studies have shown that ABO blood types are associated with COVID-19 severity *(37)(38)* and susceptibility *(39)(40)(41)*. It is still not well-known how the ABO locus regulates COVID-19 susceptibility. As ABO blood group are also expressed on endothelial cells and platelets, it has been speculated that this effect may manifest itself via elevated plasma VWF *(42)*. Our results provide evidence supporting this hypothesis, by showing that the risk alleles in the ABO locus are associated with high concentrations of VWF. Moreover, interesting associations have been found between polymorphisms in this locus and the number of various immune cell populations, especially lymphocytes, since lymphopenia is also consistently associated with severe COVID-19 *(43)*. This suggests that genetic factors are relevant to the host thrombo-inflammatory response. However, a note of caution should be mentioned, as the association between the genetics of ABO group with severity in COVID-19 Host Genetics

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20229203; this version posted May 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Initiative data did not reach a genome-wide level of significance (p value $<5\times10^{-8}$) (Table S12) (as of 21st of October 2020), and thus the association might be population specific.

Another observation is that the impact of genetic polymorphisms on the severity of COVID-19 is likely mediated through sex chromosomes, i.e. chromosome X, which is known to encode many genes related to the immune system. Indeed, men in both 500FG and 300BCG cohorts had a higher genetic risk than women, and this difference was largely lost when sex chromosomes were excluded from the PRS analysis. This finding was not able to be replicated in GenOMICC study*(30)*. (Table S13), due to fact that no significant association were reported on sex chromosomes in the other GenOMICC study. Therefore, these data suggest that at least part of the well-known increase of COVID-19 severity in men is genetically determined. The recent description by our group of rare mutations in the RNA receptor TLR7 located on chromosome X as a cause of very severe COVID-19 in young men also supports this hypothesis *(44)*.

While our study sheds further light on how COVID-19 genetic risk affects the human immune system, there are several limitations of this study: firstly, due to different sets of stimuli used in measuring cytokine production to stimulations in the two healthy cohorts, we are not able to replicate all our findings of genetic associations with cytokine responses from the 500FG cohort in the 300BCG cohort. Secondly, young adults (< 30 years) and normal weight BMI are overrepresented in both healthy cohorts (500FG and 300BCG), which may lead to a biased conclusion which cannot be generalized to the whole population, especially since the severe COVID-19 cases often occur in the elderly population.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20229203; this version posted May 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Collectively, our data demonstrate that genetic variability explains an important component of the increased susceptibility to severe COVID-19. The genetic risk for severe COVID-19 is associated with defective innate immune responses (low cytokine production), dysregulated endothelial function. These findings may contribute to the development of novel treatment and prevention strategies for severe COVID-19.

Materials and Methods

Study cohort

The cohorts involved in this study are from the Human Functional Genomics Project (HFGP)*(45)*. 500FG consists of 451 healthy individuals of European ancestry with genotype measurement. Within this cohort, immune cell counts, cytokine production upon stimulations, platelets, globulins, and gut microbiome were measured (for detailed information see *(13, 14, 46, 47)*). 300BCG consists of 313 healthy Europeans that participated in a BCG vaccination study *(48)(49)*. The basic characteristics of study populations are shown in Table S14. Within this cohort, blood was collected before vaccination and cytokine production was measured upon *ex-vivo* stimulation of PBMCs with microbial stimuli.

Genotype quality control and imputation

Genotyping on samples from 500FG and 300BCG was performed using Illumina humanOmniExpress Exome-8 v1.0 SNP chip Calling by Opticall 0.7.0*(50)* with default settings. All individuals of non-European ancestry, ambiguous sex, call rate ≤ 0.99, excess of autosomal heterozygosity (F<mean-3SD), cryptic relatedness (π>0.185) were removed. SNPs with low genotyping rate (<95%), with low minor allele frequency (<0.001), deviation from Hardy-Weinberg equilibrium ($p<10^{-4}$) were excluded. The detailed QC steps have been published in reference*(14)*. Genotype data of 500FG and 300BCG were imputed respectively. The

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20229203; this version posted May 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

imputation was performed on the Michigan imputation server*(51)*. The cohorts were phased using Eagle v2.4 with the European population of HRC 1.1 2016 hg 2019 reference panel. After imputation, variants with a MAF < 0.01, an imputation quality score R2 < 0.5, or a Hardy-Weinberg-Equilibrium $P < 10^{-12}$ were excluded. All quality control steps were performed using Plink v1.9. After imputation and quality control, 451 individuals from 500FG and 313 individuals from 300BCG were available for downstream analyses.

Immune parameter quantitative trait locus (QTL) profiles

We acquired summary statistics of cytokine QTLs *(14)*, cell proportion QTLs*(46)* and circulating mediators and metabolite QTLs*(52)* from our previous studies performed with 500FG. Metabolites were measured on the Brainshake Metabolomics/Nightingale Health metabolic platform. These samples were processed following the automated standard protocol provided by Nightingale's technology (Finland), and blood metabolites were quantified in absolute concentrations (e.g. mmol/L) and percentages using nuclear magnetic resonance (NMR) spectroscopy.We performed QTL mapping for circulating mediators and platelet traits in 500FG using an R package MatrixeQTL*(53)*. The measurement of circulating mediators including IL-18BP, resistin, leptin, adiponectin, alpha-1 antitrypsin (AAT), and IL-18 have been described previously*(13)*. Platelet traits*(54)* include Thrombin-Antithrombin Complex (TAT), Beta-thromboglobulin total, beta-thromboglobulin, fibrinogen binding, collagen-related peptide (CRP) P-selectin, CRP fibrinogen, ADP P-selectin, ADP fibrinogen, P-selectin, platelet–monocyte complex, total platelet count, and von Willebrand factor (VWF). The circulating mediator levels and platelet traits were log2 transformed. A linear model was applied to the platelet data and genetic data by taking age and sex as covariates. We considered p-value $< 5\times10^{-8}$ to be genome-wide significant.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20229203; this version posted May 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Colocalization analysis.

We performed co-localization analysis *(55)* to look at the overlapping profile between molecular QTLs, COVID-19 GWAS, and other GWAS profiles using the R package 'coloc'.

PRS calculation

Polygenic risk scores (PRS) were calculated by first intersecting the variants from the COVID-19 summary statistics*(1)* with the variants present in our samples. Clumping was done starting at the most significant variant. All variants within a 250kb window around that variant were excluded if they were in greater LD than 0.1 before continuing to the most significant variant outside of the previous window. For each sample specifically, we then multiplied the dosage of the effect allele with its effect size while substituting missing genotype data with the average dosage of that variant in the entire sample set. These values were then summed to form the PRS for each specific sample. As the GWAS summary statistics for creating PRS from Eillinghaus *et.al (1)* did not correct age and sex, we also performed a sensitivity analysis with the PRS created from the GWAS model corrected for age and sex.

PRS based correlations

Linear models were constructed using the computed PRS and various phenotype data available for each cohort. Samples within the top/bottom 10% PRS were classified as high/low-risk, respectively. Using the PRS of the samples in these risk groups, we performed a Student T-test to test for significant correlation between gender and PRS. Furthermore, we tested for enrichment of any specific gender in these risk groups using a Fisher's exact test.

Functional analysis of genomic loci

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20229203; this version posted May 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

We used the FUMA pipeline in order to identify genes linked to COVID-19 with severe respiratory failure. FUMA identified significant independent SNPs as variants with $P < 1\times10^{-5}$ that were independent from each other using an LD threshold of $r2 < 0.6$. Within these independent significant SNPs variants lead SNPs are identified as the most significant variants that are independent using an LD threshold of $r2 < 0.1$. We mapped Genes to these SNPs based on their genomic position allowing for a maximum distance of 10kb. In addition to this, genes were also mapped based on eQTL effects. Genes were selected based on significant SNP-gene pairs at FDR < 0.05 using cis- and trans-eQTLs from eQTLGen (https://www.eqtlgen.org).

As part of the FUMA pipeline we used these mapped genes in order to generate gene expression heatmaps using GTEx v8 (54 tissue types and 30 general tissue types). Gene expression values with a pseudocount of 1 were normalized across tissue types using winsorization at 50 and log2 transformed. Using the hypergeometric test, we tested for significant enrichment of our input genes in DEG sets for the different tissue types using a Bonferroni corrected P value ≤ 0.05. Finally, we tested for overrepresentation of our input genes in predetermined gene-sets using hypergeometric tests. Gene-sets were obtained from MsigDB, WikiPathways, and GWAS-catalog reported gene-sets. We used Benjamini-Hochberg FDR correction for each of the categories within these gene-set sources separately using a threshold of 0.05 for our adjusted P value.

Roadmap epigenetic state enrichment

Based on the Roadmap 15-core epigenetic state database*(20)*, we used data obtained from 23 blood samples spanning 127 epigenomes to map the QTLs in the summary statistics to their respective epigenetic states. Epigenetic state information was available for bins of 200bp. we aggregated this information into 4 categories; active enhancer states (Enh, EnhG), active

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20229203; this version posted May 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

promotor states (TssA, TssAFlnk), all enhancer states (Enh, EnhG, EnhBiv), and all promotor states (TssA, TssAFlnk, TssBiv). We tested for enrichment using a Fisher's exact test based on the number of unique 200bp bins variants mapped to. This was done after filtering the QTL's down based on their p value using different thresholds (1×10^{-5}, 1×10^{-6} and 1×10^{-7}). Enrichment P values were obtained after FDR correction.

Biomarker measurements

Hospitalized patients with presumed COVID-19 disease were included in a prospective cohort between March and April 2020 at the Radboudumc (Nijmegen, the Netherlands). Disease severity was defined based on the patient's need for intensive care at the time of sampling. The inclusion and clinical characteristics of this patient cohort has been previously described in detail*(56)*. The basic characteristics of the studies samples are shown in Table S15. Plasma samples were collected from EDTA blood and stored at -80°C. The plasma concentrations of CCL25 were determined using the commercially available Inflammation panel from Olink Proteomics AB (Uppsala, Sweden). The procedure of this immunoassay was performed as previously described*(57)*. CCL25 levels are expressed on a log2-scale as normalized protein expression (NPX) values and were normalized using control samples to correct for batch effects. Values under the detection threshold were replaced with the lower limit of detection. CCL25 levels were compared between severe COVID-19 patients (ICU ward = 18) and non-severe COVID-19 patients (non-ICU ward, N = 28) using a linear regression analysis with age and sex as covariates. CCL25 levels was measured every 2 days until an endpoint was reached (either patients left the hospital, or died of the disease).

Plasma concentrations of CXCL16 were measured using a commercially available enzyme-linked immunosorbent assays (ELISA, Invitrogen, Thermo Fisher Scientific) according to the

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20229203; this version posted May 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

manufacturer's protocol, with a lower and upper detection limits of 0.055 and 40 ng/mL, respectively. Plasma concentrations of VWF were measured using a commercially available ELISA (Abcam, ab223864) according to the manufacturer's protocol, with a lower and upper detection limits of 0.94 and 60 µg/mL, respectively. Values under or above the range of detection were replaced with the lower or upper detection limits, respectively. The basic characteristics of the studies samples are shown in Table S16. The student's *t* test was used to compare protein levels between 57 ICU patient and 102 non-ICU patients.

scRNA-seq *analysis*

Samples from patients with COVID-19 were collected during the first wave of the pandemic from Berlin between March and July 2020 in Germany. Berlin cohort consists of 25 mild and 29 severe COVID-19 patients, and 22 controls samples from publicly available scNRA-seq data. The detailed clinical characteristics of those samples have been previously described *(31)*. Gene expression levels were compared between severe patients and mild or healthy controls using FindMarkers functions in *Seurat* v3.2.2 (Stuart, Cell, 2019) with Wilcoxon Rank Sum Test. Genes at least 10% expressed in tested groups and Bonferroni-corrected p-values < 0.05 were regarded as significant differentially expressed genes.

Visualization.

R package ggplot2 was used to perform bar charts, box plots and scatter plots. We applied an online tool Locus zoom to present genes within candidate loci. We used R package pheatmap to generate heat maps.

Acknowledgments:

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20229203; this version posted May 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

We thank all volunteers in the 500FG and 300BCG cohorts of the Human Functional Genomics Project (HFGP) for their participation.

Funding

MGN was supported by an ERC Advanced Grant (833247) and a Spinoza Grant of the Netherlands Organization for Scientific Research. YL was supported by an ERC starting Grant (948207) and a Radboud University Medical Centre Hypatia Grant (2018). CJX was supported by Helmholtz Initiative and Networking Fund (1800167).

Author contributions

M.G.N., C.J.X. and Y.L. designed the study, Y.K. and X.C. performed statistical analyses supervised by Y.L. and C.J.X., M.J., S.M., V.K., V.M., C.d.B., Q.d.M. and L.A.B.J. established the cohorts, helped with the data analysis and interpretation of results. B.Z., V.K., A.d. N, I. G and N.J. performed the validation. C.J.X., M.G.N., Y.L., Y.K. and X.C. wrote the manuscript with input from all authors.

Competing interests

The authors declare no competing interests.

