medRxiv preprint doi: https://doi.org/10.1101/2020.06.17.20134015; this version posted June 19, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Analysis of Genetic Host Response Risk Factors in Severe COVID-19 Patients

Krystyna Taylor*, Sayoni Das*, Matthew Pearson*, James Kozubek, Marcin Pawlowski, Claus Erik Jensen, Zbigniew Skowron, Gert Lykke Møller, Mark Strivens*, Steve Gardner*†

PrecisionLife Ltd, Long Hanborough, Oxford, UK

* These authors contributed equally.
† Corresponding author: steve@precisionlife.com

ABSTRACT

BACKGROUND

Epidemiological studies indicate that as many as 20% of individuals who test positive for COVID-19 develop severe symptoms that can require hospitalization. These symptoms include low platelet count, severe hypoxia, increased inflammatory cytokines and reduced glomerular filtration rate. Additionally, severe COVID-19 is associated with several chronic co-morbidities, including cardiovascular disease, hypertension and type 2 diabetes mellitus.

The identification of genetic risk factors that impact differential host responses to SARS-CoV-2, resulting in the development of severe COVID-19, is important in gaining greater understanding into the biological mechanisms underpinning life-threatening responses to the virus. These insights could be used in the identification of high-risk individuals and for the development of treatment strategies for these patients.

METHODS

As of June 6, 2020, there were 976 patients who tested positive for COVID-19 and were hospitalized, indicating they had a severe response to SARS-CoV-2. There were however too few patients with a mild form of COVID-19 to use this cohort as our control population. Instead we used similar control criteria to our previous study looking at shared genetic risk factors between severe COVID-19 and sepsis, selecting controls who had not developed sepsis despite having maximum co-morbidity risk and exposure to sepsis-causing pathogens.

RESULTS

Using a combinatorial (high-order epistasis) analysis approach, we identified 68 protein-coding genes that were highly associated with severe COVID-19. At the time of analysis, nine of these genes have been linked to differential response to SARS-CoV-2. We also found many novel targets that are involved in key biological pathways associated with the development of severe COVID-19, including production of pro-inflammatory cytokines, endothelial cell dysfunction, lipid droplets, neurodegeneration and viral susceptibility factors.

CONCLUSION

The variants we found in genes relating to immune response pathways and cytokine production cascades, were in equal proportions across all severe COVID-19 patients, regardless of their co-morbidities. This suggests that such variants are not associated with any specific co-morbidity, but are common amongst patients who develop severe COVID-19.

Among the 68 severe COVID-19 risk-associated genes, we found several druggable protein targets and pathways. Nine are targeted by drugs that have reached at least Phase I clinical trials, and a further eight have active chemical starting points for novel drug development.

Several of these targets were particularly enriched in specific co-morbidities, providing insights into shared pathological mechanisms underlying both the development of severe COVID-19, ARDS and these predisposing co-morbidities. We can use these insights to identify patients who are at greatest risk of contracting severe COVID-19 and develop targeted therapeutic strategies for them, with the aim of improving disease burden and survival rates.

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/10.1101/2020.06.17.20134015; this version posted June 19, 2020. 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. It is made available under a CC-BY-NC 4.0 International license .

Introduction

Coronavirus disease 2019 (COVID-19), caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), is a major threat to public health. As of 16th June 2020, there are estimated to be over 8 million confirmed cases globally, resulting in approximately 437,000 deaths worldwide[1,2]. Although many who develop the disease present with only mild symptoms, reports from multiple international health systems have shown that up to 20% of individuals testing positive for COVID-19 go on to develop severe forms of the disease that may require hospitalization.

Significant associations of disease severity risk have been observed with epidemiological factors including age, sex, blood group and ethnicity in addition to co-morbidity with many common conditions such as cardiovascular disease, diabetes, hypertension and chronic pulmonary diseases including asthma[3]. It would be clinically useful to be able to identify the features that result in differential host responses, particularly those that predispose some patients to developing severe COVID-19. In the context of management of at-risk populations, prior to an effective and widely available vaccine, such insights could have great utility in developing new detection, protection and treatment strategies targeted specifically at these high-risk individuals.

In our previous paper, we explored the shared risk factors between sepsis (a major clinical feature in hospitalized COVID-19 patients) and severe COVID-19 disease[4]. We observed that 59% of hospitalized COVID-19 cases also have sepsis[5], and that the two diseases share similar co-morbidity risks[6]. In that study we identified 70 risk-associated genes in a sepsis population and found significant overlap in genetic risk variants between sepsis patients and those hospitalized with severe COVID-19.

As more data from COVID-19 patients has become available in the UK Biobank[7,8], we are now able to investigate the host response genetic risk factors directly, using genotype datasets from 779, 877 and 929 patients hospitalized with severe COVID-19 and comparing them against healthy controls. In this study, we have sought to identify the risk variants associated directly with severe COVID-19 patients, to gain insight of the underlying pathology and disease mechanisms in relation to this patient group.

Methods

COVID-19 test records were downloaded from the UK Biobank (data releases 18 May, 26 May and 6 June, 2020)[7,8]. The 18 May 2020 dataset included 5,657 test records relating to 3,002 individuals in UK Biobank. Of these, 1,073 patients had at least one positive COVID-19 test record, including 818 patients who were hospitalized and 255 who tested positive but were not hospitalized. In accordance with guidance from UK Biobank, we classified those 818 patients who were hospitalized after testing positive with COVID-19 as having a severe form of the disease and the 255 who were not hospitalized as having the mild form of the disease.

The analysis was conducted as a case-control study. After quality control and removal of samples with missing data, we assigned as cases 779 patients (442 males, 337 females) who had been hospitalized with severe COVID-19. Of these severe cases, 62% had one or more of the most common co-morbidities associated with high severe COVID-19 risk (Figure 1).

The most prevalent co-morbidity was hypertension (50%), followed by chronic respiratory disease (22%), diabetes (20%) and cardiovascular disease (18%). 248 (32%) patients were reported having two or more co-morbidities. 291 severe COVID-19 cases had none of these co-morbidities. In comparison, these co-morbidities were found to be less prevalent in the mild cases and an age and gender-matched random selection of the UK Biobank population (Figure 1).

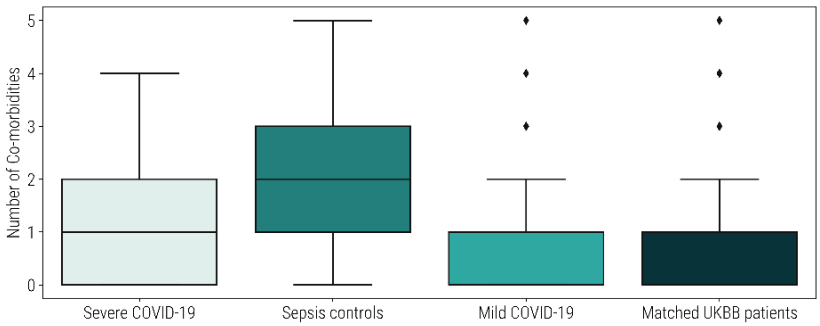


Figure 1: Boxplot showing the incidence of five co-morbidities associated with high COVID-19 risk (cardiovascular disease, hypertension, diabetes, chronic respiratory diseases and Alzheimer's disease) among the May 18, 2020 severe COVID-19 patients, sepsis controls, mild COVID-19 patients and 77,900 randomly selected UK Biobank patients that were age and gender-matched with the severe COVID-19 patients.

Due to the limited number of patients with confirmed mild disease we were unfortunately not able to use this cohort directly as our control set due to its small size and lack of statistical power. We therefore adopted similar control cohort criteria to our

medRxiv preprint doi: https://doi.org/10.1101/2020.06.17.20134015; this version posted June 19, 2020. 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. It is made available under a CC-BY-NC 4.0 International license .

previous sepsis study[4], based on evidence of shared risk factors and that severe COVID-19 patients often present with a similar pattern of co-morbid chronic conditions to those with sepsis. We selected patients who had not developed sepsis in spite of having been exposed to the most common sepsis-causing pathogens as well as having at least one of the most common chronic comorbidities known to increase a patient's risk of developing sepsis and COVID-19 (Figure 2).

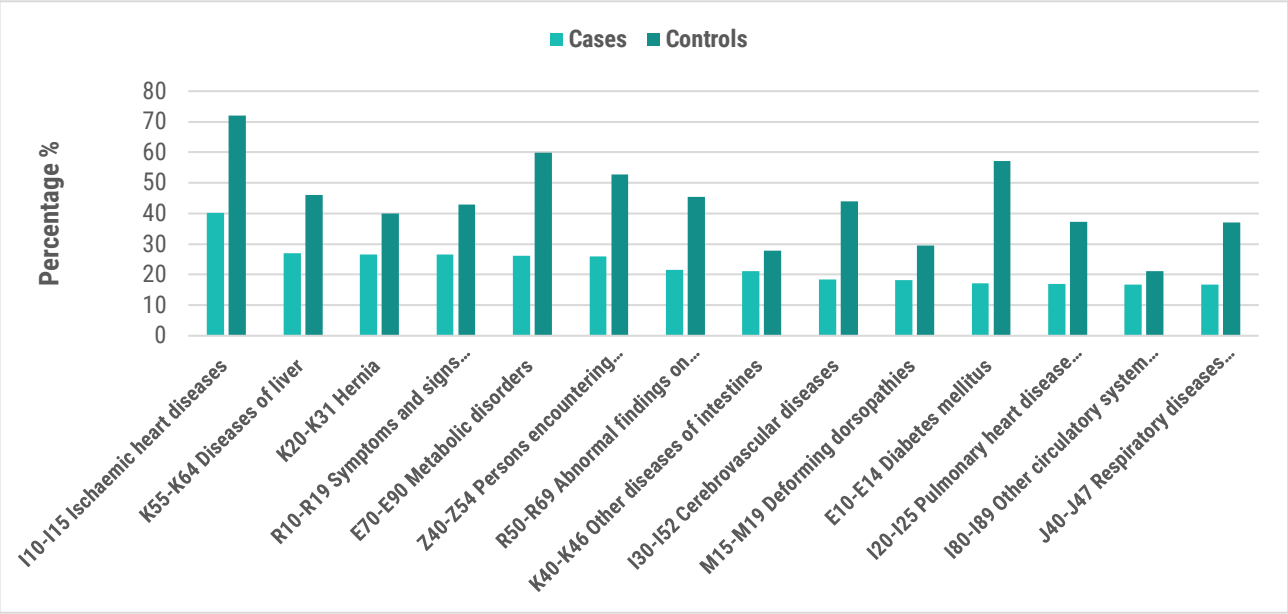


Figure 2: Incidence of co-morbidities by ICD-10 code from UK Biobank in the May 18, 2020 severe COVID-19 dataset (n= 2,332, 1:2 case control ratio with 779 cases and 1,553 controls).

We selected the oldest possible patients (as age is also a critical phenotypic risk factor for the severe form of COVID-19) with the highest number of co-morbidities. Controls were therefore selected for a lack of disease given maximum risk and exposure, and were gender-matched against the severe COVID-19 cases in the ratio 2:1. The exact control criteria and distribution of the co-morbidities in cases are described in the Appendix.

This set of patients represents the most similar available control cohort who could reasonably be expected to have some genetic protective effect against developing severe COVID-19 (or to lack any such risk factors). By selecting controls with a higher prevalence of similar chronic co-morbidities, we seek to ensure that the severe COVID-associated signal observed is not simply caused by the co-morbidities represented in the patient set but has the potential to be a true COVID-19 related enrichment.

After quality control (removal of SNPs with <95% coverage across subjects), the genotype data for the two cohorts contained 542,245 SNPs. The age distribution of cases and controls is shown in Figure 3.