Data and materials availability:

500FG data used in this project have been archived in the BBMRI-NL data infrastructure (https://hfgp.bbmri.nl/). This allows flexible data querying and download, including sufficiently rich metadata and interfaces for machine processing and using FAIR principles to optimize Findability, Accessibility, Interoperability and Reusability. scRNA-seq data generated during this study are deposited at the European Genome-phenome Archive (EGA)

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20229203; this version posted May 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

under access number EGAS00001004571. The processed count tables of the datasets used in the analysis were available at The FASTGenomics platform(fastgrnomics.org).

Ethics statement

The 300BCG (NL58553.091.16) and 500FG (NL42561.091.12) studies were approved by the Arnhem-Nijmegen Medical Ethical Committee. Experiments were conducted according to the principles expressed in the Declaration of Helsinki. Samples of venous blood were drawn after informed consent was obtained.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20229203; this version posted May 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

References

1. D. Ellinghaus, F. Degenhardt, L. Bujanda, M. Buti, A. Albillos, P. Invernizzi, J. Fernández, D. Prati, G. Baselli, R. Asselta, M. M. Grimsrud, C. Milani, F. Aziz, J. Kässens, S. May, M. Wendorff, L. Wienbrandt, F. Uellendahl-Werth, T. Zheng, X. Yi, R. de Pablo, A. G. Chercoles, A. Palom, A.-E. Garcia-Fernandez, F. Rodriguez-Frias, A. Zanella, A. Bandera, A. Protti, A. Aghemo, A. Lleo, A. Biondi, A. Caballero-Garralda, A. Gori, A. Tanck, A. Carreras Nolla, A. Latiano, A. L. Fracanzani, A. Peschuck, A. Julià, A. Pesenti, A. Voza, D. Jiménez, B. Mateos, B. Nafria Jimenez, C. Quereda, C. Paccapelo, C. Gassner, C. Angelini, C. Cea, A. Solier, D. Pestaña, E. Muñiz-Diaz, E. Sandoval, E. M. Paraboschi, E. Navas, F. García Sánchez, F. Ceriotti, F. Martinelli-Boneschi, F. Peyvandi, F. Blasi, L. Téllez, A. Blanco-Grau, G. Hemmrich-Stanisak, G. Grasselli, G. Costantino, G. Cardamone, G. Foti, S. Aneli, H. Kurihara, H. ElAbd, I. My, I. Galván-Femenia, J. Martín, J. Erdmann, J. Ferrusquía-Acosta, K. Garcia-Etxebarria, L. Izquierdo-Sanchez, L. R. Bettini, L. Sumoy, L. Terranova, L. Moreira, L. Santoro, L. Scudeller, F. Mesonero, L. Roade, M. C. Rühlemann, M. Schaefer, M. Carrabba, M. Riveiro-Barciela, M. E. Figuera Basso, M. G. Valsecchi, M. Hernandez-Tejero, M. Acosta-Herrera, M. D'Angiò, M. Baldini, M. Cazzaniga, M. Schulzky, M. Cecconi, M. Wittig, M. Ciccarelli, M. Rodríguez-Gandía, M. Bocciolone, M. Miozzo, N. Montano, N. Braun, N. Sacchi, N. Martínez, O. Özer, O. Palmieri, P. Faverio, P. Preatoni, P. Bonfanti, P. Omodei, P. Tentorio, P. Castro, P. M. Rodrigues, A. Blandino Ortiz, R. de Cid, R. Ferrer, R. Gualtierotti, R. Nieto, S. Goerg, S. Badalamenti, S. Marsal, G. Matullo, S. Pelusi, S. Juzenas, S. Aliberti, V. Monzani, V. Moreno, T. Wesse, T. L. Lenz, T. Pumarola, V. Rimoldi, S. Bosari, W. Albrecht, W. Peter, M. Romero-Gómez, M. D'Amato, S. Duga, J. M. Banales, J. R. Hov, T. Folseraas, L. Valenti, A. Franke, T. H. Karlsen, Genomewide Association Study of Severe Covid-19 with Respiratory Failure, *N Engl J Med* (2020), doi:10.1056/NEJMoa2020283.

2. F. Wu, S. Zhao, B. Yu, Y.-M. Chen, W. Wang, Z.-G. Song, Y. Hu, Z.-W. Tao, J.-H. Tian, Y.-Y. Pei, M.-L. Yuan, Y.-L. Zhang, F.-H. Dai, Y. Liu, Q.-M. Wang, J.-J. Zheng, L. Xu, E. C. Holmes, Y.-Z. Zhang, A new coronavirus associated with human respiratory disease in China, *Nature* **579**, 265–269 (2020).

3. P. Zhou, X.-L. Yang, X.-G. Wang, B. Hu, L. Zhang, W. Zhang, H.-R. Si, Y. Zhu, B. Li, C.-L. Huang, H.-D. Chen, J. Chen, Y. Luo, H. Guo, R.-D. Jiang, M.-Q. Liu, Y. Chen, X.-R. Shen, X. Wang, X.-S. Zheng, K. Zhao, Q.-J. Chen, F. Deng, L.-L. Liu, B. Yan, F.-X. Zhan, Y.-Y. Wang, G.-F. Xiao, Z.-L. Shi, A pneumonia outbreak associated with a new coronavirus of probable bat origin, *Nature* **579**, 270–273 (2020).

4. RECOVERY Collaborative Group, P. Horby, W. S. Lim, J. R. Emberson, M. Mafham, J. L. Bell, L. Linsell, N. Staplin, C. Brightling, A. Ustianowski, E. Elmahi, B. Prudon, C. Green, T. Felton, D. Chadwick, K. Rege, C. Fegan, L. C. Chappell, S. N. Faust, T. Jaki, K. Jeffery, A. Montgomery, K. Rowan, E. Juszczak, J. K. Baillie, R. Haynes, M. J. Landray, Dexamethasone in Hospitalized Patients with Covid-19 - Preliminary Report, *N Engl J Med* (2020), doi:10.1056/NEJMoa2021436.

5. C. Salama, J. Han, L. Yau, W. G. Reiss, B. Kramer, J. D. Neidhart, G. J. Criner, E. Kaplan-Lewis, R. Baden, L. Pandit, M. L. Cameron, J. Garcia-Diaz, V. Chávez, M. Mekebeb-Reuter, F. Lima de Menezes, R. Shah, M. F. González-Lara, B. Assman, J. Freedman, S. V. Mohan,

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20229203; this version posted May 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Tocilizumab in Patients Hospitalized with Covid-19 Pneumonia, *N Engl J Med* **384**, 20–30 (2020).

6. REMAP-CAP Investigators, A. C. Gordon, P. R. Mouncey, F. Al-Beidh, K. M. Rowan, A. D. Nichol, Y. M. Arabi, D. Annane, A. Beane, W. van Bentum-Puijk, L. R. Berry, Z. Bhimani, M. J. M. Bonten, C. A. Bradbury, F. M. Brunkhorst, A. Buzgau, A. C. Cheng, M. A. Detry, E. J. Duffy, L. J. Estcourt, M. Fitzgerald, H. Goossens, R. Haniffa, A. M. Higgins, T. E. Hills, C. M. Horvat, F. Lamontagne, P. R. Lawler, H. L. Leavis, K. M. Linstrum, E. Litton, E. Lorenzi, J. C. Marshall, F. B. Mayr, D. F. McAuley, A. McGlothlin, S. P. McGuinness, B. J. McVerry, S. K. Montgomery, S. C. Morpeth, S. Murthy, K. Orr, R. L. Parke, J. C. Parker, A. E. Patanwala, V. Pettilä, E. Rademaker, M. S. Santos, C. T. Saunders, C. W. Seymour, M. Shankar-Hari, W. I. Sligl, A. F. Turgeon, A. M. Turner, F. L. van de Veerdonk, R. Zarychanski, C. Green, R. J. Lewis, D. C. Angus, C. J. McArthur, S. Berry, S. A. Webb, L. P. G. Derde, Interleukin-6 Receptor Antagonists in Critically Ill Patients with Covid-19, *N Engl J Med* (2021), doi:10.1056/NEJMoa2100433.

7. M. G. Netea, E. J. Giamarellos-Bourboulis, J. Domínguez-Andrés, N. Curtis, R. van Crevel, F. L. van de Veerdonk, M. Bonten, Trained Immunity: a Tool for Reducing Susceptibility to and the Severity of SARS-CoV-2 Infection, *Cell* **181**, 969–977 (2020).

8. M. Z. Tay, C. M. Poh, L. Rénia, P. A. MacAry, L. F. P. Ng, The trinity of COVID-19: immunity, inflammation and intervention, *Nature Reviews Immunology* **20**, 363–374 (2020).

9. R. Verity, L. C. Okell, I. Dorigatti, P. Winskill, C. Whittaker, N. Imai, G. Cuomo-Dannenburg, H. Thompson, P. G. T. Walker, H. Fu, A. Dighe, J. T. Griffin, M. Baguelin, S. Bhatia, A. Boonyasiri, A. Cori, Z. Cucunubá, R. FitzJohn, K. Gaythorpe, W. Green, A. Hamlet, W. Hinsley, D. Laydon, G. Nedjati-Gilani, S. Riley, S. van Elsland, E. Volz, H. Wang, Y. Wang, X. Xi, C. A. Donnelly, A. C. Ghani, N. M. Ferguson, Estimates of the severity of coronavirus disease 2019: a model-based analysis, *The Lancet Infectious Diseases* **20**, 669–677 (2020).

10. X. Zhang, Y. Tan, Y. Ling, G. Lu, F. Liu, Z. Yi, X. Jia, M. Wu, B. Shi, S. Xu, J. Chen, W. Wang, B. Chen, L. Jiang, S. Yu, J. Lu, J. Wang, M. Xu, Z. Yuan, Q. Zhang, X. Zhang, G. Zhao, S. Wang, S. Chen, H. Lu, Viral and host factors related to the clinical outcome of COVID-19, *Nature* **583**, 437–440 (2020).

11. G. Chen, D. Wu, W. Guo, Y. Cao, D. Huang, H. Wang, T. Wang, X. Zhang, H. Chen, H. Yu, X. Zhang, M. Zhang, S. Wu, J. Song, T. Chen, M. Han, S. Li, X. Luo, J. Zhao, Q. Ning, Clinical and immunological features of severe and moderate coronavirus disease 2019, *Journal of Clinical Investigation* **130**, 2620–2629 (2020).

12. A. B. Docherty, E. M. Harrison, C. A. Green, H. E. Hardwick, R. Pius, L. Norman, K. A. Holden, J. M. Read, F. Dondelinger, G. Carson, L. Merson, J. Lee, D. Plotkin, L. Sigfrid, S. Halpin, C. Jackson, C. Gamble, P. W. Horby, J. S. Nguyen-Van-Tam, A. Ho, C. D. Russell, J. Dunning, P. J. Openshaw, J. K. Baillie, M. G. Semple, Features of 20 133 UK patients in hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: prospective observational cohort study, *BMJ* **369**, m1985 (2020).

13. R. ter Horst, M. Jaeger, S. P. Smeekens, M. Oosting, M. A. Swertz, Y. Li, V. Kumar, D. A. Diavatopoulos, A. F. M. Jansen, H. Lemmers, H. Toenhake-Dijkstra, A. E. van Herwaarden, M.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20229203; this version posted May 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Janssen, R. G. van der Molen, I. Joosten, F. C. G. J. Sweep, J. W. Smit, R. T. Netea-Maier, M. M. J. F. Koenders, R. J. Xavier, J. W. M. van der Meer, C. A. Dinarello, N. Pavelka, C. Wijmenga, R. A. Notebaart, L. A. B. Joosten, M. G. Netea, Host and Environmental Factors Influencing Individual Human Cytokine Responses, *Cell* **167**, 1111-1124.e13 (2016).

14. Y. Li, M. Oosting, S. P. Smeekens, M. Jaeger, R. Aguirre-Gamboa, K. T. T. Le, P. Deelen, I. Ricaño-Ponce, T. Schoffelen, A. F. M. Jansen, M. A. Swertz, S. Withoff, E. van de Vosse, M. van Deuren, F. van de Veerdonk, A. Zhernakova, J. W. M. van der Meer, R. J. Xavier, L. Franke, L. A. B. Joosten, C. Wijmenga, V. Kumar, M. G. Netea, A Functional Genomics Approach to Understand Variation in Cytokine Production in Humans, *Cell* **167**, 1099-1110.e14 (2016).

15. K. Watanabe, E. Taskesen, A. van Bochoven, D. Posthuma, Functional mapping and annotation of genetic associations with FUMA, *Nature Communications* **8**, 1826 (2017).

16. X. Xu, X. Chang, H. Pan, H. Su, B. Huang, M. Yang, D. Luo, M. Weng, L. Ma, X. Nie, [Pathological changes of the spleen in ten patients with coronavirus disease 2019(COVID-19) by postmortem needle autopsy], *Zhonghua bing li xue za zhi = Chinese journal of pathology* **49**, 576–582 (2020).