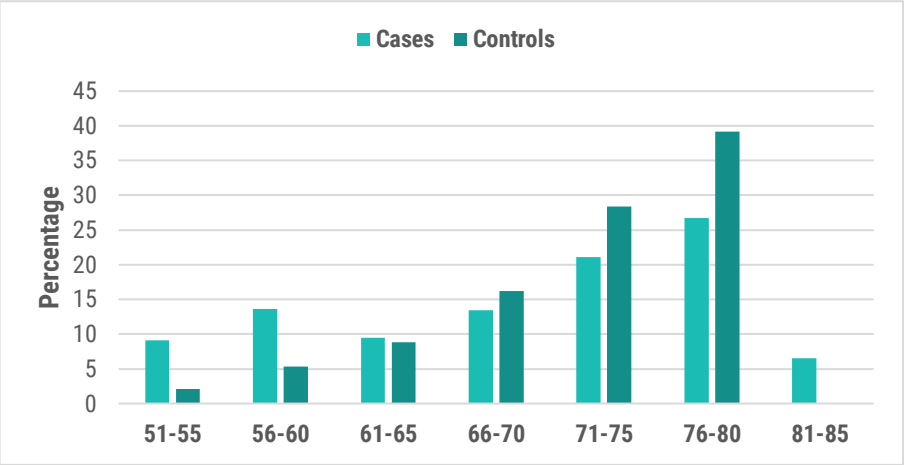


Figure 3: Age distribution of cases vs controls from UK Biobank in the May 18, 2020 severe COVID-19 dataset (779 cases and 1,553 controls).

Having generated our case-control dataset, we analyzed it using the **precision**life platform to identify risk-associated SNPs and genes that were strongly associated in the severe COVID-19 disease cohorts. This platform identifies high-order, disease-associated combinations of multi-modal (e.g. SNPs, transcriptomic, epidemiological or clinical) features at whole genome

medRxiv preprint doi: https://doi.org/10.1101/2020.06.17.20134015; this version posted June 19, 2020. 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. It is made available under a CC-BY-NC 4.0 International license .

resolution in large patient cohorts. It has been validated across multiple different disease populations[9,10,11]. This type of analysis is intractable to other existing methods due the combinatorial explosion posed by the analysis of large numbers of patients with combinatorial non-linear additive combinations of features per patient.

When applied to genomic data, **precision**life finds high-order epistatic interactions (multi-SNP genotype signatures - typically of combinatorial order between 3-8) that are significantly more predictive of patients' phenotype than those identified using existing single SNP based methods. When the individual SNPs making up these signatures are assessed individually across the whole population, they may fall below the GWAS (genome-wide association study) significance thresholds. However, we have demonstrated that when evaluated in combination with each other using multiple statistical validation techniques, these SNPs can be highly significant in particular disease sub-populations. The phenotype to which the signatures are associated in this context might be disease status, progression rate, therapy response or other, depending on the data available and study design.

We used the platform to find and statistically validate combinations of SNPs that together are strongly associated with the severe COVID-19 disease diagnosis. Analysis and annotation of these COVID-19-associated combinatorial genomic signatures took less than a day to complete, running on a dual CPU, 4 GPU compute server. The signatures identified by the analysis were then mapped to the human reference genome[12] to identify disease-associated and clinically relevant target genes. A semantic knowledge graph derived from multiple public and private data sources was used to annotate the SNP and gene targets, including relevant tissue expression, chemical activity/tractability for gene targets, functional assignment and disease-associated literature. This provides contextual information to test the targets against the 5Rs criteria of early drug discovery[13] and allows us efficiently to form strong, testable hypotheses for their mechanism of action in driving severe, life-threatening host responses to COVID-19 infection.

All of the significant disease signatures were traced back to the cases in which they were found and were associated with selected attributes such as case co-morbidities and COVID-19 test records. This generated a high-resolution stratification of severe COVID-19 patient subgroups and enabled further analysis of the different underlying factors relating to their specific forms of the disease.

Subsequently, UK Biobank added the COVID-19 test records of 1,508 individuals in data release 26 May, 2020 and 1,607 individuals in data release 6 June, 2020[8]. Of these, 401 patients were tested positive and 158 of them were hospitalized. We repeated our analysis on these two updated datasets with a higher number of controls to make the analysis more robust (see Appendix). 50 (~70%) of the genes identified in our initial analysis were identified in more than one subsequent analysis (highlighted in bold in the gene tables) and therefore have an additional replication.

Results

Using the severe COVID-19 dataset to perform a standard PLINK[14] GWAS analysis revealed no significant SNPs (Figure 4) using a genome-wide significance threshold of *p*<5e-08. The lowest SNP PLINK *p*-value reported was 9.02e-08.

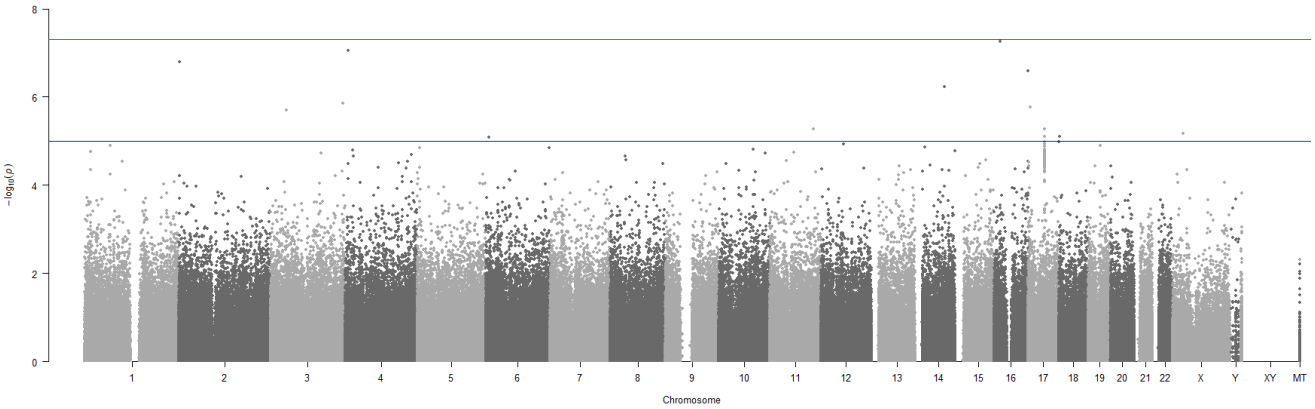


Figure 4: Manhattan plot generated using PLINK of genome-wide *p*-values of association for the May 18, 2020 severe COVID-19 UK Biobank cohort. The horizontal red and blue lines represent the genome-wide significance threshold at *p*<5e-08 and *p*<1e-05 respectively.

When the same May 18, 2020 severe COVID-19 patient dataset was run using the **precision**life platform, we identified 3,515 combinations of SNP genotypes representing different combinations of SNP genotypes that were highly associated with the severe COVID-19 patient cohort (Table 1). The majority (n = 3,494) of SNPs were found in combinations with 5 or more SNP genotypes, and as such could not have been found using standard GWAS analysis methods.

medRxiv preprint doi: https://doi.org/10.1101/2020.06.17.20134015; this version posted June 19, 2020. 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. It is made available under a CC-BY-NC 4.0 International license .

Table 1: Summary of May 18, 2020 Severe COVID-19 Cases vs Sepsis Controls disease study.

	Severe COVID19/Sepsis Study
FDR	5%
Disease signatures	3,515
SNPs in all Disease signatures	5,402
Penetrance (cases represented by all signatures)	100%
RF scored SNPs	156
RF scored Genes	71

All of the SNP genotypes and their combinations were scored using a Random Forest (RF) algorithm based on a 5-fold cross-validation method to evaluate the accuracy with which the SNP genotype combinations predict the observed case: control split. 156 SNPs were scored by the RF algorithm, indicating that they accurately predict the differences between cases and controls (Figure 5A). The chromosomal distribution of the critical SNPs is shown in aggregate in Figure 5B.

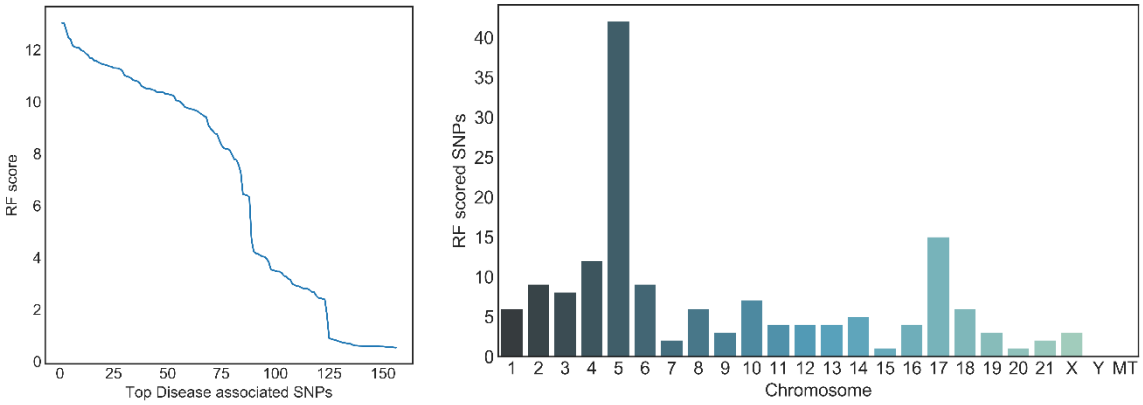


Figure 5: Distributions of (A) RF scores and (B) chromosomal locations for critical disease associated SNPs

Clustering the SNPs by the patients in whom they co-occur allows us to generate a disease architecture of severe COVID-19 patients, providing useful insights into patient stratification. We can use this to find genes and biological pathways that are associated with patient sub-populations and co-morbidities, enabling the development of disease biomarkers and precision medicine strategies.

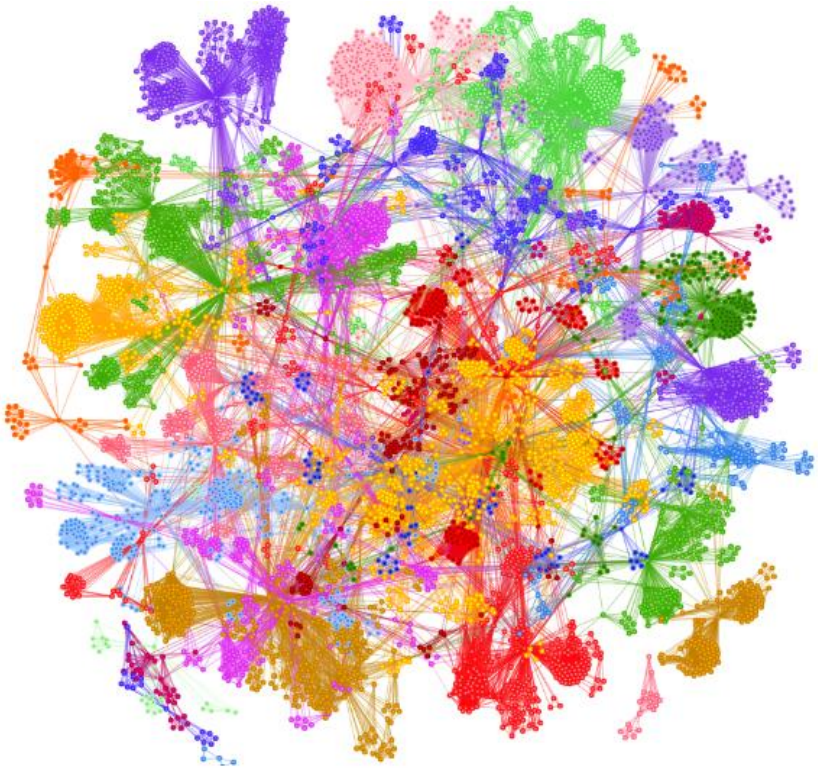
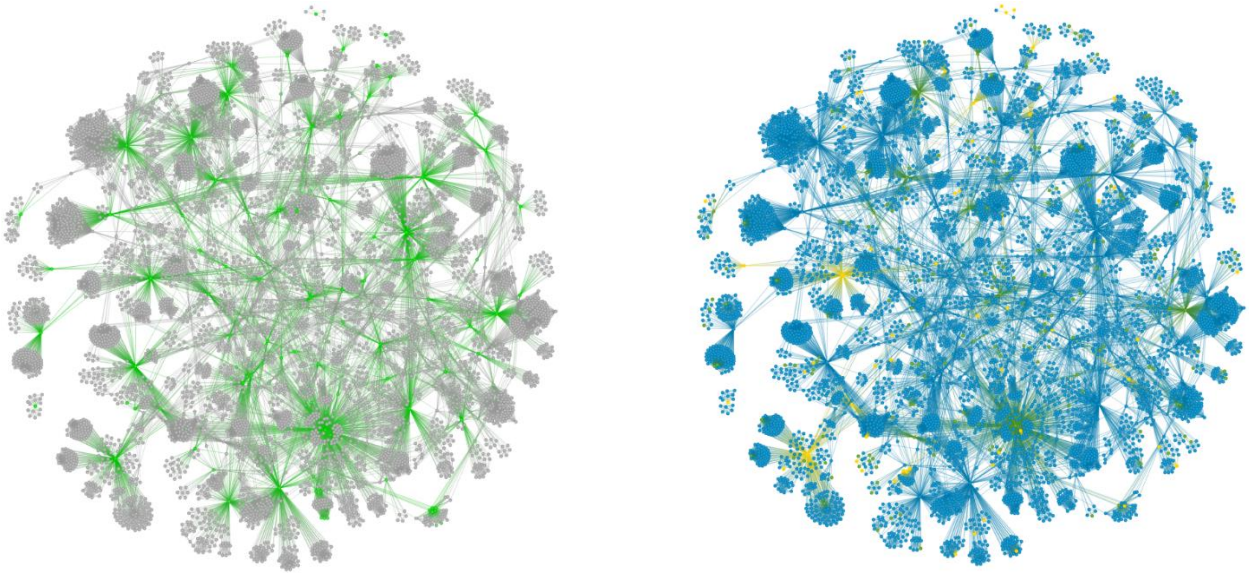


Figure 6: Disease architecture of the severe COVID-19 patient population generated by the **precision**life platform. Each circle represents a disease-associated SNP genotype, edges represent co-association in patients, and colours represent distinct patient sub-populations.

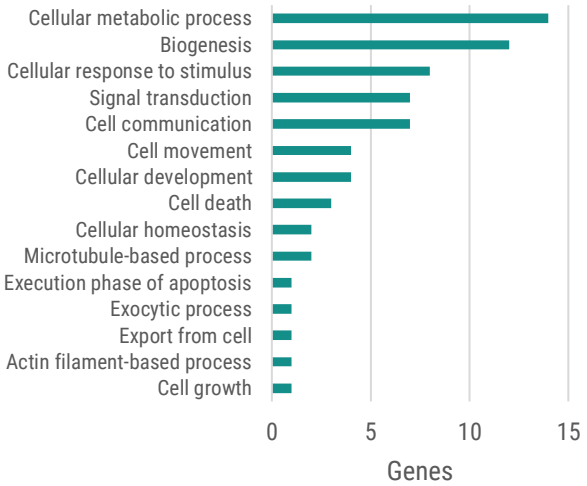
medRxiv preprint doi: https://doi.org/10.1101/2020.06.17.20134015; this version posted June 19, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .



(A) (B)

Figure 7: Disease architecture of the severe COVID-19 patient population highlighting (A) the critical disease SNPs (green) and (B) showing SNP genotypes (right - homozygous major allele = blue, heterozygous = green, homozygous minor allele = gold).

(A)



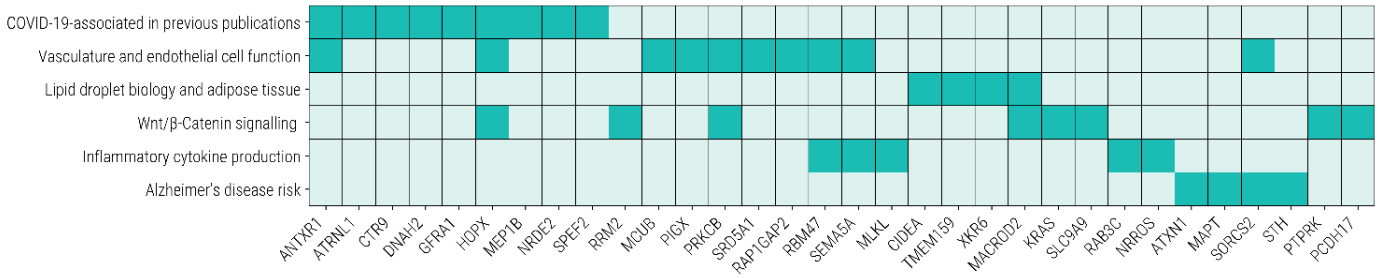
(B)

Figure 8: (A) Distributions of chromosomal locations for RF scored Genes. (B) Functional categories of genes defined by high-level Gene Ontology[15] terms.

medRxiv preprint doi: https://doi.org/10.1101/2020.06.17.20134015; this version posted June 19, 2020. 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. It is made available under a CC-BY-NC 4.0 International license .

Our analysis identified 68 protein-coding genes that were strongly associated with the disease phenotype in patients who developed severe COVID-19. To date, nine of these genes have already been linked to differential COVID-19 host responses in various studies, providing validation for our hypothesis-free approach comparing severe COVID-19 patients against sepsis-free controls. We found several biological pathways and processes that were common in across the 68 COVID-associated genes, including T cell regulation and host pathogenic responses, inflammatory cytokine production, and lipid formation and endothelial cell function (Figure 9).

Figure 9: Biological pathways and processes known to be associated with some of the genes replicated across the datasets used in this study



We identified twelve genes that are associated with host immune response to viral infections, including SARS-CoV-2 (Table 2). Variants in several of these genes have been associated with increased infectivity to these strains. Genes marked in bold represent those that were further validated in subsequent analyses using the updated severe COVID-19 cases.

Table 2: Severe COVID-19 risk-associated genes associated with host immune response to virus infection, listed alphabetically. Bold text indicates genes found in multiple study populations.

GENE	FUNCTION	MECHANISM OF ACTION
ANTXR1[16,17]	Anthrax toxin receptor 1 (TEM8), type I transmembrane protein	ANTXR1 is used by anthrax toxin and Seneca valley virus (SVV) to gain entry into host cells
ATRNL1[18]	Attractin-like 1	Phenotypes associated with decreased *Salmonella* invasion, hepatitis C virus replication and vaccinia virus infection
CTR9	Paf/RNA polymerase II complex component	Acts as a host restriction factor, suppressing HIV replication
DNAH2[19]	Dynein heavy chain	Expressed in lung cilia, may be responsible for effective pathogenic clearance
GFRA1[20]	GDNF family receptor alpha	SNPs associated with susceptibility to hepatitis C and *Staphylococcus aureus* infection
HOPX	Hop homeodomain protein, cofactor	High expression in expanded CD8+ T cells, improving SARS-CoV-2 viral clearance
IKZF2	Ikaros family member, Helios	Helios expression is high in regulatory T cells, necessary for self-tolerance and prevention of autoimmunity
ITK	IL-2 inducible T cell kinase, Th2 differentiation	Altered expression associated with SARS-CoV-2 infectivity
MEP1B	Meprin A subunit beta, membrane metallopeptidase	Macrophages in *Mep1b-/-* mice have greater phagocytic function. Biomarker for high tuberculosis burden.
NRDE2[21]	RNA interference	RNA interference is a mechanism utilized by the innate immune response. NRDE2 is used by Kaposi's sarcoma-associated herpesvirus (KSHV) for late gene expression
SPEF2[22]	Sperm flagellar 2	Loss of *Spef2* in mice results in increased inflammatory response to *Streptococcus pneumonia* infection due to defects in pulmonary cilia.
TBC1D2	TBC1 domain family member, regulates Rab7A activation	SNP associated with severe H1N1 infection

Our analysis also revealed five genes regulating pro-inflammatory pathways such as necroptosis, reactive oxygen species (ROS) production and cytokine signalling (Table 3).

medRxiv preprint doi: https://doi.org/10.1101/2020.06.17.20134015; this version posted June 19, 2020. 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. It is made available under a CC-BY-NC 4.0 International license .

Table 3: Severe COVID-19 risk-associated genes that have a role in regulation of inflammatory cytokine production

GENE	FUNCTION	MECHANISM OF ACTION
MLKL[23]	Mixed lineage pseudokinase	Activator of TNF-induced necroptosis, high expression results in necrotic cell death following coronavirus infection
NRROS[24]	Negative regulator of reactive oxygen species (ROS)	Limits phagocytic ROS and pro-inflammatory cytokine production during host immune response
RAB3C[25]	Rab GTPase	RAB3C increases cells' ability to release IL-6 and activate the JAK2-STAT3 pathway.
RBM47[26]	RNA binding motif protein	Promotes IL-10 production in B cells, repressed by TGF-β
SEMA5A[27]	Semaphorin 5A	High levels of Sema5A are associated with high IFN-Y and low IL-4 expression

There were also six genes associated with severe COVID-19 patients that play central roles in lipid droplet biology, as well as having high expression in adipose tissue, correlating with serum lipid levels and coronary artery disease (Table 4).

Table 4: Severe COVID-19 risk-associated genes that are highly expressed in adipose tissue and relate to lipid storage and signalling

GENE	FUNCTION	MECHANISM OF ACTION
CIDEA[28]	Cell death inducing DFFA like effector	Co-localizes with perilipins, highly expressed in brown adipose tissue
MACROD2[29]	Mono-ADP ribosylhydrolase	Highly associated with VAP-1 in human adipose tissue
PLIN4[30]	Perilipin 4	Lipid droplet protein, highly expressed in white adipose tissue, high expression results in cardiac steatosis
SLIT3[31]	Slit guidance ligand 3	High expression in adipose tissue controlled by *Prdm16* expression. Upregulated in obese patients.
TMEM159[32,33]	Promethin	Activated by peroxisome proliferator-activated receptor γ1 (PPARγ1), partners with seipin to regulate lipid droplet organization
XKR6[34]	XK-related 6	SNP association with serum lipid levels and risk of CAD

In addition to this, our analysis revealed twelve more genes that may be implicated in the vascular complications seen in COVID-19 (Table 5). These genes are all highly expressed in the cardiovascular system, modulating cardiac function and endothelial cell homeostasis.

Table 5: Severe COVID-19 risk-associated genes that are implicated in cardiovascular and endothelial cell function

GENE	FUNCTION	MECHANISM OF ACTION
ANTXR1[35]	Anthrax toxin receptor 1, type I transmembrane protein (TEM8)	Loss of ANTXR1 results in endothelial basement membrane loss and leaky blood vessels
HOPX[36]	Hop homeodomain protein, cofactor	Acts via serum response factor (SRF), modulating the expression of cardiac genes and stress-induced cardiac hypertrophy
MCUB[37]	Mitochondrial calcium uniporter	Role in cardiac homeostasis, response to cardiac stress including ischemia
PIGX[38]	Phosphatidylinositol glycan anchor biosynthesis class X	Associated with hypertension. Forms a complex with PIG-M that regulates several processes, including maintenance of arterial blood pressure. Acts via apelin receptor signalling.
PRKCB[39]	Protein kinase C beta form, protein kinase	Variety of different cellular functions, including endothelial cell proliferation and insulin signaling
PLS3[40]	Plastin-3	Induced by angiotensin II in endothelial cells to promote cell migration
RAP1GAP2[41,42]	GTPase activator for RAP-1A	Regulates platelet granule secretion and aggregation
RBM47[43]	RNA binding motif protein	Missense variant is associated with hypertension
SEMA5A[44]	Semaphorin 5A	High Sema5A expression increased endothelial cell proliferation and angiogenesis and decreased apoptosis
SLIT3[45]	Slit guidance ligand 3	Upregulation associated with endothelial cell dysfunction in pulmonary hypertension
SORCS2[46]	Sortilin-related VPS10 domain-containing receptor 2 (SorCS2), oxidative stress response gene	SorCS2 releases endostatin, an endogenous inhibitor of endothelial cell proliferation and angiogenesis
SRD5A1[47]	Steroid 5 alpha-reductase 1, catalyzes the production of androgen dihydrotestosterone	Androgens have been linked to peripheral artery disease and macrophage modulation

medRxiv preprint doi: https://doi.org/10.1101/2020.06.17.20134015; this version posted June 19, 2020. 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. It is made available under a CC-BY-NC 4.0 International license .