17. F. L. van de Veerdonk, N. A. F. Janssen, I. Grondman, A. H. de Nooijer, V. A. C. M. Koeken, V. Matzaraki, C. K. Boahen, V. Kumar, M. Kox, H. J. P. M. Koenen, R. L. Smeets, I. Joosten, R. J. M. Br&uumlggemann, I. J. E. Kouijzer, H. G. van der Hoeven, J. A. Schouten, T. Frenzel, M. Reijers, W. Hoefsloot, A. S. M. Dofferhoff, A. P. M. Kerckhoffs, M. J. T. Blaauw, K. Veerman, C. Maas, A. H. Schoneveld, I. E. Hoefer, L. P. G. Derde, L. Willems, E. Toonen, M. van Deuren, J. W. M. van der Meer, R. van Crevel, E. J. Giamarellos-Bourboulis, L. A. B. Joosten, M. M. van den Heuvel, J. Hoogerwerf, Q. de Mast, P. Pickkers, M. G. Netea, A systems approach to inflammation identifies therapeutic targets in SARS-CoV-2 infection, *medRxiv* , 2020.05.23.20110916 (2020).

18. C. Robba, D. Battaglini, P. Pelosi, P. R. M. Rocco, Multiple organ dysfunction in SARS-CoV-2: MODS-CoV-2, *Expert Review of Respiratory Medicine* , 1–4 (2020).

19. C. Esche, C. Stellato, L. A. Beck, Chemokines: Key Players in Innate and Adaptive Immunity, *Journal of Investigative Dermatology* **125**, 615–628 (2005).

20. Roadmap Epigenomics Consortium, A. Kundaje, W. Meuleman, J. Ernst, M. Bilenky, A. Yen, A. Heravi-Moussavi, P. Kheradpour, Z. Zhang, J. Wang, M. J. Ziller, V. Amin, J. W. Whitaker, M. D. Schultz, L. D. Ward, A. Sarkar, G. Quon, R. S. Sandstrom, M. L. Eaton, Y.-C. Wu, A. R. Pfenning, X. Wang, M. Claussnitzer, Y. Liu, C. Coarfa, R. A. Harris, N. Shoresh, C. B. Epstein, E. Gjoneska, D. Leung, W. Xie, R. D. Hawkins, R. Lister, C. Hong, P. Gascard, A. J. Mungall, R. Moore, E. Chuah, A. Tam, T. K. Canfield, R. S. Hansen, R. Kaul, P. J. Sabo, M. S. Bansal, A. Carles, J. R. Dixon, K.-H. Farh, S. Feizi, R. Karlic, A.-R. Kim, A. Kulkarni, D. Li, R. Lowdon, G. Elliott, T. R. Mercer, S. J. Neph, V. Onuchic, P. Polak, N. Rajagopal, P. Ray, R. C. Sallari, K. T. Siebenthall, N. A. Sinnott-Armstrong, M. Stevens, R. E. Thurman, J. Wu, B. Zhang, X. Zhou, A. E. Beaudet, L. A. Boyer, P. L. De Jager, P. J. Farnham, S. J. Fisher, D. Haussler, S. J. M. Jones, W. Li, M. A. Marra, M. T. McManus, S. Sunyaev, J. A. Thomson, T. D. Tlsty, L.-H. Tsai, W. Wang, R. A. Waterland, M. Q. Zhang, L. H. Chadwick, B. E. Bernstein, J. F.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20229203; this version posted May 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Costello, J. R. Ecker, M. Hirst, A. Meissner, A. Milosavljevic, B. Ren, J. A. Stamatoyannopoulos, T. Wang, M. Kellis, Integrative analysis of 111 reference human epigenomes, *Nature* **518**, 317–330 (2015).

21. E. J. Giamarellos-Bourboulis, M. G. Netea, N. Rovina, K. Akinosoglou, A. Antoniadou, N. Antonakos, G. Damoraki, T. Gkavogianni, M.-E. Adami, P. Katsaounou, M. Ntaganou, M. Kyriakopoulou, G. Dimopoulos, I. Koutsodimitropoulos, D. Velissaris, P. Koufargyris, A. Karageorgos, K. Katrini, V. Lekakis, M. Lupse, A. Kotsaki, G. Renieris, D. Theodoulou, V. Panou, E. Koukaki, N. Koulouris, C. Gogos, A. Koutsoukou, Complex Immune Dysregulation in COVID-19 Patients with Severe Respiratory Failure, *Cell Host & Microbe* **27**, 992-1000.e3 (2020).

22. P. Mehta, D. F. McAuley, M. Brown, E. Sanchez, R. S. Tattersall, J. J. Manson, COVID-19: consider cytokine storm syndromes and immunosuppression, *The Lancet* **395**, 1033–1034 (2020).

23. J. T. Sun, Z. Chen, P. Nie, H. Ge, L. Shen, F. Yang, X. L. Qu, X. Y. Ying, Y. Zhou, W. Wang, M. Zhang, J. Pu, Lipid Profile Features and Their Associations With Disease Severity and Mortality in Patients With COVID-19, *Front. Cardiovasc. Med.* **7** (2020), doi:10.3389/fcvm.2020.584987.

24. M. Franchini, F. Capra, G. Targher, M. Montagnana, G. Lippi, Relationship between ABO blood group and von Willebrand factor levels: from biology to clinical implications, *Thrombosis Journal* **5**, 14 (2007).

25. G. Goshua, A. B. Pine, M. L. Meizlish, C.-H. Chang, H. Zhang, P. Bahel, A. Baluha, N. Bar, R. D. Bona, A. J. Burns, C. S. Dela Cruz, A. Dumont, S. Halene, J. Hwa, J. Koff, H. Menninger, N. Neparidze, C. Price, J. M. Siner, C. Tormey, H. M. Rinder, H. J. Chun, A. I. Lee, Endotheliopathy in COVID-19-associated coagulopathy: evidence from a single-centre, cross-sectional study, *The Lancet Haematology* **7**, e575–e582 (2020).

26. Aksenova A.Y., Von Willebrand factor and endothelial damage: a possible association with COVID-19, *EcoGen* **18**, 135–138 (2020).

27. A. V. Rapkiewicz, X. Mai, S. E. Carsons, S. Pittaluga, D. E. Kleiner, J. S. Berger, S. Thomas, N. M. Adler, D. M. Charytan, B. Gasmi, J. S. Hochman, H. R. Reynolds, Megakaryocytes and platelet-fibrin thrombi characterize multi-organ thrombosis at autopsy in COVID-19: A case series, *EClinicalMedicine* **24** (2020), doi:10.1016/j.eclinm.2020.100434.

28. C. Kawecki, P. J. Lenting, C. V. Denis, von Willebrand factor and inflammation, *Journal of Thrombosis and Haemostasis* **15**, 1285–1294 (2017).

29. A. V. Khera, M. Chaffin, K. G. Aragam, M. E. Haas, C. Roselli, S. H. Choi, P. Natarajan, E. S. Lander, S. A. Lubitz, P. T. Ellinor, S. Kathiresan, Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations, *Nature Genetics* **50**, 1219–1224 (2018).

30. E. Pairo-Castineira, S. Clohisey, L. Klaric, A. D. Bretherick, K. Rawlik, D. Pasko, S. Walker, N. Parkinson, M. H. Fourman, C. D. Russell, J. Furniss, A. Richmond, E. Gountouna, N.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20229203; this version posted May 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Wrobel, D. Harrison, B. Wang, Y. Wu, A. Meynert, F. Griffiths, W. Oosthuyzen, A. Kousathanas, L. Moutsianas, Z. Yang, R. Zhai, C. Zheng, G. Grimes, R. Beale, J. Millar, B. Shih, S. Keating, M. Zechner, C. Haley, D. J. Porteous, C. Hayward, J. Yang, J. Knight, C. Summers, M. Shankar-Hari, P. Klenerman, L. Turtle, A. Ho, S. C. Moore, C. Hinds, P. Horby, A. Nichol, D. Maslove, L. Ling, D. McAuley, H. Montgomery, T. Walsh, A. C. Pereira, A. Renieri, J. Millar, A. Nichol, T. Walsh, M. Shankar-Hari, C. Ponting, J. Meikle, P. Finernan, E. Mcmaster, A. Law, J. K. Baillie, T. Paterson, T. Wackett, R. Armstrong, R. Clark, A. Coutts, L. Donnelly, T. Gilchrist, K. Hafezi, L. Macgillivray, A. Maclean, S. McCafferty, K. Morrice, J. Weaver, C. Boz, A. Golightly, M. Ward, H. Mal, H. Szoor-McElhinney, A. Brown, R. Hendry, A. Stenhouse, L. Cullum, D. Law, S. Law, R. Law, M. Swets, N. Day, F. Taneski, E. Duncan, N. Parkinson, D. Collier, S. Wood, A. Zak, C. Borra, M. Matharu, P. May, Z. Alldis, O. Mitchelmore, R. Bowles, A. Easthope, F. Bibi, I. Lancoma-Malcolm, J. Gurasashvili, J. Pheby, J. Shiel, M. Bolton, M. Patel, M. Taylor, O. Zongo, P. Ebano, P. Harding, R. Astin-Chamberlain, Y. Choudhury, A. Cox, D. Kallon, M. Burton, R. Hall, S. Blowes, Z. Prime, J. Biddle, O. Prysyazhna, T. Newman, C. Tierney, J. Kassam, M. Shankar-Hari, M. Ostermann, S. Campos, A. Bociek, R. Lim, N. Grau, T. O. Jones, C. Whitton, M. Marotti, G. Arbane, S. Bonner, K. Hugill, J. Reid, I. Welters, V. Waugh, K. Williams, D. Shaw, J. F. Roman, M. L. Martinez, E. Johnson, A. Waite, B. Johnston, D. Hamilton, S. Mulla, M. McPhail, J. Smith, J. K. Baillie, L. Barclay, D. Hope, C. McCulloch, L. McQuillan, S. Clark, J. Singleton, K. Priestley, N. Rea, M. Callaghan, R. Campbell, G. Andrew, L. Marshall, S. McKechnie, P. Hutton, A. Bashyal, N. Davidson, C. Summers, P. Polgarova, K. Stroud, N. Pathan, K. Elston, S. Agrawal, C. Battle, L. Newey, T. Rees, R. Harford, E. Brinkworth, M. Williams, C. Murphy, I. White, M. Croft, N. Bandla, M. Gellamucho, J. Tomlinson, H. Turner, M. Davies, A. Quinn, I. Hussain, C. Thompson, H. Parker, R. Bradley, R. Griffiths, J. Scriven, A. Nilsson, M. Bates, J. Dasgin, J. Gill, A. Puxty, S. Cathcart, D. Salutous, L. Turner, K. Duffy, K. Puxty, A. Joseph, R. Herdman-Grant, R. Simms, A. Swain, A. Naranjo, R. Crowe, K. Sollesta, A. Loveridge, D. Baptista, E. Morino, M. Davey, D. Golden, J. Jones, J. Moreno Cuesta, A. Haldeos, D. Bakthavatsalam, R. Vincent, M. Elhassan, K. Xavier, A. Ganesan, D. Purohit, M. Abdelrazik, J. Morgan, L. Akeroyd, S. Bano, T. Lawton, D. Warren, M. Bromley, K. Sellick, L. Gurr, B. Wilkinson, V. Nagarajan, P. Szedlak, J. Cupitt, E. Stoddard, L. Benham, S. Preston, S. Laha, N. Slawson, Z. Bradshaw, J. Brown, M. Caswell, S. Melling, P. Bamford, M. Faulkner, K. Cawley, H. Jeffrey, E. London, H. Sainsbury, I. Nagra, F. Nasir, C. Dunmore, R. Jones, A. Abraheem, M. Al-Moasseb, R. Girach, G. Padden, J. Egan, C. Brantwood, P. Alexander, J. Bradley-Potts, S. Allen, T. Felton, S. Manna, S. Farnell-Ward, S. Leaver, J. Queiroz, E. Maccacari, The GenOMICC Investigators, GenOMICC Consortium, GenOMICC co-investigators, Central management and laboratory team, Data analysis team, L. Barts Health NHS Trust UK, L. Guy's & St Thomas' Hospital UK, M. James Cook University Hospital UK, L. The Royal Liverpool University Hospital UK, L. King's College Hospital UK, E. Royal Infirmary of Edinburgh UK, O. John Radcliffe Hospital UK, C. Addenbrooke's Hospital UK, S. Morriston Hospital UK, S. Ashford & St Peter's Hospital UK, S. Royal Stoke University Hospital UK, B. Queen Elizabeth Hospital UK, G. Glasgow Royal Infirmary UK, S. Kingston Hospital UK, K. The Tunbridge Wells Hospital & Maidstone Hospital UK, L. North Middlesex University Hospital NHS Trust UK, B. Bradford Royal Infirmary UK, B. Blackpool Victoria Hospital UK, C. Countess of Chester Hospital UK, M. Wythenshawe Hospital UK, L. St George's Hospital UK, Genetic mechanisms of critical illness in COVID-19, *Nature* **591**, 92–98 (2021).