Amongst other cancer-associated genes, we identified nine risk-associated genes that directly interacted with the Wnt/β-catenin signaling pathway (Table 6). With the exception of KRAS and SLC9A9, all of the proteins encoded by these genes act as endogenous inhibitors of the pathway.

Table 6: Severe COVID-19 risk-associated genes that directly interact with the Wnt/β-Catenin signalling pathway

GENE	FUNCTION	MECHANISM OF ACTION
HOPX	Hop homeodomain protein, cofactor	Promotes BMP-mediated inhibition of Wnt signaling pathway
KRAS[48]	KRAS proto-oncogene, GTPase	Activates the Wnt/β-catenin pathway
MACROD2[49]	Mono-ADP ribosylhydrolase	Suppresses GSK-3β/β-catenin signaling
PCDH17[50]	Protocadherin 17, member of the cadherin superfamily	Acts as a tumour suppressor in breast cancer, inhibiting the Wnt signaling pathway
PRKCB[51]	Protein kinase C beta form, protein kinase	Promotes the phosphorylation of β-catenin
PTPRK[52]	Protein tyrosine phosphatase (PTP) receptor, type K	Redistributes and inhibits the transcriptional activity of β-catenin
RRM2[53]	Ribonucleotide reductase regulatory subunit M2	Downregulation of RRM2 suppresses the activity of β-catenin and the Wnt signaling pathway
SLC9A9[54]	Solute carrier family 9	Upregulates beta-catenin
SLIT3[55]	Slit guidance ligand 3	Suppresses GSK3β/β-catenin pathway

Finally, our analysis identified four genes that have previously been associated with increased risk of developing Alzheimer's disease (AD), including *MAPT,* the key gene underlying tau pathology (Table 7).

Table 7: Severe COVID-19 risk-associated genes that have also been found to confer increased Alzheimer's disease risk

GENE	FUNCTION	MECHANISM OF ACTION
ATXN1	Ataxin-1, unknown function	Loss of ataxin-1 causes increased BACE1 expression and amyloid precursor protein (APP) cleavage
MAPT	Microtubule-associated protein tau, stabilizes microtubules	Tau pathology is a hallmark of AD and other neurodegeneration diseases
SORCS2	Sortilin-related VPS10 domain-containing receptor 2, oxidative stress response gene	SNPs are associated with AD, result in altered APP processing
STH	Saitohin, unknown function	Q7R polymorphism increases risk of AD and other neurodegenerative diseases. Interacts with tau.

Out of the total 68 genes, nine are targeted by clinical candidates that have been evaluated in Phase I clinical trials and beyond (see Appendix). These could potentially form the basis of repurposing therapies, after evaluating factors such as safety and pharmacology data. A further eight targets have active chemistry in ChEMBL[56], meaning they have active chemical starting points for novel drug development (see Appendix). We can also extend our search for repurposing candidates by looking into known gene interactions to find other more tractable targets in implicated in the same biological pathway.

medRxiv preprint doi: https://doi.org/10.1101/2020.06.17.20134015; this version posted June 19, 2020. 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. It is made available under a CC-BY-NC 4.0 International license .

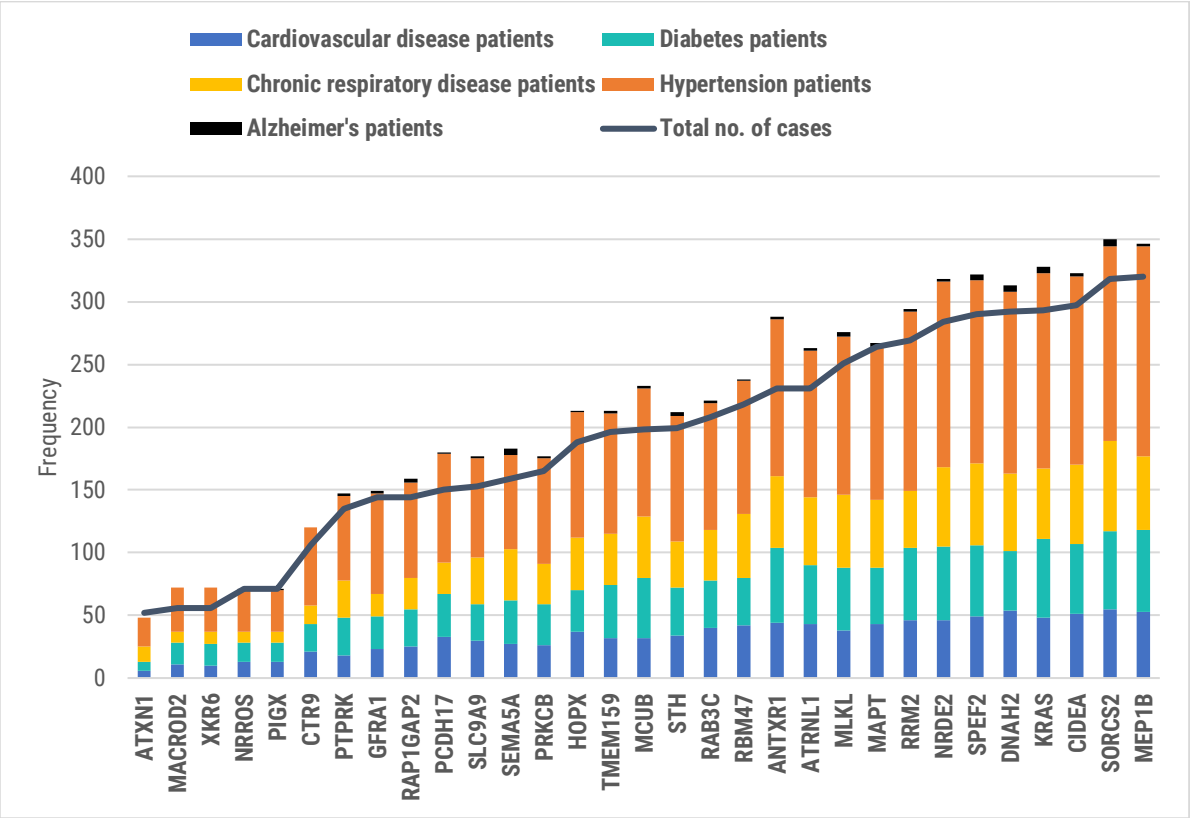


Figure 10: Stacked bar plots showing the number of cases with different co-morbidities (cardiovascular disease, hypertension, diabetes, chronic respiratory diseases and Alzheimer's disease) associated with the severe form of COVID-19 who are affected by the risk-associated genes identified by the **precision**life platform. The line plot shows the total unique number of cases who are affected for each gene.

medRxiv preprint doi: https://doi.org/10.1101/2020.06.17.20134015; this version posted June 19, 2020. 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. It is made available under a CC-BY-NC 4.0 International license .

Discussion

Grouping the 68 genes by common biological functions revealed that many are involved in processes that have also been linked to aberrant host responses to COVID-19, such as the pro-inflammatory cytokine storm and immune system dysregulation, as well as lipid droplet formation and endothelial cell dysfunction.

HOST RESPONSE FACTORS AND INFLAMMATION

The host immune response must maintain a balance between effective viral clearance and limiting the immune response to prevent chronic inflammation and collateral tissue damage. In many patients who develop severe COVID-19 reactions and acute respiratory distress syndrome (ARDS) there is evidence of dysregulated cytokine production, resulting in increased levels of pro-inflammatory mediators including IL-1, IL-6, IL-8, CXCL-10[57], causing pathological features such as inflammatory cell infiltration, pulmonary edema and sepsis[58].

From the number of genes we found relating to immune response pathways and cytokine production cascades, it is clear that patients who develop severe COVID-19 and ARDS may have innate genetic variants that prevent this balance from being struck. These variants were found in equal proportions across all severe COVID-19 patients, regardless of their co-morbidities. This suggests that variants in these genes may not be associated with any specific co-morbidity, but are common amongst many patients who develop severe COVID-19.

HOPX regulates a variety of different cellular processes, including cardiac development and myogenesis[36]. However, it was reported in a recent COVID-19 study as part of a selection of genes upregulated in expanded CD8+ T cells in patients with mild COVID-19[59]. These expanded CD8+ effector T cells likely represent SARS-CoV-2-specific T cells, indicating greater efficiency in viral clearance in those patients. However, it is also necessary for Th-1 persistence and resistance to apoptosis, driving chronic inflammation and autoimmune mechanisms[60]. This exemplifies the delicate balance between establishing an effective host immune response necessary for viral clearance and preventing the development of chronic inflammation that results in collateral tissue damage. It seems likely that patients who develop severe COVID-19 may have an inherent imbalance in these factors.

Similarly, we identified SNP variants in *ITK* that were highly associated with severe COVID-19. High *ITK* expression is associated with Th2-driven diseases such as allergic asthma, causing pro-inflammatory cytokine production, eosinophil infiltration and production of mucus[61]. A range of potent and selective inhibitors (both small molecule and biologic) of ITK have been developed by several pharmaceutical companies[62,63]. No selective ITK inhibitors have progressed beyond preclinical testing to date, although ibrutinib – a joint ITK and BTK inhibitor – is currently licensed for use in B cell malignancies and as also demonstrated efficacy in the treatment of leishmaniasis. Data collected from several clinical trials indicate that ibrutinib is reasonably well-tolerated by patients[64]. Inhibition of ITK diminishes lung injury, cytokine production and inflammation in a mouse model of asthma[65]. Many of the pro-inflammatory cytokines, such as IFNγ, IL-2, and IL-17, blocked by ITK inhibition are raised in patients with ARDS[66,67]. Furthermore, use of a selective ITK inhibitor blocked HIV cell entry, transcription and particle formation, effectively reducing viral replication[68].

A study has found that MLKL is implicated in necrotic cell death following infection with a neurovirulent human coronavirus (HCoV)[69]. Whilst MLKL-induced necrosis may be useful in limiting viral replication, increased necroptosis can also lead to increased inflammation and tissue damage. SNP variants in *MLKL* could be predisposing patients to severe COVID-19 in one of two ways; with under-functioning protein resulting in poor viral clearance, or through over-expression and over-activation of the necroptotic inflammatory response resulting in organ damage. The SNP variant identified in *MLKL* was found in the highest scoring 20% of the significant genes in the severe COVID-19 dataset and is present in 251 cases. It is not highly associated with any of the most common co-morbidities found in this dataset. Necrosulfonamide (NSA) is a specific inhibitor of MLKL that potently suppresses necroptosis[70]. Inhibition of this pathway using NSA decreases pro-inflammatory cytokines such as IL-1β, IL-6 and IL-17A in a way that was protective in a model of psoriasis. High levels of IL-1β and IL-6 have both been observed in cases with severe COVID-19[71]. Unfortunately, NSA has only ever been evaluated in preclinical trials, so there is no safety or toxicology data in humans available.

NRROS (LRRC3) is an inhibitor of multiple toll-like receptors (TLR) and subsequent NF-κB signalling, acting as a brake on pro-inflammatory cytokine production[72]. NRROS also limits the amount of reactive oxygen species (ROS) produced by phagocytes during the innate immune response, thereby limiting tissue damage caused whilst defending against invading pathogens. It could therefore be hypothesized that SNP variants in this gene, limiting its activity, could make individuals more susceptible to COVID-19-induced inflammatory damage[73]. This is supported by the evidence that *Lrrc33-/-* mice suffered from increased organ damage as a result of greater pro-inflammatory signalling when challenged with LPS. There are currently no specific small molecule activators of NRROS in publicly-available databases, however a number of chemical agents have been shown to increase the

medRxiv preprint doi: https://doi.org/10.1101/2020.06.17.20134015; this version posted June 19, 2020. 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. It is made available under a CC-BY-NC 4.0 International license .

expression of both the protein and mRNA forms of this gene which may provide chemical starting points for novel drug discovery[74].

RAB3C may be driving inflammation by inducing the release of IL-6 and activation of the IL-6/JAK2/STAT3 pathway, driving the production of pro-inflammatory cytokines[25,75]. Ruxolitinib – a JAK2 inhibitor – has been shown to mitigate the effects of RAB3C-induced IL-6 release[25]. Ruxolitinib is currently being trialled as a treatment for respiratory distress caused by SARS-CoV-2[76].

VASCULAR INFLAMMATION, AUTOPHAGY AND CARDIOVASCULAR DYSFUNCTION

Endothelial dysfunction and vascular inflammation, resulting in neutrophilic infiltration, endothelial cell apoptosis and tissue edema, is seen in patients who develop ARDS from SARS-CoV-2 infection[77]. Many of the common co-morbidities associated with severe COVID-19 development, - such as diabetes, hypertension and cardiovascular disease – are already associated with vascular inflammation and endothelial cell dysfunction pathological features[78,79,80]. It is therefore unsurprising that these patients are at higher risk of developing severe COVID-19.

Many of the risk-associated SNPs that we found that were related to these pathological pathways were found in severe COVID-19 patients who had at least one of these co-morbidities. However, we controlled for this effect by co-morbidity-matching these cases against the sepsis-free controls, meaning that these SNPs are not likely to be just be artefacts of the co-associated co-morbidity, but a differentiating factor in the development of severe COVID-19 and ARDS.

Helios (*IZKF2*) expression is used as a marker for regulatory T cells (Tregs), and therefore has an important role in self-tolerance and regulating autoimmunity[81]. Decreased levels of Helios+ Tregs have been observed in patients with hypertension and rheumatoid arthritis, with lower expression levels correlating with increased inflammatory markers[82]. Tregs may protect against hypertension by limiting vascular inflammation through suppression of effector T cells[83,84]. Furthermore, our analysis revealed that variants in *IZKF2* were disproportionately co-associated with patients with hypertension (Figure 10). This adds further evidence to the association between vascular inflammation, severe COVID-19 and cardiovascular hypertensive co-morbidities.

MCUB encodes one of the pore-forming subunits of the mitochondrial Ca^{2+} uniporter (MCU). MCUB is a necessary part of a protective response against mitochondrial Ca^{2+} overload during cardiac injury and ischemia[85]. *Mcub -/-* mice displayed increased cardiac remodelling and ischemic injury as a result of increased mitochondrial Ca^{2+} uptake[86]. Although more research is required to fully understand the role of MCUB in severe COVID-19 pathogenesis, a decreased level of serum calcium has been suggested as a biomarker for increased COVID-19 severity and ARDS[87].

PKCβ (*PRKCB*) is a serine/threonine protein kinase that is activated by calcium. It has a range of functions, including B cell activation, endothelial cell proliferation and activation of apoptosis[35]. However, PKCβ has also been linked to a number of different vascular diseases, including atherosclerosis, diabetes and hypertension[88,89]. We find that in our case population, the risk-associated SNP found in *PRKCB* was found in 165 severe COVID-19 cases (penetrance = 20.2%), and was present in 45% of patients with cardiovascular disease and 51% of patients with hypertension. High levels of PKCβ result in increased vascular inflammation, endothelial dysfunction and oxidative stress, all of which have been found in patients with severe COVID-19[90,91]. PKCβ also drives the accumulation of cholesterol in macrophages, leading to foam cell development and macrophage dysregulation[92]. Therefore, inhibition of PKCβ could help to reverse some of the vascular-related pathology, contributing to sepsis development and multi-organ failure, that is seen in severe COVID-19 patients. Ruboxistaurin (LY333531) is a selective PKCβ inhibitor that has been trialled as an anti-diabetic drug to reduce vascular and retinal complications that certain diabetic patients develop[93]. It has reasonable pharmacokinetic and toxicology data, as it can be orally administered and is well-tolerated by patients[94,95]. However, it was discontinued and has failed to progress beyond Phase III clinical trials[96].

We also found genes, including *CIDEA* and *PLIN4*, that protect against insulin resistance (IR). IR results in increased cardiovascular inflammation and oxidative stress, contributing to atherosclerotic plaque formation and hypertension[97]. High Cidea expression decreases circulating fatty acid levels by increasing the level of triglycerides stored in lipid droplets. This helps to protect against insulin resistance[98]. The anti-diabetic drug rosiglitazone may have insulin sensitising effects by increasing the expression of both Cidea and perilipins[28]. Our analysis identified *CIDEA* and a perilipin, *PLIN4*, as being highly associated with severe COVID-19 patients. SNPs in *CIDEA* were found in one of the highest proportions in severe COVID-19 patients, particularly in patients with cardiovascular and hypertensive diseases (Figure 10). This indicates that insulin resistance is a contributing factor to vascular inflammatory pathologies that may predispose patients to developing ARDS in response to SARS-CoV-2.

Finally, variants in *GFRA1* are within the highest 10% of genes most associated with hypertensive patients in our case population. GFRα1, encoded by *GFRA1*, plays a key role in glial cell line-derived neurotrophic factor (GDNF)-mediated signalling[99]. GFRα1 induces autophagy via activation of the RET/AMPK signalling pathway independently from its interaction with GDNF. In the early stages of sepsis, autophagy is a protective mechanism employed by cells to remove damaged proteins, reduce mitochondrial dysfunction-induced inflammation and eliminate pathogens[100,101]. Furthermore, inhibition of autophagy in mouse models of

medRxiv preprint doi: https://doi.org/10.1101/2020.06.17.20134015; this version posted June 19, 2020. 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. It is made available under a CC-BY-NC 4.0 International license .

sepsis results in increased mortality. Selective agonists of GFRα1 are under development[102], which may help to protect against sepsis-induced tissue damage in the early stages of the disease.

LIPID DROPLET BIOLOGY

There is increasing evidence that lipid dysfunction plays a role in COVID-19 pathogenesis. Lipid droplets (LD) play a key role in viral pathogenesis, as intracellular lipid stores are crucial for viral replication and assembly[103]. LD proteins have been associated with particle load and pathogenicity in other virus strains, including hepatitis C, dengue and rotavirus[104]. It has also been shown that coronaviruses use host cellular lipid machinery during replication[105].

We identified six genes associated with lipids and adipose tissue, with three genes directly involved in lipid droplet formation (Table 4). A different member of the perilipin family (PLIN3) is used by hepatitis C virus (HCV) for steatosis development and viral assembly[106], with PLIN3 inhibition resulting in decreased viral particle release from host cells[107]. Promethin (*TMEM159*) has also been recently identified as lipid droplet organiser (LDO) in partnership with seipin[33]. Although promethin has not been studied in the context of viral disease, overexpression of its co-associated protein seipin results in significantly reduced viral particle secretion and infectivity in a model of hepatitis C[108].

WNT/β-CATENIN SIGNALLING PATHWAY

Our analysis identified nine genes that regulate either the Wnt or β-catenin components of the Wnt/β-catenin pathway.

There is a long-established role of the Wnt gene-family and the associated signally pathway in embryogenesis and therefore its relationship to developmental disorders and carcinogenesis. However recent work[109,110] shows an emerging and complex picture of the role of Wnt genes in host cell defence mechanisms, the modulation of inflammatory cytokine production and connections between innate and adaptive immune systems. The role of these ligands varies within the family members and their relation to β-catenin.

Wnt/β-catenin signalling has both anti- and pro-inflammatory effects in different contexts, dependent on its interaction with NF-κB[109]. In several studies, activation of β-catenin was shown to inhibit IL-1β and the subsequent production of IL-6 and matrix metalloproteinases (MMPs), resulting in an anti-inflammatory effect[110,111]. Although its role in sepsis pathogenesis is yet to be fully elucidated, patients with severe sepsis and sepsis-driven ARDS had increased levels of Wnt5A in lung tissue and serum[112].

However, in bronchial epithelial cells, inhibition of beta-catenin's activation of NF-κB resulted in fewer proinflammatory cytokines and cell injury[113]. In a previous unpublished analysis using the same combinatorial approach on a Sjögren's Syndrome cohort found in the UK Biobank, we identified a significant number of genes involved in the Wnt/β-catenin pathway. As with severe COVID-19, patients with Sjögren's syndrome also present with elevated cytokines and inflammation-driving pathology[114].

The Wnt/ β -catenin pathway has also been implicated in promoting viral replication. In a model of influenza (H1N1), activation of this pathway resulted in higher viral replication and inhibition of Wnt/ β -catenin signalling using iCRT14 reduced viral production and improved clinical symptoms[115]. It has been demonstrated that Rift Valley Virus exploits the proliferative cell state that activation of Wnt signalling promotes to enable viral replication through easier trafficking of viral proteins within the cell. It has also been shown that coronaviruses use this mechanism in proliferative cells[116]. Therefore, inhibition of the Wnt/β-catenin pathway may help limit SARS-CoV-2 replication and reduce viral load in this way. This theory is supported by the fact that niclosamide – an anti-helminthic drug – limits coronavirus replication in a model of SARS[117]. The study did not investigate how this drug inhibited viral replication, but a subsequent study has revealed that niclosamide is an inhibitor of the Wnt signalling pathway[118]. Due to the significant number of genes found to regulate this pathway in our analysis, we believe that modulation of this pathway could be of benefit to a large number of patients who develop severe COVID-19.

NEURODEGENERATIVE DISEASE ASSOCIATIONS

There is emerging evidence that key genes associated with Alzheimer's Disease (AD) risk such as *APOE4* are associated with increased risk of severe COVID-19 disease, although the reasons for this link are still unclear[119]. Our analysis identified a strong conserved signal for four genes that also confer greater risk of developing AD, including *MAPT*.

We also found three additional genes that were significant in a previous neurodegenerative disease study we performed. This unpublished study identified 33 genes variants in which were highly associated in cases with sporadic amyotrophic lateral sclerosis (ALS). We have also shown that increasing the activity of two of these genes targets in a mixed cell neuronal assay containing SOD1 reactive astrocytes significantly improves motor neurone survival. This could indicate shared biological pathways that drive both severe COVID-19 and neurodegeneration, potentially through pro-inflammatory, neurotoxic mechanisms. However, the SNP variants found in these genes were not highly enriched in the few severe COVID-19 patients who had been diagnosed with AD. This is probably due to the relatively low numbers of cases diagnosed with AD in UK Biobank as a whole.

medRxiv preprint doi: https://doi.org/10.1101/2020.06.17.20134015; this version posted June 19, 2020. 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. It is made available under a CC-BY-NC 4.0 International license .

Conclusion

We performed this analysis using only genetic data from patients found in the UK Biobank, identifying 68 protein coding genes that are highly associated with the development of severe COVID-19. These targets would not have been found using standard analytical approaches such as GWAS on the same population. In addition to this, we have identified 29 drugs and clinical candidates and a further eight targets with chemical starting points that could be used in the development of treatment strategies that improve clinical outcomes in severe COVID-19 patients.

It appears that the variants we found in genes relating to immune response pathways and cytokine production cascades were in equal proportions across all severe COVID-19 patients, regardless of their co-morbidities. This suggests that such variants are not associated with any specific co-morbidity, but are common amongst patients who develop severe COVID-19. In contrast there were small deviations in the penetrance of some of the SNPs identified in different co-morbidity cohorts. While the pattern of these deviations were similar across cardiovascular, diabetes and hypertension co-morbidity cohorts, the pattern across the respiratory co-morbidity cohort was somewhat different. These differences, while suggestive of the need for further study, were not yet significant enough to be reported in detail.

As more test records and additional medical data become available in the UK Biobank and other data sources, we will be able to fully ascertain the severe and mild COVID-19 cases and provide an additional layer of validation to the results from this study.

One limitation of the UK Biobank dataset is that the ethnicity distribution of the participants is heavily skewed to white British participants and it has consequently not been possible to fully investigate additional risk factors in BAME patients. We are actively seeking to investigate severe COVID-19 risk factors in other datasets with more ethnically diverse populations. The addition of more phenotypic and clinical data in our analyses may also be used to gain greater understanding into the association of other observed epidemiological risk factors such as ethnicity, socioeconomic status and prescription medication history with development of severe COVID-19.