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20229203; this version posted May 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

31. J. Schulte-Schrepping, N. Reusch, D. Paclik, K. Baßler, S. Schlickeiser, B. Zhang, B. Krämer, T. Krammer, S. Brumhard, L. Bonaguro, E. De Domenico, D. Wendisch, M. Grasshoff, T. S. Kapellos, M. Beckstette, T. Pecht, A. Saglam, O. Dietrich, H. E. Mei, A. R. Schulz, C. Conrad, D. Kunkel, E. Vafadarnejad, C.-J. Xu, A. Horne, M. Herbert, A. Drews, C. Thibeault, M. Pfeiffer, S. Hippenstiel, A. Hocke, H. Müller-Redetzky, K.-M. Heim, F. Machleidt, A. Uhrig, L. Bosquillon de Jarcy, L. Jürgens, M. Stegemann, C. R. Glösenkamp, H.-D. Volk, C. Goffinet, M. Landthaler, E. Wyler, P. Georg, M. Schneider, C. Dang-Heine, N. Neuwinger, K. Kappert, R. Tauber, V. Corman, J. Raabe, K. M. Kaiser, M. T. Vinh, G. Rieke, C. Meisel, T. Ulas, M. Becker, R. Geffers, M. Witzenrath, C. Drosten, N. Suttorp, C. von Kalle, F. Kurth, K. Händler, J. L. Schultze, A. C. Aschenbrenner, Y. Li, J. Nattermann, B. Sawitzki, A.-E. Saliba, L. E. Sander, A. Angelov, R. Bals, A. Bartholomäus, A. Becker, D. Bezdan, E. Bonifacio, P. Bork, T. Clavel, M. Colome-Tatche, A. Diefenbach, A. Dilthey, N. Fischer, K. Förstner, J.-S. Frick, J. Gagneur, A. Goesmann, T. Hain, M. Hummel, S. Janssen, J. Kalinowski, R. Kallies, B. Kehr, A. Keller, S. Kim-Hellmuth, C. Klein, O. Kohlbacher, J. O. Korbel, I. Kurth, M. Landthaler, Y. Li, K. Ludwig, O. Makarewicz, M. Marz, A. McHardy, C. Mertes, M. Nöthen, P. Nürnberg, U. Ohler, S. Ossowski, J. Overmann, S. Peter, K. Pfeffer, A. R. Poetsch, A. Pühler, N. Rajewsky, M. Ralser, O. Rieß, S. Ripke, U. Nunes da Rocha, P. Rosenstiel, A.-E. Saliba, L. E. Sander, B. Sawitzki, P. Schiffer, E.-C. Schulte, J. L. Schultze, A. Sczyrba, O. Stegle, J. Stoye, F. Theis, J. Vehreschild, J. Vogel, M. von Kleist, A. Walker, J. Walter, D. Wieczorek, J. Ziebuhr, Severe COVID-19 Is Marked by a Dysregulated Myeloid Cell Compartment, *Cell* **182**, 1419-1440.e23 (2020).

32. S. Zhou, G. Butler-Laporte, T. Nakanishi, D. R. Morrison, J. Afilalo, M. Afilalo, L. Laurent, M. Pietzner, N. Kerrison, K. Zhao, E. Brunet-Ratnasingham, D. Henry, N. Kimchi, Z. Afrasiabi, N. Rezk, M. Bouab, L. Petitjean, C. Guzman, X. Xue, C. Tselios, B. Vulesevic, O. Adeleye, T. Abdullah, N. Almamlouk, Y. Chen, M. Chassé, M. Durand, C. Paterson, J. Normark, R. Frithiof, M. Lipcsey, M. Hultström, C. M. T. Greenwood, H. Zeberg, C. Langenberg, E. Thysell, M. Pollak, V. Mooser, V. Forgetta, D. E. Kaufmann, J. B. Richards, A Neanderthal OAS1 isoform protects individuals of European ancestry against COVID-19 susceptibility and severity, *Nature Medicine* (2021), doi:10.1038/s41591-021-01281-1.

33. W. Tian, N. Zhang, R. Jin, Y. Feng, S. Wang, S. Gao, R. Gao, G. Wu, D. Tian, W. Tan, Y. Chen, G. F. Gao, C. C. L. Wong, Immune suppression in the early stage of COVID-19 disease, *Nature Communications* **11**, 5859 (2020).

34. Y. Dai, J. Wang, H.-H. Jeong, W. Chen, P. Jia, Z. Zhao, Association of CXCR6 with COVID-19 severity: Delineating the host genetic factors in transcriptomic regulation, *bioRxiv* , 2021.02.17.431554 (2021).

35. H. Zeberg, S. Pääbo, The major genetic risk factor for severe COVID-19 is inherited from Neanderthals, *Nature* (2020), doi:10.1038/s41586-020-2818-3.

36. L. Wei, W. Wang, D. Chen, B. Xu, Dysregulation of the immune response affects the outcome of critical COVID-19 patients, *Journal of Medical Virology* **92**, 2768–2776 (2020).

37. R. L. Hoiland, N. A. Fergusson, A. R. Mitra, D. E. G. Griesdale, D. V. Devine, S. Stukas, J. Cooper, S. Thiara, D. Foster, L. Y. C. Chen, A. Y. Y. Lee, E. M. Conway, C. L. Wellington, M. S. Sekhon, The association of ABO blood group with indices of disease severity and multiorgan dysfunction in COVID-19, *Blood Advances* **4**, 4981–4989 (2020).

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20229203; this version posted May 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

38. M. Hultström, B. Persson, O. Eriksson, M. Lipcsey, R. Frithiof, B. Nilsson, Blood type A associates with critical COVID-19 and death in a Swedish cohort, *Critical Care* **24**, 496 (2020).

39. M. B. Barnkob, A. Pottegård, H. Støvring, T. M. Haunstrup, K. Homburg, R. Larsen, M. B. Hansen, K. Titlestad, B. Aagaard, B. K. Møller, T. Barington, Reduced prevalence of SARS-CoV-2 infection in ABO blood group O, *Blood Advances* **4**, 4990–4993 (2020).

40. Y. Wu, Z. Feng, P. Li, Q. Yu, Relationship between ABO blood group distribution and clinical characteristics in patients with COVID-19, *Clinica Chimica Acta* **509**, 220–223 (2020).

41. S. Dzik, K. Eliason, E. B. Morris, R. M. Kaufman, C. M. North, COVID-19 and ABO blood groups, *Transfusion* **60**, 1883–1884 (2020).

42. J. M. O'Sullivan, S. Ward, H. Fogarty, J. S. O'Donnell, More on 'Association between ABO blood groups and risk of SARS-CoV-2 pneumonia,' *British Journal of Haematology* **190**, 27–28 (2020).

43. I. Huang, R. Pranata, Lymphopenia in severe coronavirus disease-2019 (COVID-19): systematic review and meta-analysis, *Journal of Intensive Care* **8**, 36 (2020).

44. C. I. van der Made, A. Simons, J. Schuurs-Hoeijmakers, G. van den Heuvel, T. Mantere, S. Kersten, R. C. van Deuren, M. Steehouwer, S. V. van Reijmersdal, M. Jaeger, T. Hofste, G. Astuti, J. Corominas Galbany, V. van der Schoot, H. van der Hoeven, W. Hagmolen of ten Have, E. Klijn, C. van den Meer, J. Fiddelaers, Q. de Mast, C. P. Bleeker-Rovers, L. A. B. Joosten, H. G. Yntema, C. Gilissen, M. Nelen, J. W. M. van der Meer, H. G. Brunner, M. G. Netea, F. L. van de Veerdonk, A. Hoischen, Presence of Genetic Variants Among Young Men With Severe COVID-19, *JAMA* **324**, 663–673 (2020).

45. Y. Li, M. Oosting, P. Deelen, I. Ricaño-Ponce, S. Smeekens, M. Jaeger, V. Matzaraki, M. A. Swertz, R. J. Xavier, L. Franke, C. Wijmenga, L. A. B. Joosten, V. Kumar, M. G. Netea, Inter-individual variability and genetic influences on cytokine responses to bacteria and fungi, *Nature Medicine* **22**, 952–960 (2016).

46. R. Aguirre-Gamboa, I. Joosten, P. C. M. Urbano, R. G. van der Molen, E. van Rijssen, B. van Cranenbroek, M. Oosting, S. Smeekens, M. Jaeger, M. Zorro, S. Withoff, A. E. van Herwaarden, F. C. G. J. Sweep, R. T. Netea, M. A. Swertz, L. Franke, R. J. Xavier, L. A. B. Joosten, M. G. Netea, C. Wijmenga, V. Kumar, Y. Li, H. J. P. M. Koenen, Differential Effects of Environmental and Genetic Factors on T and B Cell Immune Traits, *Cell Reports* **17**, 2474–2487 (2016).

47. M. Schirmer, S. P. Smeekens, H. Vlamakis, M. Jaeger, M. Oosting, E. A. Franzosa, R. ter Horst, T. Jansen, L. Jacobs, M. J. Bonder, A. Kurilshikov, J. Fu, L. A. B. Joosten, A. Zhernakova, C. Huttenhower, C. Wijmenga, M. G. Netea, R. J. Xavier, Linking the Human Gut Microbiome to Inflammatory Cytokine Production Capacity, *Cell* **167**, 1125-1136.e8 (2016).

48. V. A. C. M. Koeken, L. C. J. de Bree, V. P. Mourits, S. J. C. F. M. Moorlag, J. Walk, B. Cirovic, R. J. W. Arts, M. Jaeger, H. Dijkstra, H. Lemmers, L. A. B. Joosten, C. S. Benn, R. van Crevel, M. Netea, BCG vaccination in humans inhibits systemic inflammation in a sex-dependent manner, *J Clin Invest* (2020), doi:10.1172/JCI133935.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20229203; this version posted May 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

49. V. P. Mourits, V. A. C. M. Koeken, L. C. J. de Bree, S. J. C. F. M. Moorlag, W. C. Chu, X. Xu, H. Dijkstra, H. Lemmers, L. A. B. Joosten, Y. Wang, R. van Crevel, M. G. Netea, A. González, Ed. BCG-Induced Trained Immunity in Healthy Individuals: The Effect of Plasma Muramyl Dipeptide Concentrations, *Journal of Immunology Research* **2020**, 5812743 (2020).

50. T. S. Shah, J. Z. Liu, J. A. B. Floyd, J. A. Morris, N. Wirth, J. C. Barrett, C. A. Anderson, optiCall: a robust genotype-calling algorithm for rare, low-frequency and common variants, *Bioinformatics* **28**, 1598–1603 (2012).

51. P.-R. Loh, P. Danecek, P. F. Palamara, C. Fuchsberger, Y. A Reshef, H. K Finucane, S. Schoenherr, L. Forer, S. McCarthy, G. R. Abecasis, R. Durbin, A. L Price, Reference-based phasing using the Haplotype Reference Consortium panel, *Nature Genetics* **48**, 1443–1448 (2016).

52. X. Chu, M. Jaeger, O. B. Bakker, R. Aguirre-Gamboa, M. Oosting, S. P. Smeekens, S. Withoff, R. T. Netea-Maier, H. J. P. M. Koenen, I. Joosten, R. J. Xavier, L. Franke, L. a. B. Joosten, S. Sanna, V. Kumar, C. Wijmenga, M. G. Netea, Y. Li, in *European Journal of Human Genetics*, (Nature Publishing Group, 2019), vol. 27, pp. 787–787.

53. A. A. Shabalin, Matrix eQTL: ultra fast eQTL analysis via large matrix operations, *Bioinformatics.* **28** (2012), doi:10.1093/bioinformatics/bts163.

54. R. N. Tunjungputri, Y. Li, P. G. de Groot, C. A. Dinarello, S. P. Smeekens, M. Jaeger, M. Doppenberg-Oosting, M. Cruijsen, H. Lemmers, H. Toenhake-Dijkstra, R. Aguirre-Gamboa, V. Kumar, T. N. Wijmenga, L. A. B. Joosten, M. G. Netea, A. van der Ven, Q. de Mast, The inter-relationship of platelets with interleukin-1 β-mediated inflammation in humans, *Thrombosis and Haemostasis* **118**, 2112–2125 (2018).

55. C. Giambartolomei, D. Vukcevic, E. E. Schadt, L. Franke, A. D. Hingorani, C. Wallace, V. Plagnol, Bayesian Test for Colocalisation between Pairs of Genetic Association Studies Using Summary Statistics, *PLOS Genetics* **10**, e1004383 (2014).

56. N. A. F. Janssen, I. Grondman, A. H. de Nooijer, C. K. Boahen, V. A. C. M. Koeken, V. Matzaraki, V. Kumar, X. He, M. Kox, H. J. P. M. Koenen, R. L. Smeets, I. Joosten, R. J. M. Brüggemann, I. J. E. Kouijzer, H. G. van der Hoeven, J. A. Schouten, T. Frenzel, M. Reijers, W. Hoefsloot, A. S. M. Dofferhoff, M. J. van Apeldoorn, M. J. T. Blaauw, K. Veerman, C. Maas, A. H. Schoneveld, I. E. Hoefer, L. P. G. Derde, M. van Deuren, J. W. M. van der Meer, R. van Crevel, E. J. Giamarellos-Bourboulis, L. A. B. Joosten, M. M. van den Heuvel, J. Hoogerwerf, Q. de Mast, P. Pickkers, M. G. Netea, F. L. van de Veerdonk, Dysregulated innate and adaptive immune responses discriminate disease severity in COVID-19, *The Journal of Infectious Diseases* (2021), doi:10.1093/infdis/jiab065.