Acknowledgements

We would like to acknowledge UK Biobank for providing us access to the data under Application ID 44288 and Bugbank for linking infection data from Public Health England to improve the study of infection in the UK Biobank cohort.

medRxiv preprint doi: https://doi.org/10.1101/2020.06.17.20134015; this version posted June 19, 2020. 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. It is made available under a CC-BY-NC 4.0 International license .

References

1. Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect Dis;* published online Feb 19, 2020. https://doi.org/10.1016/S1473-3099(20)30120-1.Accessed from https://coronavirus.jhu.edu/map.html
2. World Health Organization, WHO Coronavirus Disease (COVID-19) Dashboard, available from https://covid19.who.int/, last accessed 1 June 2020
3. Semple, M.G. *et al.* Features of 16,749 hospitalised UK patients with COVID-19 using the ISARIC WHO Clinical Characterisation Protocol (preprint posted April 28, 2020) medRxiv preprint doi: https://doi.org/10.1101/2020.04.23.20076042
4. Das, S., Taylor, K., Pearson, M. et al. Identification and Analysis of Shared Risk Factors in Sepsis and High Mortality Risk COVID-19 Patients (preprint posted May 5, 2020, medRxiv preprint doi: https://doi.org/10.1101/2020.05.05.20091918
5. Zhou, F. et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study *Lancet* (Mar 11, 2020) **395**(10229):1054-1062 DOI:https://doi.org/10.1016/S0140-6736(20)30566-3
6. Hajj J, Blaine N, Salavaci J, Jacoby D. The "Centrality of Sepsis": A Review on Incidence, Mortality, and Cost of Care. *Healthcare (Basel).* 2018;**6**(3):90. Published 2018 Jul 30. doi:10.3390/healthcare6030090
7. Bycroft, C., Freeman, C., Petkova, D., Band, G., Elliott, L.T., Sharp, K., Motyer, A., Vukcevic, D., Delaneau, O., O'Connell, J. and Cortes, A., 2018. The UK Biobank resource with deep phenotyping and genomic data. *Nature*, **562**(7726), pp.203-209.
8. Armstrong, J., Rudkin, J., Allen, N., Crook, D., Wilson, D., Wyllie, D., and O'Connell, A. 2020. Dynamic linkage of COVID-19 test results between Public Health England's Second Generation Surveillance System and UK Biobank. Microbial Genomics, doi: 10.1099/mgen.0.000397. Accessed on 18 May 2020 and 8 June 2020.
9. Taylor K, Das S, Pearson M, Kozubek J, Strivens M, Gardner S. Systematic drug repurposing to enable precision medicine: A case study in breast cancer. *Digit Med* 2019;**5**:180-6.
10. Gardner, S., Das, S. & Taylor, K. AI Enabled Precision Medicine: Patient Stratification, Drug Repurposing & Combination Therapies (ebook *'AI in Oncology Drug Discovery & Development'* in press)
11. Mellerup E, Andreassen OA, Bennike B, Dam H, Djurovic S, Jorgensen MB, et al. Combinations of genetic variants associated with bipolar disorder. *PLoS One* 2017;**12**:e0189739
12. Church DM, Schneider VA, Graves T, *et al.* Modernizing reference genome assemblies. *PLoS Biol.* 2011;**9**(7):e1001091. doi:10.1371/journal.pbio.1001091
13. Morgan, P., Brown, D.G., Lennard, S., Anderton, M.J., Barrett, J.C., Eriksson, U., Fidock, M., Hamren, B., Johnson, A., March, R.E. and Matcham, J., 2018. Impact of a five-dimensional framework on R&D productivity at AstraZeneca. *Nat. Rev. Drug Disc.*, **17**(3), p.167
14. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet.* 2007 Sep;**81**(3):559-75.
15. The Gene Ontology Consortium. The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Res.* Jan 2019;**47**(D1):D330-D338.
16. Young AT, Collier RJ. Anthrax Toxin: Receptor Binding, Internalization, Pore Formation, and Translocation. *Annu Rev Biochem.* 2007;**76**:243-265. https://doi.org/10.1146/annurev.biochem.75.103004.142728
17. Miles LA, Burga LN, Gardner EE, Bostina M, Poirier JT, Rudin CM. Anthrax toxin receptor 1 is the cellular receptor for Seneca Valley virus. *J Clin Invest*. 2017;**127**(8):2957-2967. doi:10.1172/JCI93472Liu L, Oliveira NM, Cheney KM, *et al.* A whole genome screen for HIV restriction factors. *Retrovirology*. 2011;**8**:94. doi:10.1186/1742-4690-8-94
18. GenomeRNAi. Phenotype details for gene 26033 (Homo sapiens, ATRNL1). Available from: http://www.genomernai.org/v17/fullPagePhenotypes/gene/26033 [Accessed 12th June, 2020]
19. Beg, J, Evans J, Leigh M, *et al.* Next generation massively parallel sequencing of targeted exomes to identify genetic mutations in primary ciliary dyskinesia. Implications for application to clinical testing. *Genet Med.* 2011;**13**:218-229. https://doi.org/10.1097/GIM.0b013e318203cff2
20. Kato N, Ji G, Wang Y, *et al.* Large-scale search of single nucleotide polymorphisms for hepatocellular carcinoma susceptibility genes in patients with hepatitis C. *Hepatology*. 2005;**42**(4):846-853. doi:10.1002/hep.20860
21. Ruiz JC, Devlin A, Kim J, Conrad NK. Kaposi's sarcoma-associated herpesvirus fine-tunes the temporal expression of late genes by manipulating a host RNA quality control pathway. *J Virol.* [published online ahead of print, 2020 May 6]. 2020;JVI.00287-20. doi:10.1128/JVI.00287-20
22. McKenzie CW, Klonoski JM, Maier T, et al. Enhanced response to pulmonary Streptococcus pneumoniae infection is associated with primary ciliary dyskinesia in mice lacking Pcdp1 and Spef2. *Cilia*. 2013;**2**(1):18. doi:10.1186/2046-2530-2-18
23. Meessen-Pinard M, Le Coupanec A, Desforges M, Talbot PJ. Pivotal Role of Receptor-Interacting Protein Kinase 1 and Mixed Lineage Kinase Domain-Like in Neuronal Cell Death Induced by the Human Neuroinvasive Coronavirus OC43. *J Virol*. 2016;91(1):e01513-16. doi:10.1128/JVI.01513-16

medRxiv preprint doi: https://doi.org/10.1101/2020.06.17.20134015; this version posted June 19, 2020. 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. It is made available under a CC-BY-NC 4.0 International license .

24. Noubade R, Wong K, Ota N, *et al*. NRROS negatively regulates reactive oxygen species during host defence and autoimmunity. *Nature*. 2014;**509**:235-239. https://doi.org/10.1038/nature13152
25. Chang YC, Su CY, Chen MH, *et al.* Secretory RAB GTPase 3C modulates IL6-STAT3 pathway to promote colon cancer metastasis and is associated with poor prognosis. *Mol Cancer*. 2017;**16**(1):135. doi:10.1186/s12943-017-0687-7
26. Wei Y, Zhang F, Zhang Y, et al. Post-transcriptional regulator Rbm47 elevates IL-10 production and promotes the immunosuppression of B cells. *Cell Mol Immunol*. 2019;**16**(6):580-589. doi:10.1038/s41423-018-0041-z
27. Lyu M, Li Y, Hao Y, et al. Elevated Semaphorin 5A correlated with Th1 polarization in patients with chronic immune thrombocytopenia. *Thromb Res*. 2015;**136**(5):859-864. doi:10.1016/j.thromres.2015.07.032
28. Barneda D, Planas-Iglesias J, Gaspar ML, *et al*. The brown adipocyte protein CIDEA promotes lipid droplet fusion via a phosphatidic acid-binding amphipathic helix. *Elife*. 2015;**4**:e07485. doi:10.7554/eLife.07485
29. Chang YC, Hee SW, Lee WJ, *et al*. Genome-wide scan for circulating vascular adhesion protein-1 levels: MACROD2 as a potential transcriptional regulator of adipogenesis. *J Diabetes Investig*. 2018;**9**(5):1067-1074. doi:10.1111/jdi.12805
30. Chen W, Chang B, Wu X, Li L, Sleeman M, Chan L. Inactivation of Plin4 downregulates Plin5 and reduces cardiac lipid accumulation in mice. *Am J Physiol Endocrinol Metab*. 2013;**304**(7):E770-E779. doi:10.1152/ajpendo.00523.2012
31. Svensson KJ, Long JZ, Jedrychowski MP, et al. A Secreted Slit2 Fragment Regulates Adipose Tissue Thermogenesis and Metabolic Function. Cell Metab. 2016;23(3):454-466. doi:10.1016/j.cmet.2016.01.008
32. Castro IG, Eisenberg-Bord M, Persiani E, Rochford JJ, Schuldiner M, Bohnert M. Promethin Is a Conserved Seipin Partner Protein. *Cells*. 2019;**8**(3):268. doi:10.3390/cells8030268
33. Bohnert M. New friends for seipin – implications of seipin partner proteins in the life cycle of lipid droplets. *Semin Cell Dev Biol*. Published online May 8th, 2020. https://doi.org/10.1016/j.semcdb.2020.04.012
34. Zheng PF, Yin RX, Deng GX, Guan YZ, Wei BL, Liu CX. Association between the XKR6 rs7819412 SNP and serum lipid levels and the risk of coronary artery disease and ischemic stroke. *BMC Cardiovasc Disord*. 2019;**19**(1):202 doi:10.1186/s12872-019-1179-z
35. Besschetnova TY, Ichimura T, Katebi N, St Croix B, Bonventre JV, Olsen BR. Regulatory mechanisms of anthrax toxin receptor 1-dependent vascular and connective tissue homeostasis. *Matrix Biol*. 2015;**42**:56-73. doi:10.1016/j.matbio.2014.12.002
36. Liu Y, Zhang W. The role of HOPX in normal tissues and tumor progression. *Biosci Rep*. 2020;**40**(1):BSR20191953. doi:10.1042/BSR20191953
37. Pallafacchina G, Zanin S, Rizzuto R. Recent advances in the molecular mechanism of mitochondrial calcium uptake. *F1000Res*. 2018;**7**:F1000 Faculty Rev-1858. doi:10.12688/f1000research.15723.1
38. Zhu Y. Zhuo J, Li C, *et al.* Regulatory network analysis of hypertension and hypotension microarray data from mouse model. *Clin Exp Hypertens*. 2017;**40**(7):631-636. https://doi.org/10.1080/10641963.2017.1416120
39. Kawakami T, Kawakami Y, Kitaura J. Protein kinase C beta (PKC beta): normal functions and diseases. *J Biochem*. 2002;**132**(5):677-682. doi:10.1093/oxfordjournals.jbchem.a003273
40. Pan L, Lin Z, Tang X, *et al.* S-Nitrosylation of Plastin-3 Exacerbates Thoracic Aortic Dissection Formation via Endothelial Barrier Dysfunction. *Arterioscler Thromb Vasc Biol*. 2020;**40**(1):175-188. doi:10.1161/ATVBAHA.119.313440
41. Schultess J, Danielewski O, Smolenski AP. Rap1GAP2 is a new GTPase-activating protein of Rap1 expressed in human platelets. *Blood*. 2005;**105**(8):3185-3192. doi:10.1182/blood-2004-09-3605
42. Neumüller O, Hoffmeister M, Babica J, *et al.* Synaptotagmin-like protein 1 interacts with the GTPase-activating protein Rap1GAP2 and regulates dense granule secretion in platelets. *Blood*. 2009;**114**(7):1396-1404. doi:10.1182/blood-2008-05-155234
43. Surendran P, Drenos F, Young R, *et al.* Trans-ancestry meta-analyses identify rare and common variants associated with blood pressure and hypertension. Nat Genet. 2016;**48**(10):1151-1161. doi:10.1038/ng.3654
44. Sadanandam A, Rosenbaugh EG, Singh S, Varney M, Singh RK. Semaphorin 5A promotes angiogenesis by increasing endothelial cell proliferation, migration, and decreasing apoptosis. *Microvasc Res*. 2010;**79**(1):1-9. doi:10.1016/j.mvr.2009.10.005
45. Sa S, Gu M, Chappell J, *et al.* Induced Pluripotent Stem Cell Model of Pulmonary Arterial Hypertension Reveals Novel Gene Expression and Patient Specificity. *Am J Respir Crit Care Med*. 2017;**195**(7):930-941. doi:10.1164/rccm.201606-1200OC
46. Malik AR, Lips J, Gorniak-Walas M, *et al*. SorCS2 facilitates release of endostatin from astrocytes and controls post-stroke angiogenesis. *Glia*. 2020;**68**(6):1304-1316. doi:10.1002/glia.23778
47. Signorelli SS, Barresi V, Musso N, *et al.* Polymorphisms of steroid 5-alpha-reductase type I (SRD5A1) gene are associated to peripheral arterial disease. *J Endocrinol Invest*. 2008;**31**(12):1092-1097. doi:10.1007/BF03345658
48. Lemieux E, Cagnol S, Beaudry K, Carrier J, Rivard N. Oncogenic KRAS signalling promotes the Wnt/β-catenin pathway through LRP6 in colorectal cancer. *Oncogene*. 2015;**34**(38):4914-4927. doi:10.1038/onc.2014.416
49. Shou Z, Luo C, Xin H, *et al.* MACROD2 deficiency promotes hepatocellular carcinoma growth and metastasis by activating GSK-3β/β-catenin signaling. *npj Genom Med*. 2020;**5**(15); https://doi.org/10.1038/s41525-020-0122-7
50. Yin X, Xiang T, Mu J, *et al.* Protocadherin 17 functions as a tumor suppressor suppressing Wnt/β-catenin signaling and cell metastasis and is frequently methylated in breast cancer. *Oncotarget*. 2016;**7**(32):51720-51732. doi:10.18632/oncotarget.10102