57. E. Assarsson, M. Lundberg, G. Holmquist, J. Björkesten, S. B. Thorsen, D. Ekman, A. Eriksson, E. R. Dickens, S. Ohlsson, G. Edfeldt, A.-C. Andersson, P. Lindstedt, J. Stenvang, M. Gullberg, S. Fredriksson, Homogenous 96-Plex PEA Immunoassay Exhibiting High Sensitivity, Specificity, and Excellent Scalability, *PLOS ONE* **9**, e95192 (2014).

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20229203; this version posted May 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Figure Captions

Fig. 1. Study overview. Firstly, we performed a functional mapping and annotation (FUMA) to link COVID-19 SNPs to gene expression and identified important pathways and tissues contributing to the pathophysiology of COVID-19. Secondly, we utilized the cytokine quantitative trait loci (QTL), metabolite QTL and plate QTL from Human Functional Genomics Projects (HFGP) 500FG data (n=451) to test if specific loci are associated with immune functions. Thirdly, we linked PRS score with gender and BMI in 500FG (n=451) and 300BCG (n=313) cohorts. Lastly, we validated our findings in disease cohorts in single-cell transcriptomics data from Berlin (n=76), and proteomics data from Nijmegen (n=46, n=159). SNPs, single-nucleotide polymorphisms.

Fig.2. Functional annotation of COVID-19 loci using the FUMA pipeline and association 3p21.31 loci with immune traits. A) MAGMA Tissue expression results on 30 general tissues type (GTEx v8). FUMA analysis was done based on genes identified after using their genomic location, eQTL associations, and histone activity. B) The heatmap showing the assoication between 3p21.31 loci with cytokine production upon *in vitro* stimulations. Red color in heatmap indicates higher cytokine production leaded by risk allele in COVID-19 GWAS profiles, Blue color indicates lower cytokine production leaded by risk allele in COVID-19 GWAS profiles. C) a boxplot showing COVID-19 risk allele(rs6441930-C) associated with reduced IL6 production with influenza stimulation of PBMC for 24 hours (p-value = 0.026).

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20229203; this version posted May 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Fig 3. Functional annotation of ABO loci. A) locus zoom plot showing the significant association between ABO loci and VWF level. B) a boxplot showing COVID-19 risk allele(rs687621-G) associated with increasing VWF level (p-value = 9.58×10^{-20}). C) a barplot showing consistent negative correlations between VWF levels and T cell-derived cytokines D) scatter plots showing colocalization between ABO loci with VWF, lymphocytes, monocytes, neutrophils and white blood cell counts.

Fig. 4. Correlation of COVID-19 PRS with gender and BMI. A) bar plot representing the ratio of low risk versus high risk in 500FG. The X-axis shows the range of different quantiles (e.g.,10% corresponds to those individuals with PRS between 0th and 10th percentile of the population), and the Y-axis shows the odds ratio when comparing low PRS risk and high PRS risk in the male and female group from different quantiles. B) bar plot representing the ratio of low risk versus high risk in 300BCG. C) Bar plot representing the ratio of low versus high PRS based risk between men and women in 500FG calculated without including the sex chromosomes. D). Bar plot representing the ratio of low versus high PRS based risk between men and women in 300BCG calculated without including the sex chromosomes. E). Scatter plot showing the correlation between PRS with BMI in 500FG F) in Scatter plot showing the correlation between PRS with BMI in 300BCG.

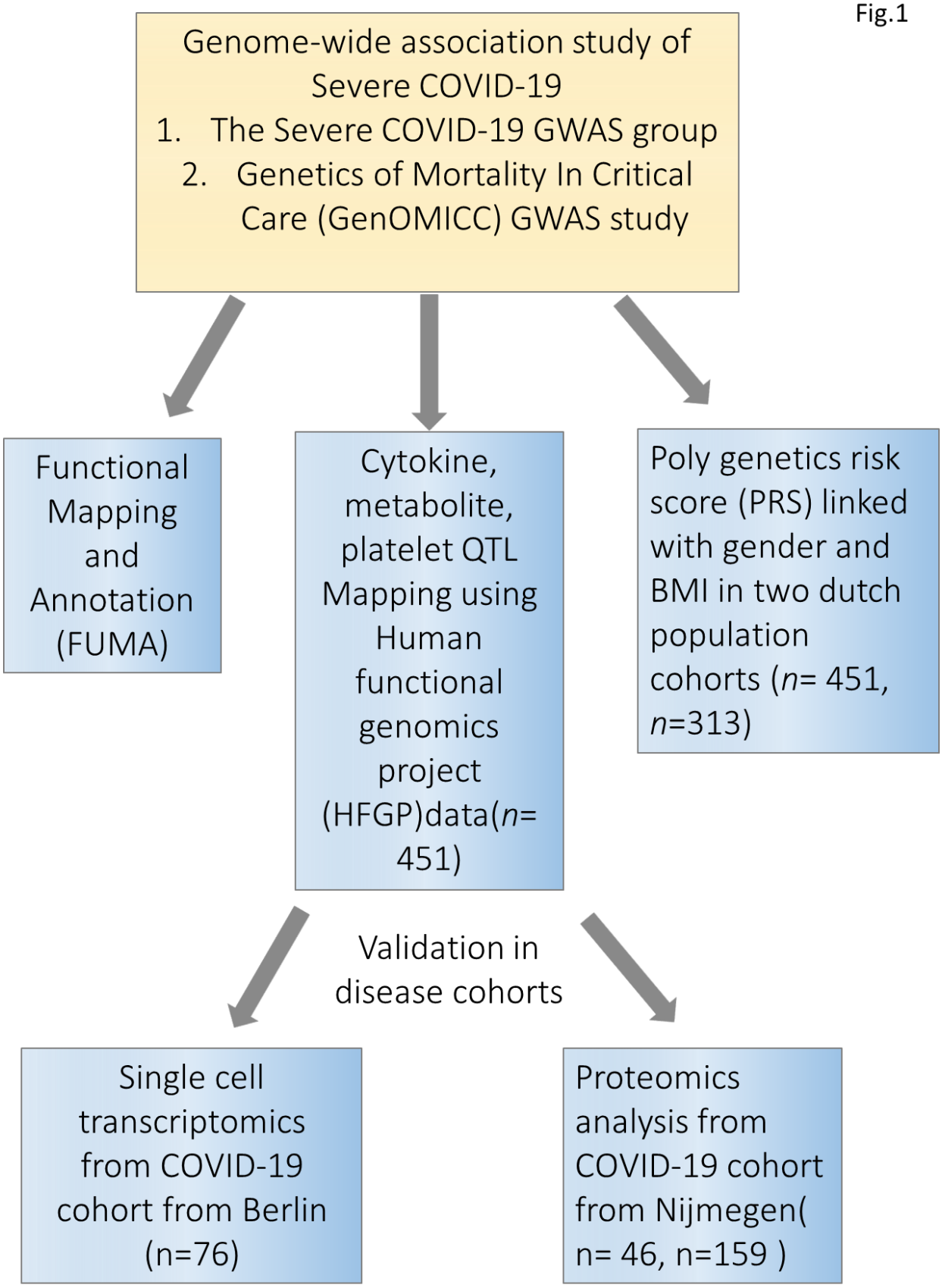
Fig. 5 Replication and validation. A) Heatmap showing the association between 3p21.31 loci and immune traits can be replicated in an independent cohort. The red color corresponds to higher cytokine production leaded by risk allele in COVID-19 GWAS profiles, whereas blue color indicates lower cytokine production leaded by risk allele in COVID-19 GWAS profiles. B) Dot plots of expression of GWAS genes in single-cell transcriptomics of COVID-19 patients. The GWAS genes were selected from the Severe COVID-19 GWAS*(1)* and

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20229203; this version posted May 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

GenOMICC study*(30)*. C) Violin plots of the expression of monocyte-derived cytokine genes and T cell derived cytokine genes in COVID-19 Berlin cohort based on single cell RNA-seq data. D) A boxplot of the differential protein levels of CCL25 between 18 ICU and 28 non-ICU COVID-19 patients from another independent cohort. E) A boxplot of the differential protein levels of CCL25 between 57 ICU and 102 non-ICU COVID-19 patients. F) A boxplot of the differential expression of VWF between 57 ICU and 102 non-ICU COVID-19 patients.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20229203; this version posted May 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

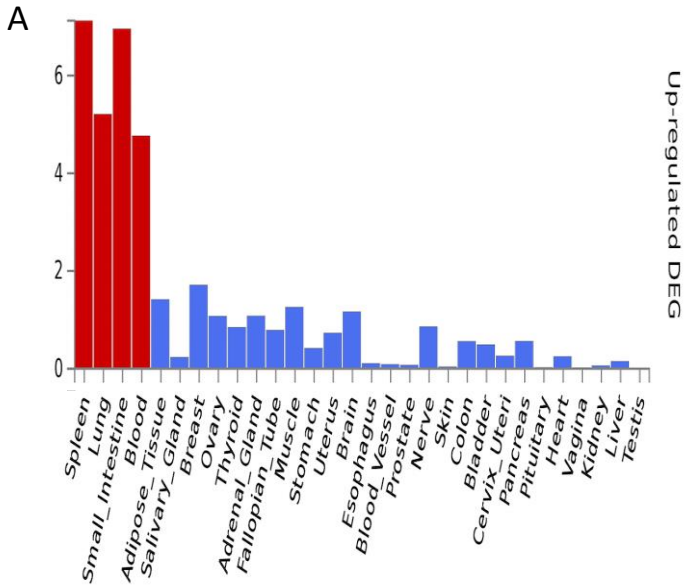
Fig.1



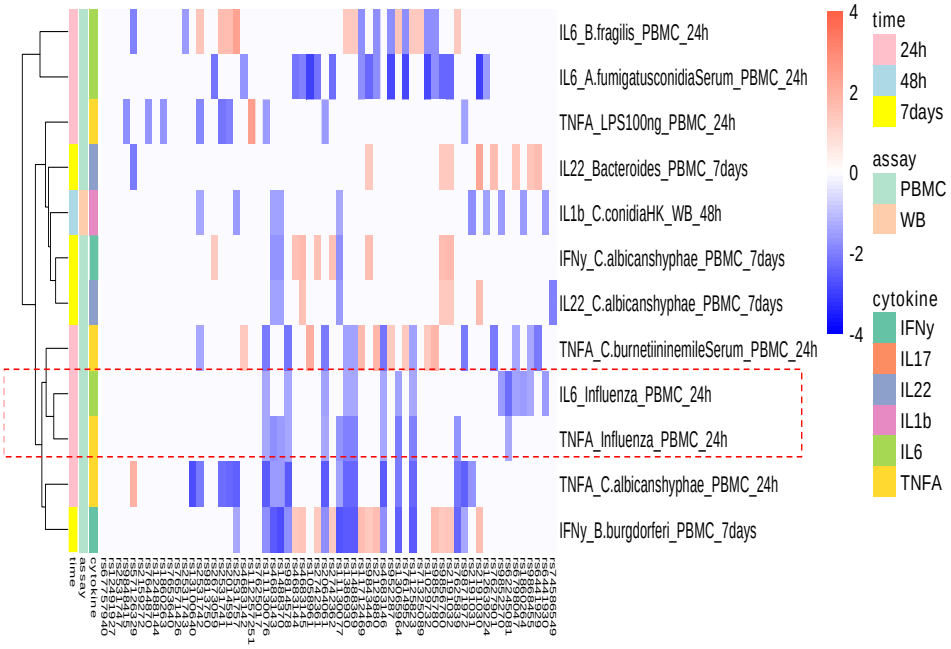
medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20229203; this version posted May 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Fig.2

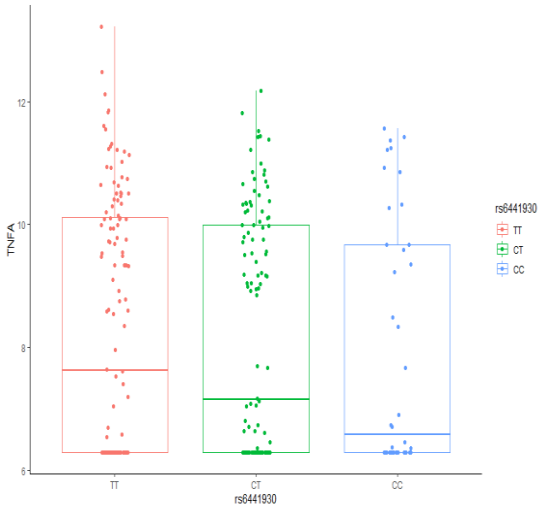
A



B



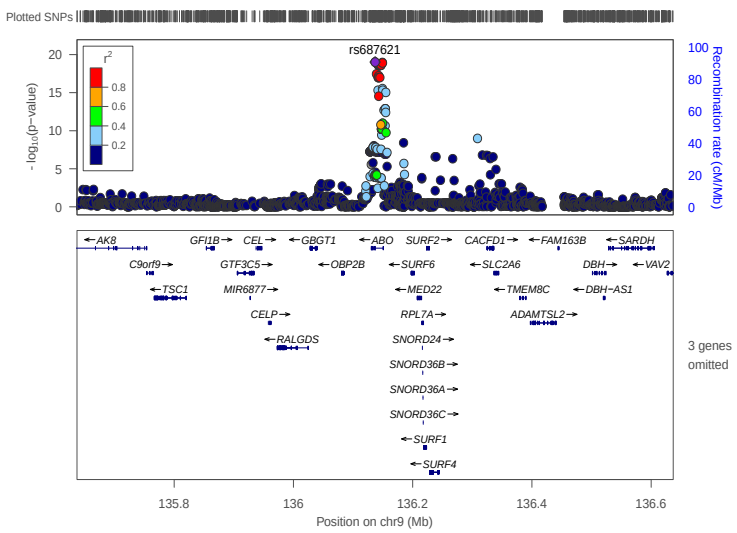
C



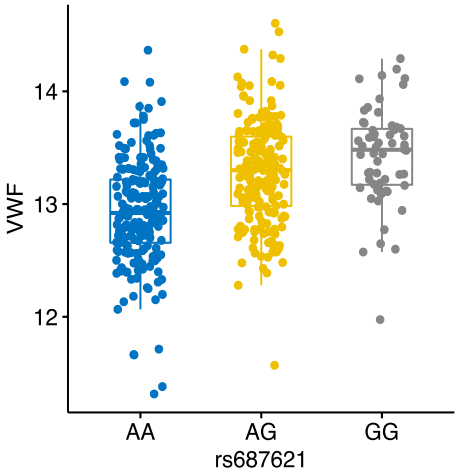
medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20229203; this version posted May 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Fig.3

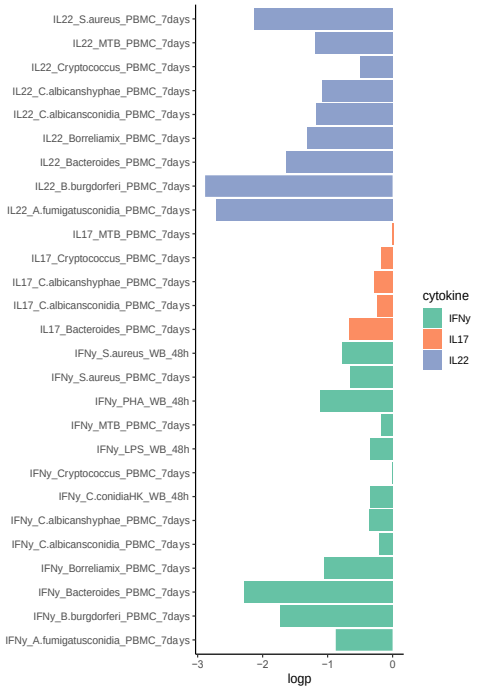
A



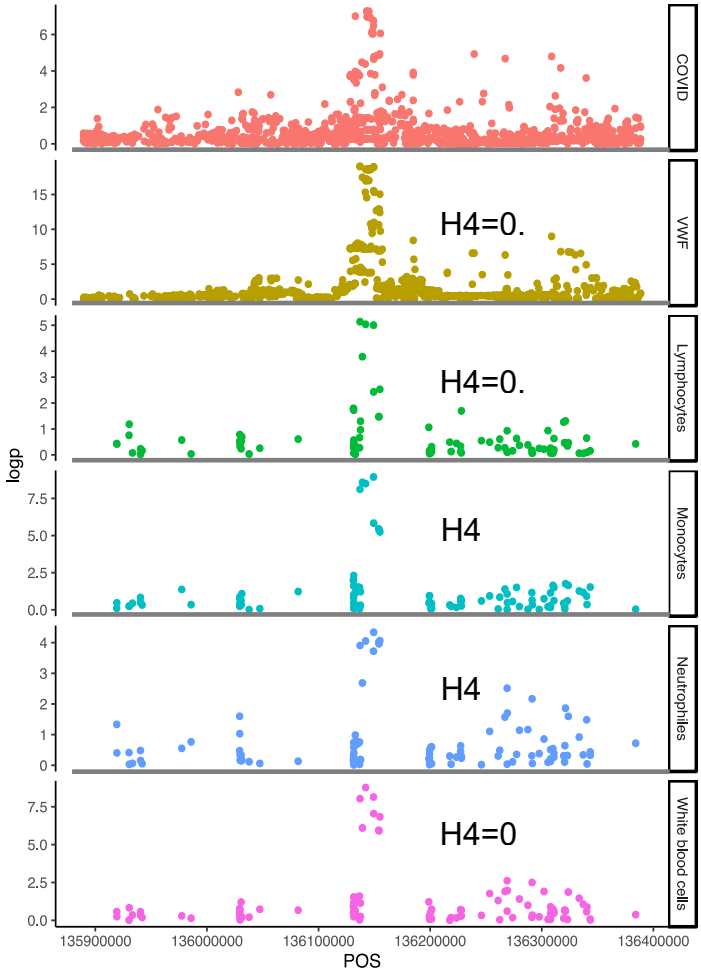
B



C



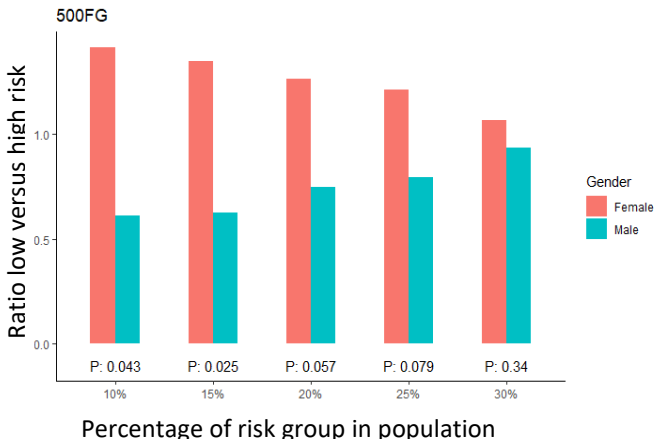
D



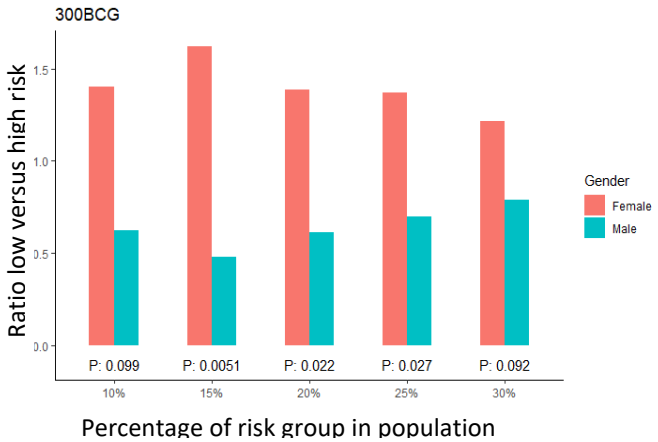
medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20229203; this version posted May 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Fig.4

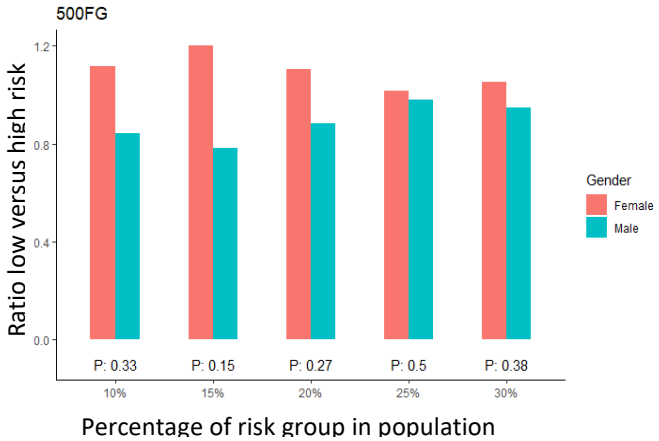
A



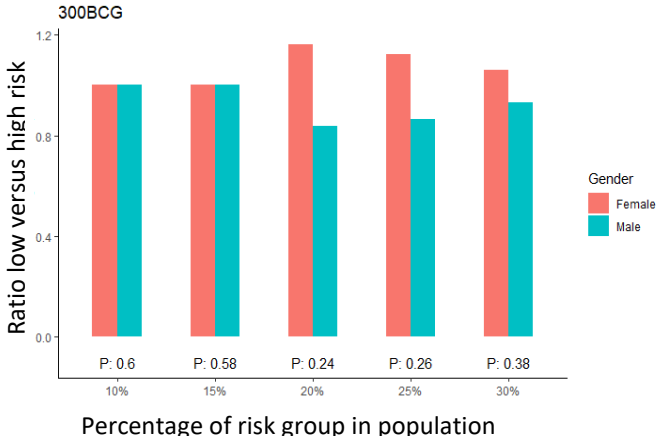
B



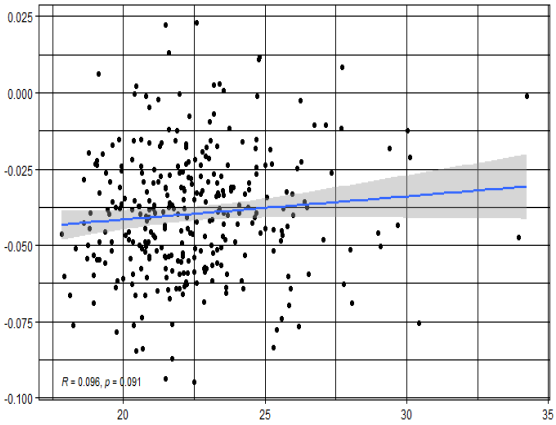
C



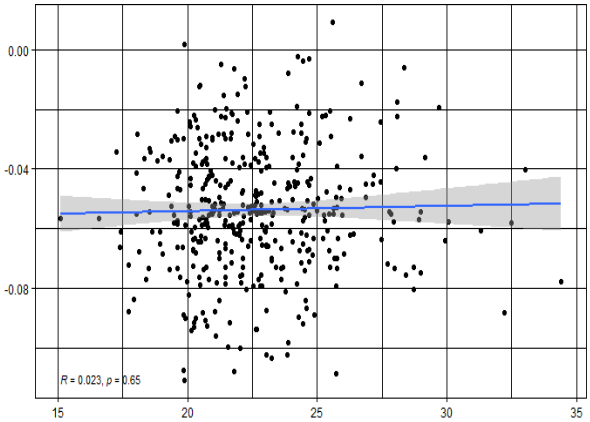
D



E



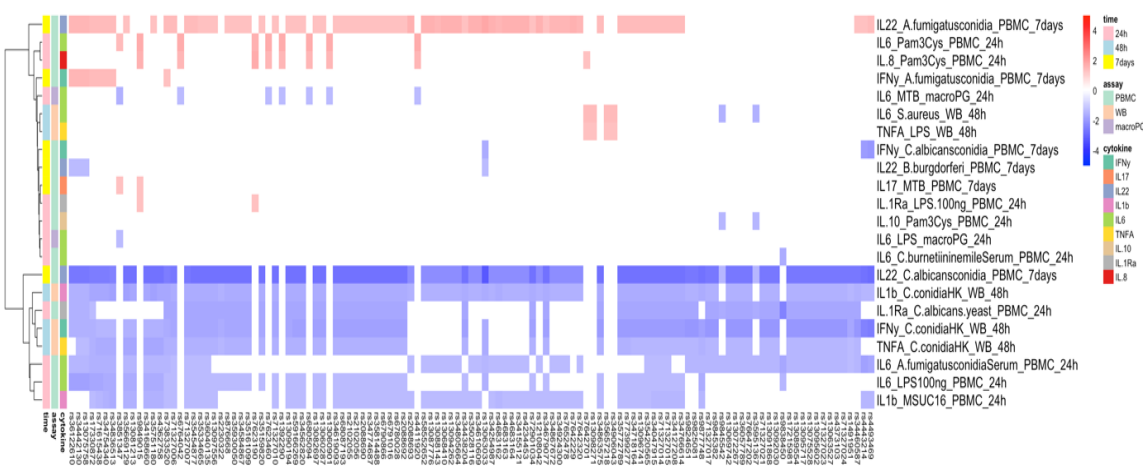
F



medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20229203; this version posted May 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