medRxiv preprint doi: https://doi.org/10.1101/2020.06.17.20134015; this version posted June 19, 2020. 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. It is made available under a CC-BY-NC 4.0 International license .

51. Gwak J, Cho M, Gong SJ, *et al*. Protein-kinase-C-mediated beta-catenin phosphorylation negatively regulates the Wnt/beta-catenin pathway. *J Cell Sci*. 2006;**119**(Pt 22):4702-4709. doi:10.1242/jcs.03256
52. Novellino L, De Filippo A, Deho P, *et al*. PTPRK negatively regulates transcriptional activity of wild type and mutated oncogenic beta-catenin and affects membrane distribution of beta-catenin/E-cadherin complexes in cancer cells. *Cell Signal*. 2008;**20**(5):872-883. doi:10.1016/j.cellsig.2007.12.024
53. Liu X, Peng J, Zhou Y, Xie B, Wang J. Silencing RRM2 inhibits multiple myeloma by targeting the Wnt/β-catenin signaling pathway. *Mol Med Rep*. 2019;**20**(3):2159-2166. doi:10.3892/mmr.2019.10465
54. Chen J, Yang H, Wen J, *et al*. NHE9 induces chemoradiotherapy resistance in esophageal squamous cell carcinoma by upregulating the Src/Akt/β-catenin pathway and Bcl-2 expression. Oncotarget. 2015;**6**(14):12405-12420. doi:10.18632/oncotarget.3618
55. Ng L, Chow AKM, Man JHW, et al. Suppression of Slit3 induces tumor proliferation and chemoresistance in hepatocellular carcinoma through activation of GSK3β/β-catenin pathway. BMC Cancer. 2018;18:621 https://doi.org/10.1186/s12885-018-4326-5
56. Gaulton A, Bellis LJ, Bento AP, et al. ChEMBL: a large-scale bioactivity database for drug discovery. *Nucleic Acids Res.* 2012;**40**(Database issue):D1100-D1107. doi:10.1093/nar/gkr777
57. de Wilde AH, Snijder EJ, Kikkert M, van Hemert MJ. Host Factors in Coronavirus Replication. Curr Top *Microbiol Immunol*. 2018;**419**:1-42. doi: 10.1007/82_2017_25
58. Channappanavar R, Zhao J, Perlman S. T cell-mediated immune response to respiratory coronaviruses. *Immunol Res*. 2014;**59**(1-3):118-128. doi:10.1007/s12026-014-8534-z
59. Liao M, Yuan LJ, Wen, Y *et al*. The landscape of lung broncheoalveolar immune cells in COVID-19 revealed by single-cell RNA sequencing. *medRxiv*. https://doi.org/10.1101/2020.02.23.20026690
60. Albrecht I, Niesner U, Janke M, *et al*. Persistence of effector memory Th1 cells is regulated by Hopx. *Eur J Immunol.* 2010;**40**(11):2993-3006. https://doi.org/10.1002/eji.201040936
61. Mueller C, August A. Attenuation of immunological symptoms of allergic asthma in mice lacking the tyrosine kinase ITK. *J Immunol*. 2003;**170**(10):5056-5063. https://doi.org/10.4049/jimmunol.170.10.5056
62. Lo HY. Itk inhibitors: a patent review. *Expert Opin Ther Pat* 2010;**20**(4):459-469. https://doi.org/10.1517/13543771003674409
63. Dubovsky JA, Beckwith KA, Natarajan G, *et al.* Ibrutinib is an irreversible molecular inhibitor of ITK driving a Th1-selective pressure in T lymphocytes. *Blood*. 2013;**122**(15):2539-2549. doi:10.1182/blood-2013-06-507947
64. Brien SO, Hillmen P, Coutre S, *et al*. Safety Analysis of Four Randomized Controlled Studies of Ibrutinib in Patients With Chronic Lymphocytic Leukemia/Small Lymphocytic Lymphoma or Mantle Cell Lymphoma. *Clin Lymphoma Myeloma Leuk.* 2018;**18**(10):648-657. https://doi.org/10.1016/j.clml.2018.06.016
65. Lin TA, McIntyre KW, Das J, *et al.* Selective Itk inhibitors block T-cell activation and murine lung inflammation. *Biochemistry*. 2004;**43**(34):11056-11062. doi:10.1021/bi049428r
66. Harling JD, Deakin AM, Campos S, et al. Discovery of novel irreversible inhibitors of interleukin (IL)-2-inducible tyrosine kinase (Itk) by targeting cysteine 442 in the ATP pocket. J Biol Chem. 2013;288(39):28195-28206. doi:10.1074/jbc.M113.474114
67. Wong JJM, Leong JY, Lee JH, Albani S, Yeo JG. Insights into the immuno-pathogenesis of acute respiratory distress syndrome. Ann Transl Med. 2019;7(19):504. doi:10.21037/atm.2019.09.28
68. Readinger JA, Schiralli GM, Jiang JK, *et al.* Selective targeting of ITK blocks multiple steps of HIV replication. *Proc Natl Acad Sci U S A*. 2008;**105**(18):6684-6689. doi:10.1073/pnas.0709659105
69. Meessen-Pinard M, Le Coupanec A, Desforges M, Talbot PJ. Pivotal Role of Receptor-Interacting Protein Kinase 1 and Mixed Lineage Kinase Domain-Like in Neuronal Cell Death Induced by the Human Neuroinvasive Coronavirus OC43. *J Virol*. 2016;91(1):e01513-16. doi:10.1128/JVI.01513-16
70. Liao D, Sun L, He S, Wang X, Lei Xiaoguang L. Necrosulfamide inhibits necroptosis by selectively targeting the mixed lineage kinase domain-like protein. *Med Chem Comm.* 2014;**5**(3):333-337 https://doi.org/10.1039/C3MD00278K
71. Ingraham NE, Lotfi-Emran S, Thielen BK. Immunomodulation in COVID-19. *Lancet Respir Med*. Published online May 4, 2020. https://doi.org/10.1016/S2213-2600(20)30226-5
72. Liu J, Zhang Z, Chai L, *et al*. Identification and characterisation of a unique leucine-rich repeat protein (LRRC33) that inhibits Toll-like receptor-mediated NF-κB activation. *Biochem Biophys Res Commun.* 2013;**434**(1):28-34. https://doi.org/10.1016/j.bbrc.2013.03.071
73. Janowski AM, Sutterwala FS. Editorial: dialing down inflammation: LRRC33 regulates Toll-like receptor signaling. *J Leukoc Biol*. 2014;**96**(1):4-6. doi:10.1189/jlb.3CE0214-119R
74. PubChem Identifier: GENEID 375387. URL: https://pubchem.ncbi.nlm.nih.gov/gene/375387
75. Bromberg J, Wang TC. Inflammation and cancer: IL-6 and STAT3 complete the link. *Cancer Cell.* 2009;**15**(2):79-80. doi:10.1016/j.ccr.2009.01.009

medRxiv preprint doi: https://doi.org/10.1101/2020.06.17.20134015; this version posted June 19, 2020. 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. It is made available under a CC-BY-NC 4.0 International license .