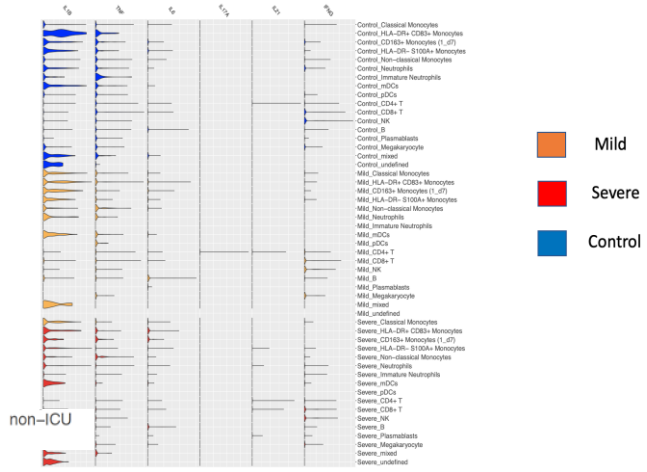
Fig.5

A



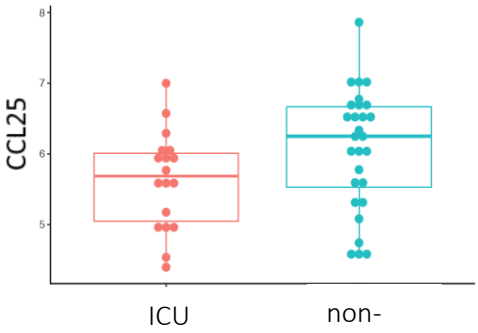
B

C



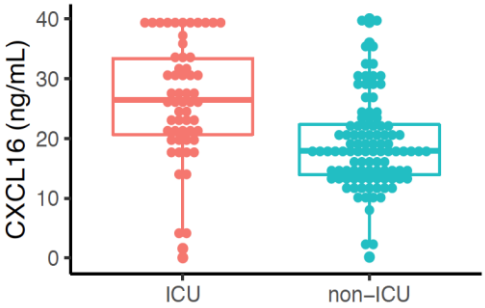
D

CCL25
P: 0.0255



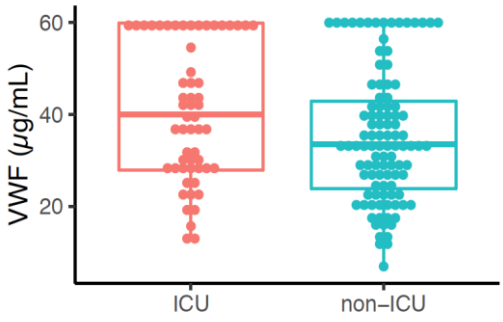
E

CXCL16 (incl sample >UDL)



F

VWF (incl sample >UDL)
P<0.0259



medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20229203; this version posted May 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

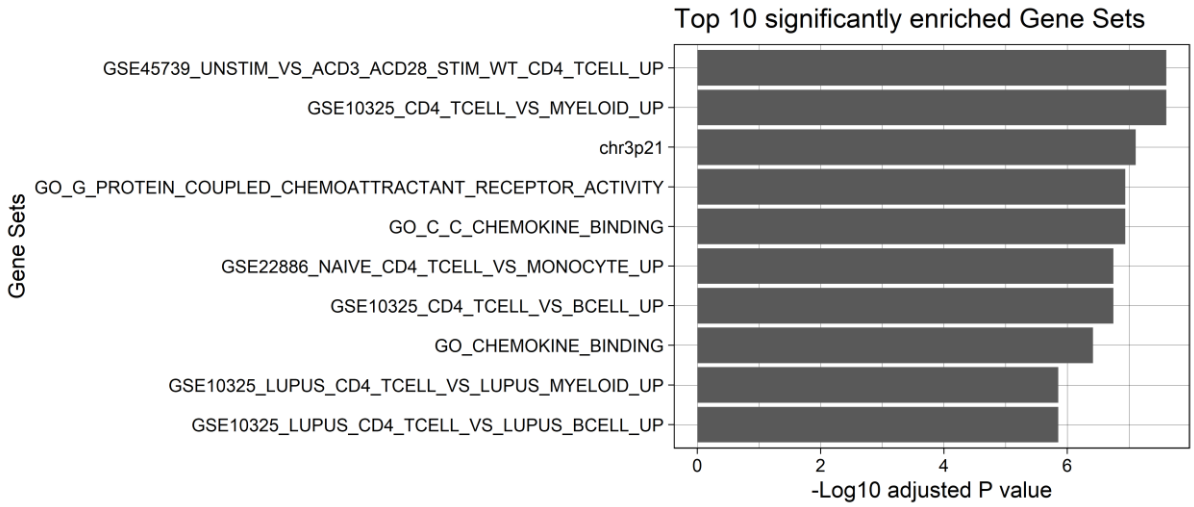


Fig. S1 A bar plot showing the top 10 significant enriched gene sets using functional annotation of COVID-19 loci by FUMA pipeline.

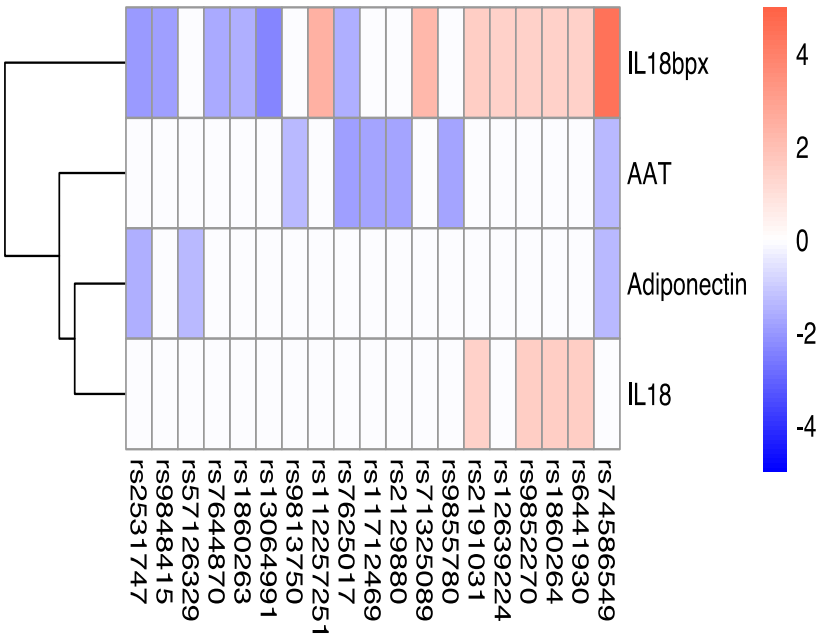


Fig. S2 Heatmap of the genetic associations between 3p21.31 loci and circulating mediator.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20229203; this version posted May 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

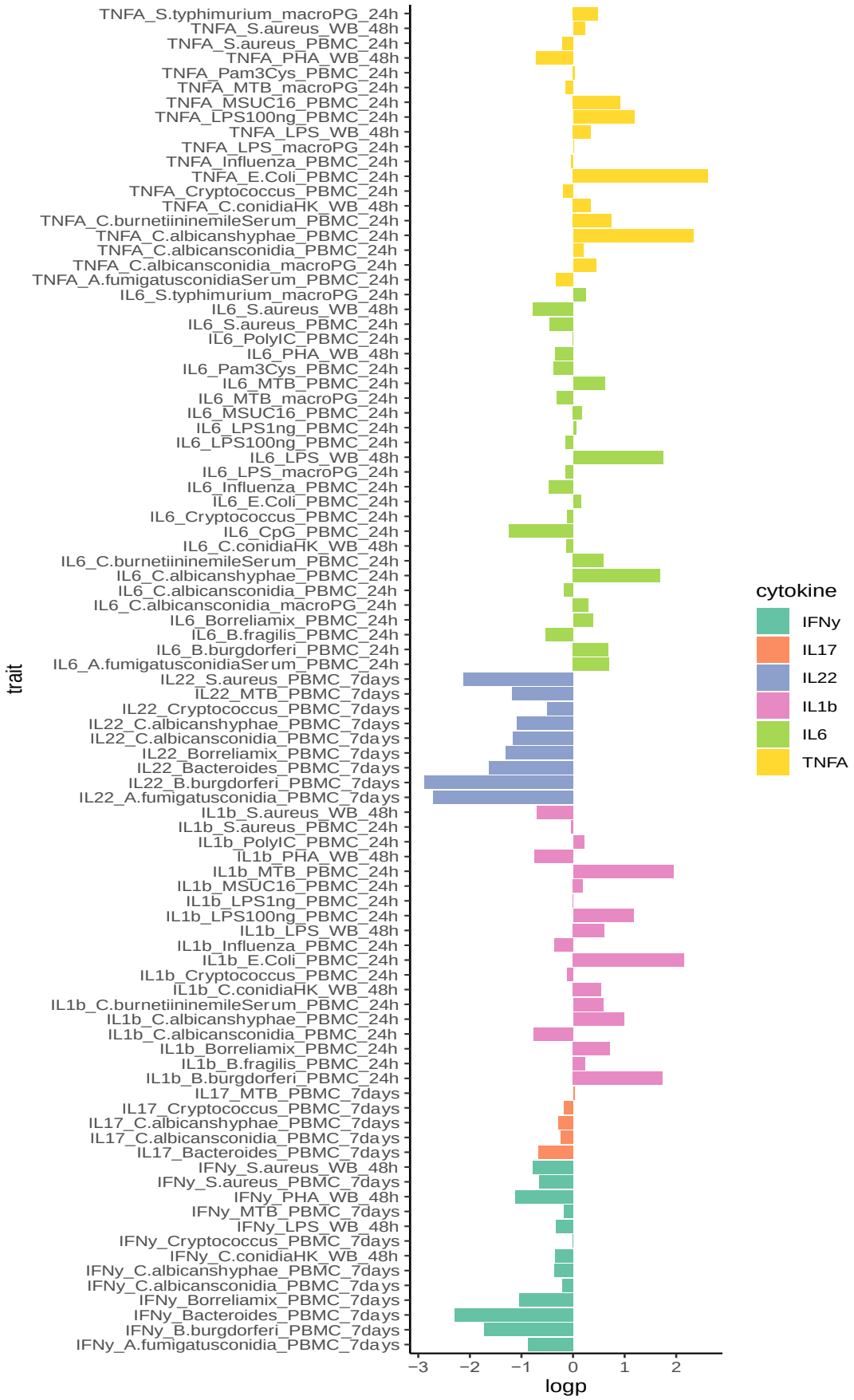


Fig. S3 A barplot showing associations between VWF levels and cytokine

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20229203; this version posted May 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

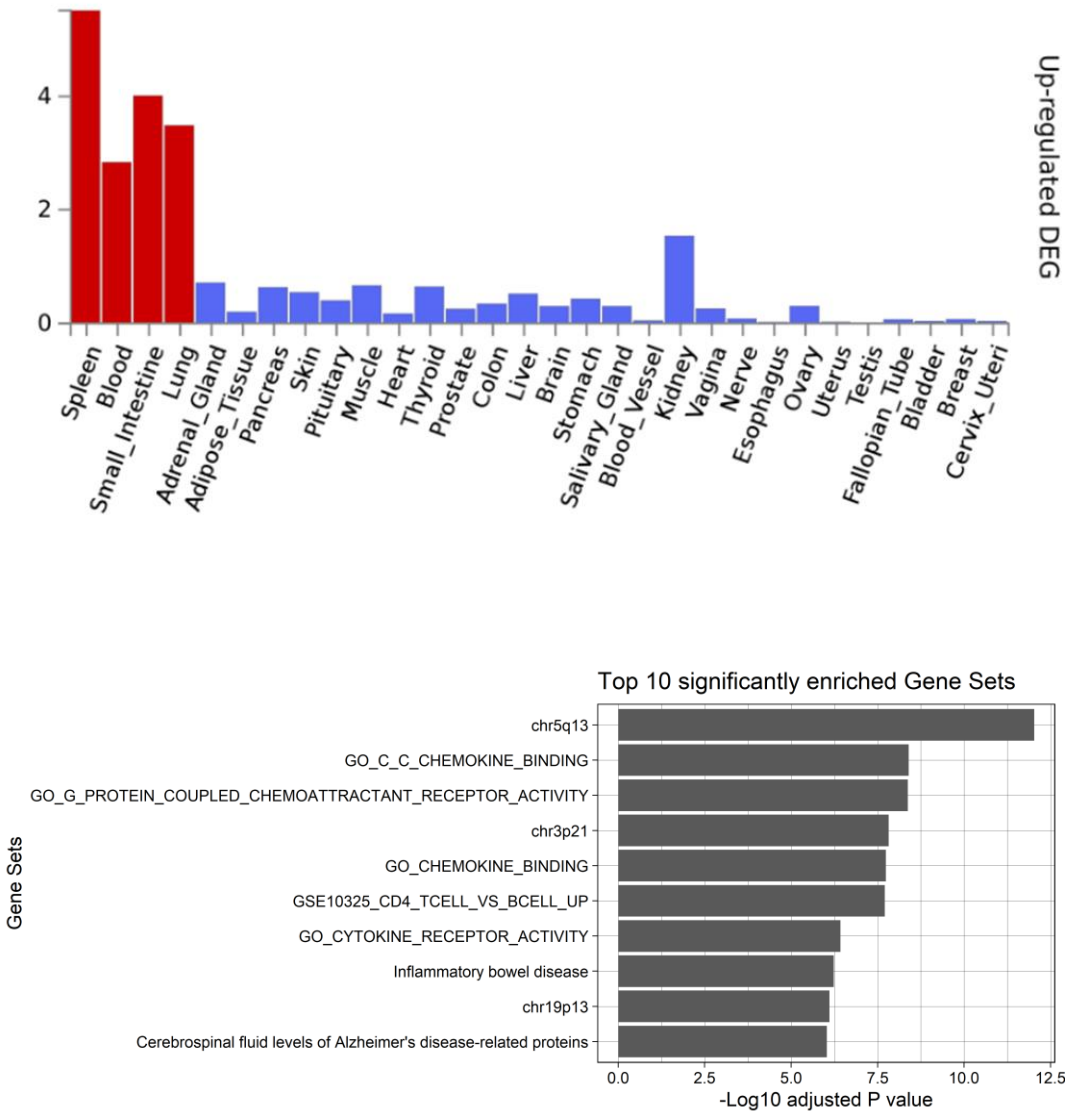


Fig.S4**.** Functional annotation of COVID-19 loci from GenOMICC study of European ancestry group using the FUMA pipeline. This was done based on genes identified after using FUMA to map QTLs based on their genomic location, eQTL associations, and histone activity. A) MAGMA Tissue expression results on 30 general tissues type (GTEx v8), B) The top 10 significant enriched gene sets.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20229203; this version posted May 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

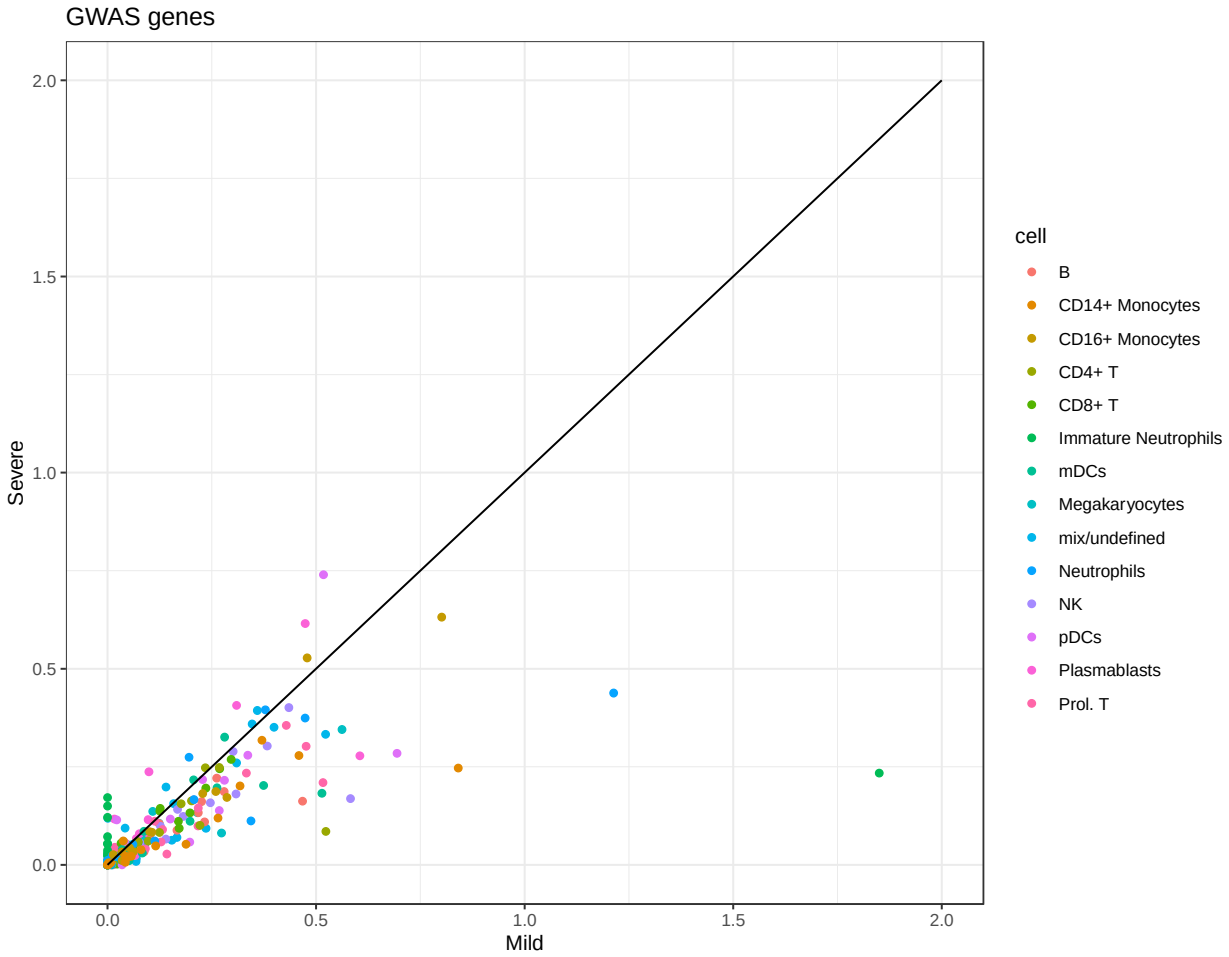


Fig.S5**.** A scatter plot showing the average cell type-specific expression of GWAS genes between mild and severe COVID-19 patients.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20229203; this version posted May 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

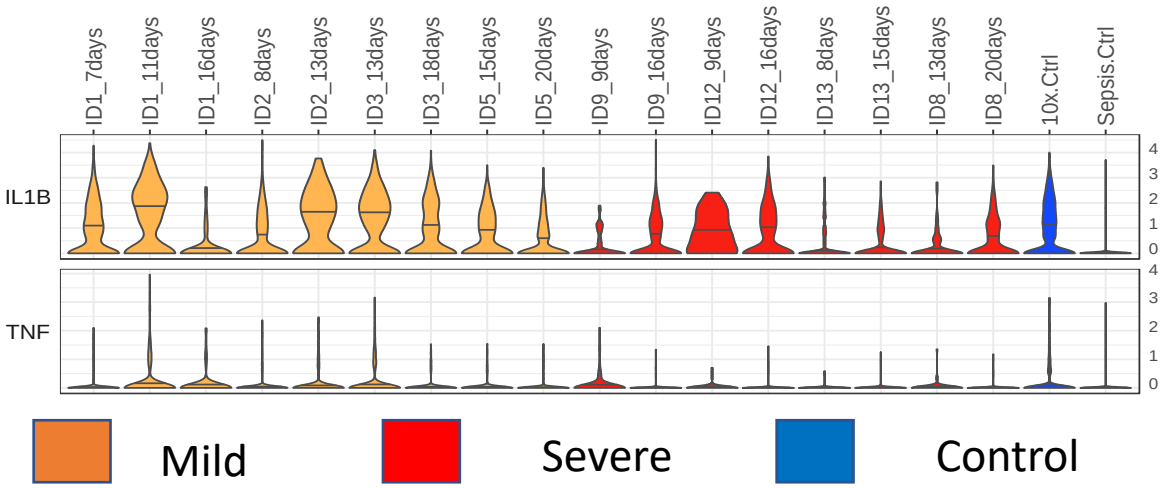


Fig. S6. The longitudinal change of *IL-1β* and *TNF-α* expression (violin-plots) in monocytes of COVID-19 Berlin cohort.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20229203; this version posted May 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

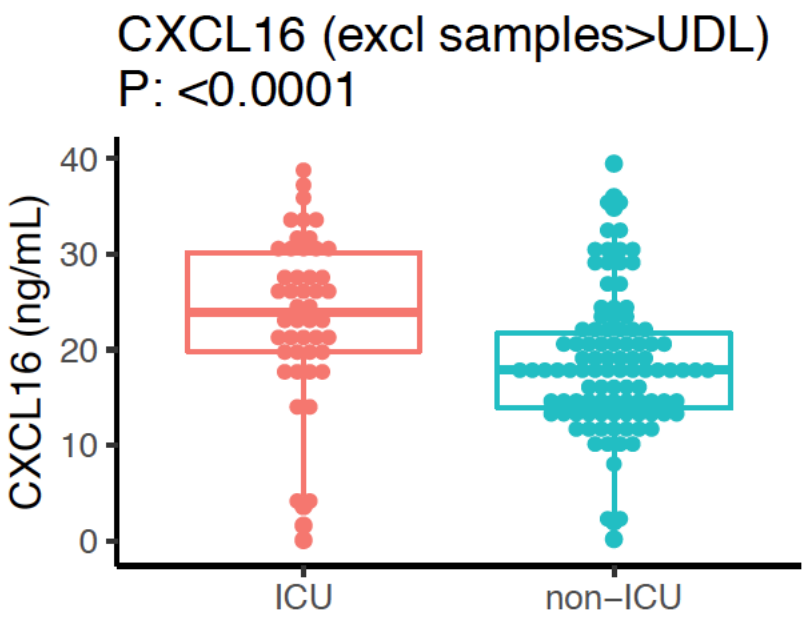


Fig. S7) A boxplot of the differential expression of CXCL16 between ICU and non-ICU COVID-19 patients after excluding samples which are above upper detection limit (UDL).

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20229203; this version posted May 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

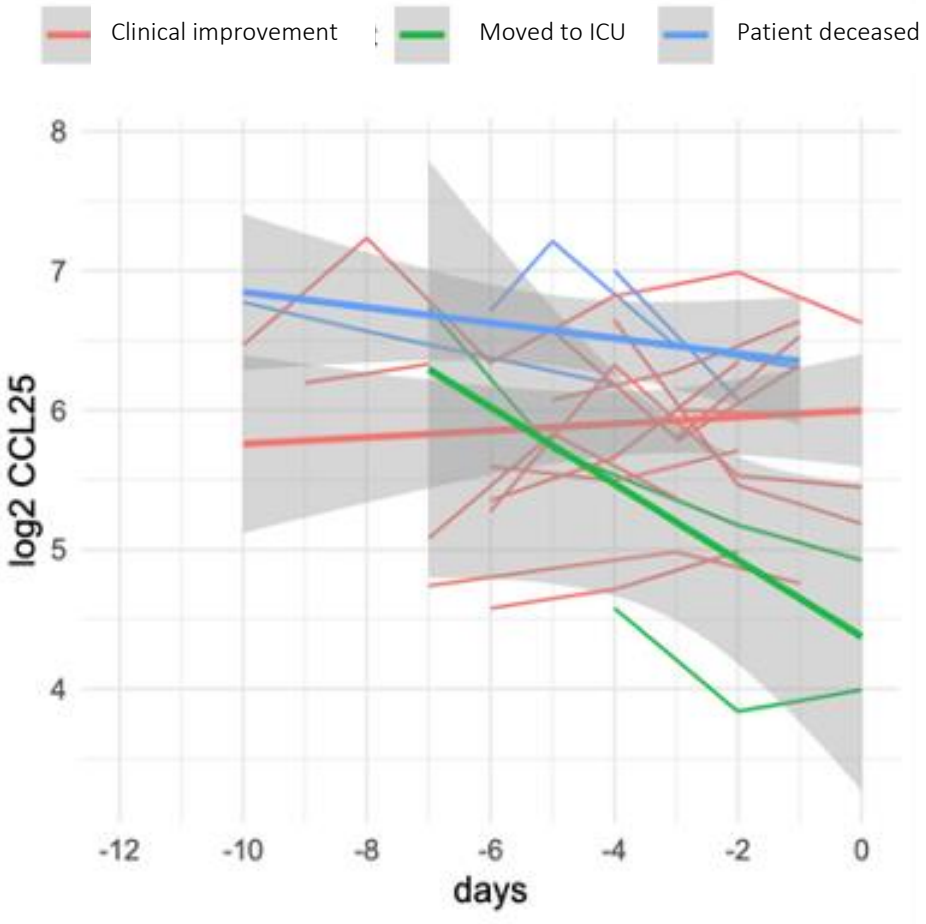


Fig. S8 The expression of CCL25 at different time points in three clinical groups (clinical improvement, moved to ICU and patient deceased).

medRxiv preprint doi: https://doi.org/10.1101/2020.11.10.20229203; this version posted May 14, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

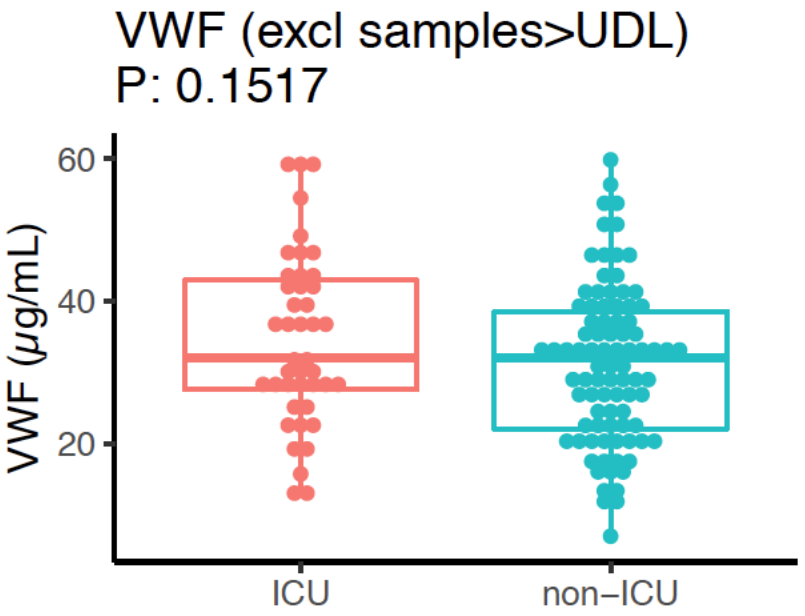


Fig. S9 A boxplot of the differential expression of VWF between ICU and non-ICU COVID-19 patients after excluding samples which are above upper detection limit (UDL).