76. ClinicalTrials.gov. Treatment of SARS Caused by COVID-19 with Ruxolitinib. ClinicalTrials.gov identifier: NCT04334044. Available from: https://clinicaltrials.gov/ct2/show/NCT04334044
77. Varga Z, Flammer AJ, Steiger P, *et al*. Endothelial cell infection and endotheliitis in COVID-19. *The Lancet*. 2020, **395**: 1417-1418. https://doi.org/10.1016/ S0140-6736(20)30917-X
78. Savoia C, Schiffrin EL. Vascular inflammation in hypertension and diabetes: molecular mechanisms and therapeutic interventions. *Clin Sci (Lond)*. 2007;**112**(7):375-384. https://doi.org/10.1042/CS20060247
79. Taddei S, Virdis A, Ghiadoni L. *et al*. Endothelial dysfunction in hypertension. *J Cardiovasc Pharmacol.* 2001;**38**:S11-S14.
80. Madjid M, Safavi-Naeini P, Solomon SD. Potential effects of coronavirus on the cardiovascular system. *JAMA Cardiol.* Published online March 27, 2020. doi:10.1001/jamacardio.2020.1286
81. Thornton AM, Shevach EM. Helios: still behind the clouds. *Immunology*. 2019;**158**(3):161-170. https://doi.org/10.1111/imm.13115
82. Yang M, Liu Y, Mo B, *et al*. Helios but not CD226, TIGIT and Foxp3 is a Potential Marker for $CD4^+$ Treg Cells in Patients with Rheumatoid Arthritis. *Cell Physiol Biochem*. 2019;**52**(5):1178-1192. doi:10.33594/000000080
83. Chen ZY, Chen F, Wang YG. Down-regulation of Helios Expression in Tregs from Patients with Hypertension. *Curr Med Sci.* 2018;**38**:58-63. https://doi.org/10.1007/s11596-018-1846-9
84. Kassan M, Wecker A, Kadowitz P, Trebak M, Matrougui K. CD4+CD25+Foxp3 regulatory T cells and vascular dysfunction in hypertension. *J Hypertens*. 2013;**31**(10):1939-1943. doi:10.1097/HJH.0b013e328362feb7
85. Lambert JP, Luongo TS, Tomar D, *et al*. MCUB Regulates the Molecular Composition of the Mitochondrial Calcium Uniporter Channel to Limit Mitochondrial Calcium Overload During Stress. *Circulation.* 2019;**140**(21):1720-1733. doi:10.1161/CIRCULATIONAHA.118.037968
86. Huo J, Lu S, Kwong JQ *et al.* MCUb Induction Protects the Heart from Post-Ischemic Remodelling. *Circ Res.* 2020;0. https://doi.org/10.1161/CIRCRESAHA.119.316369
87. Sun JK, Zhang WH, Zou L. Serum calcium as a biomarker of clinical severity and prognosis in patients with coronavirus disease 2019: a retrospective cross-sectional study. 2020;0. *Critical Care & Emergency Medicine*. doi: 10.21203/rs.3.rs-17575/v1
88. Osto E, Kouroedov A, Mocharla P *et al*. Inhibition of Protein Kinase Cβ Prevents Foam Cell Formation by Reducing Scavenger Receptor A Expression in Human Macrophages. *Circulation*. 2008;**118**:2174-2182. https://doi.org/10.1161/CIRCULATIONAHA.108.789537
89. Cotter MA, Jack AM, Cameron NE. Effects of the protein kinase C beta inhibitor LY333531 on neural and vascular function in rats with streptozotocin-induced diabetes. *Clin Sci (Lond)*. 2002;**103**(3):311-321. doi:10.1042/cs1030311
90. Chakrabarty B, Das D, Bulusu G, Roy A. *ChemRxiv.* Published online 20th March, 2020. Available from: https://chemrxiv.org/articles/Network-Based_Analysis_of_Fatal_Co-morbidities_of_COVID-19_and_Potential_Therapeutics/12136470/1
91. Sheak JR, Yan S, Weise-Cross L, *et al*. PKCβ and reactive oxygen species mediate enhanced pulmonary vasoconstrictor reactivity following chronic hypoxia in neonatal rats. *Am J Physiol Heart Circ Physiol*. 2020;**318**(2):H470-H483. doi:10.1152/ajpheart.00629.2019
92. Ma HT, Lin WW, Zhao B, *et al.* Protein kinase C beta and delta isoenzymes mediate cholesterol accumulation in PMA-activated macrophages. *Biochem Biophys Res Commun*. 2006;**349**(1):214-220. doi:10.1016/j.bbrc.2006.08.018
93. Ruboxistaurin: LY 333531. *Drugs R D*. 2007;8(3):193-199. doi:10.2165/00126839-200708030-00007
94. Aiello LP, Clermont A, Arora V, *et al.* Inhibition of PKC beta by oral administration of ruboxistaurin is well tolerated and ameliorates diabetes-induced retinal hemodynamic abnormalities in patients. *Invest Ophthalmol Vis Sci*. 2006;**47**(1):86-92. doi:10.1167/iovs.05-0757
95. McGill JB, King GL, Berg PH, *et al.* Clinical safety of the selective PKC-beta inhibitor, ruboxistaurin. *Expert Opin Drug Saf.* 2006;**5**(6):835-845. doi:10.1517/14740338.5.6.835
96. PharmaTimes. Eli Lilly and Takeda pull the plug on ruboxistaurin studies (2007). Available from: http://www.pharmatimes.com/news/eli_lilly_and_takeda_pull_the_plug_on_ruboxistaurin_studies_990757
97. Ormazabal V, Nair S, Elfeky O, Aguayo C, Salomon C, Zuñiga FA. Association between insulin resistance and the development of cardiovascular disease. *Cardiovasc Diabetol.* 2018;**17**(1):122. Published 2018 Aug 31. doi:10.1186/s12933-018-0762-4
98. Puri V, Ranjit S, Konda S, *et al.* Cidea is associated with lipid droplets and insulin sensitivity in humans. *Proc Natl Acad Sci U S A*. 2008;**105**(22):7833-7838. doi:10.1073/pnas.0802063105
99. Kim M, Kim DJ. GFRA1: A Novel Molecular Target for the Prevention of Osteosarcoma Chemoresistance. *Int J Mol Sci.* 2018;**19**(4):1078. Published 2018 Apr 4. doi:10.3390/ijms19041078
100. Xing W, Yang L, Peng Y, et al. Ginsenoside Rg3 attenuates sepsis-induced injury and mitochondrial dysfunction in liver via AMPK-mediated autophagy flux. *Biosci Rep.* 2017;**37**(4):BSR20170934. Published 2017 Aug 30. doi:10.1042/BSR20170934

medRxiv preprint doi: https://doi.org/10.1101/2020.06.17.20134015; this version posted June 19, 2020. 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. It is made available under a CC-BY-NC 4.0 International license .

101. Yin X, Xin H, Mao S, Wu G, Guo L. The Role of Autophagy in Sepsis: Protection and Injury to Organs. *Front Physiol*. 2019;**10**:1071. doi:10.3389/fphys.2019.01071
102. Ivanova, L, Tamiiku-Taul J, Sidrova Y, *et al.* Small-molecule ligands as potential GDNF family receptor agonists. ACS Omega. 2018;**3**:1022-1030. DOI: 10.1021/acsomega.7b01932
103. Zhang J, Lan Y, Sanyal S. Modulation of Lipid Droplet Metabolism-A Potential Target for Therapeutic Intervention in *Flaviviridae* Infections. *Front Microbiol*. 2017;**8**:2286. doi:10.3389/fmicb.2017.02286
104. Camus G, Vogt DA, Kondratowicz AS, Ott M. Chapter 10 – Lipid Droplets and Viral Infections. *Methods Cell Biol.* 2013;**116**:167-190. https://doi.org/10.1016/B978-0-12-408051-5.00009-7
105. Müller C, Hardt M, Schwudke D, Neuman BW, Pleschka S, Ziebuhr J. Inhibition of Cytosolic Phospholipase $A_2\alpha$ Impairs an Early Step of Coronavirus Replication in Cell Culture. *J Virol*. 2018;**92**(4):e01463-17. doi:10.1128/JVI.01463-17.
106. Zhang J, Lan Y, Sanyal S. Modulation of Lipid Droplet Metabolism-A Potential Target for Therapeutic Intervention in *Flaviviridae* Infections. *Front Microbiol*. 2017;**8**:2286. doi:10.3389/fmicb.2017.02286
107. Ploen D, Hafirassou ML, Himmelsbach K *et al*. TIP47 plays a crucial role in the life cycle of hepatitis C virus. *J Hepatol*. 2013;**58**(6):1081-1088. DOI:https://doi.org/10.1016/j.jhep.2013.01.022
108. Clément S, Fauvelle C, Branche E. Role of seipin in lipid droplet morphology and hepatitis C virus life cycle. *J Gen Virol*. 2013;**94**(10):2208-2214. https://doi.org/10.1099/vir.0.054593-0
109. Ma B, Hottiger MO. Crosstalk between Wnt/β-Catenin and NF-κB Signaling Pathway during Inflammation. *Front Immunol*. 2016;**7**:378. doi:10.3389/fimmu.2016.00378
110. Duan Y, Liao AP, Kuppireddi S, Ye Z, Ciancio MJ, Sun J. beta-Catenin activity negatively regulates bacteria-induced inflammation. Lab Invest. 2007;87(6):613-624. doi:10.1038/labinvest.3700545
111. Ma B, van Blitterswijk CA, Karperien M. A Wnt/β-catenin negative feedback loop inhibits interleukin-1-induced matrix metalloproteinase expression in human articular chondrocytes. Arthritis Rheum. 2012;64(8):2589-2600. doi:10.1002/art.34425
112. Ljungberg JK, Kling JC, Tran TT, Blumenthal A. Functions of the WNT Signaling Network in Shaping Host Responses to Infection. *Front Immunol.* 2019;**10**:2521 https://doi.org/10.3389/fimmu.2019.02521
113. Ma B, Hottiger MO. Crosstalk between Wnt/β-Catenin and NF-κB Signaling Pathway during Inflammation. *Front Immunol*. 2016;7:378. Published 2016 Sep 22. doi:10.3389/fimmu.2016.00378
114. Chen X, Aqrawi LA, Utheim TP *et al.* Elevated cytokine levels in tears and saliva of patients with primary Sjögren's syndrome correlate with clinical ocular and oral manifestations. *Sci Rep.*2019;**9**:7319. https://doi.org/10.1038/s41598-019-43714-5
115. More S, Yang X, Zhu Z, *et al.* Regulation of influenza virus replication by Wnt/β-catenin signaling. *PLoS One*. 2018;**13**(1):e0191010. doi:10.1371/journal.pone.0191010
116. Cawood R, Harrison SM, Dove BK, Reed ML, Hiscox JA. Cell cycle dependent nucleolar localization of the coronavirus nucleocapsid protein. *Cell Cycle*. 2007;**6**(7):863-867. doi:10.4161/cc.6.7.4032
117. Wu CJ, Jan JT, Chen CM, *et al*. Inhibition of Severe Acute Respiratory Syndrome Coronavirus Replication by Niclosamide. *Antimicrob Agents Chemother.* 2004;**48**(7):2693-2696. doi: 10.1128/AAC.48.7.2693–2696.2004
118. Ono M, Yin P, Navarro A *et al.* Inhibition of canonical WNT signaling attenuates human leiomyoma cell growth. *Fertil Steril.* 2014;**101**(5):1441-1449. https://doi.org/10.1016/j.fertnstert.2014.01.017
119. Kuo CL, Pilling LC, Atkins JL, *et al.* APOE e4 genotype predicts severe COVID-19 in the UK Biobank community cohort. *J Gerontol A.* https://doi.org/10.1093/gerona/glaa131

medRxiv preprint doi: https://doi.org/10.1101/2020.06.17.20134015; this version posted June 19, 2020. 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. It is made available under a CC-BY-NC 4.0 International license .

Appendix

SEPSIS CONTROL CRITERIA

1. Controls to exclude any patients with the following ICD-10 codes:

ICD-10	Disease
A02.1	Salmonella septicemia
A22.7	Anthrax septicemia
A40.x	Streptococcal septicemia
A41.x	Other septicemia
B37.7	Candidal septicemia
O35	Puerperal sepsis
R57.2	Septic shock

2. Controls to include at least one of the following ICD-10 codes:

ICD-10	Disease
J12-J18	Pneumonia
J20-J22	Lower respiratory infection
B95	Streptococcus
B95	Staphylococcus

3. Controls to include least one of the following ICD-10 codes:

ICD-10	Disease
E10-E14	Diabetes
N00-N19	Kidney Disease
K70-K77	Liver Disease
J40-J44	COPD

Table 8: Pairwise comparison of co-morbidities prevalent in severe COVID-19 patient population

	All Severe COVID-19	Cardiovascular Disease	Hypertension	Respiratory Disease	Diabetes
All Severe COVID-19	779				
Cardiovascular Disease	136				
Hypertension	385	115			
Respiratory Disease	170	39	105		
Diabetes	143	55	117	38	
Alzheimer's Disease	10	1	5	1	2

IDENTIFICATION OF CASES WITH SPECIFIC CO-MORBIDITIES

Cases with co-morbidities were identified using the following disease codes:

Disease	ICD-10	Self-Reported codes	Operation codes
Cardiovascular disease	I20-I25	1070, 1075, 1095	K40-41, K45, K49, K50.2, K75
Hypertension	I10, I11, R03	1065	
Diabetes	E10-14	1220, 1222, 1223	
Chronic respiratory diseases	J40-47	1111, 1112	
Alzheimer's disease	G30		

medRxiv preprint doi: https://doi.org/10.1101/2020.06.17.20134015; this version posted June 19, 2020. 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. It is made available under a CC-BY-NC 4.0 International license .

Table 9: Table of severe COVID-19-associated genes with chromosomal location and RF score (genes in bold have been validated in further studies)

GENE	CHROMOSOME	RF SCORE
RPL7AP57	11	7.80
TMEM159	16	2.41
CDC7	1	1.83
REX1BD	19	1.26
PSMC1	14	1.16
STH	17	1.10
KRAS	12	0.95
CIDEA	18	0.82
PLS3	X	0.74
COQ6	14	0.72
PLIN4	19	0.67
HOPX	4	0.67
PCDH17	13	0.67
SRD5A1	5	0.61
MLKL	16	0.57
TBC1D2	9	0.53
MEP1B	18	0.49
B3GLCT	13	0.48
CRLF1	19	0.46
MCUB	4	0.45
SAMD3	6	0.44
ENTPD5	14	0.42
LINC02210-CRHR1	17	0.39
L3MBTL3	6	0.36
PEX14	1	0.36
SPEF2	5	0.35
SLC16A10	6	0.35
LINC02210	17	0.34
ERICH1	8	0.29
MAPT	17	0.28
SMCHD1	18	0.26
RRM2	2	0.26
NRDE2	14	0.25
ANTXR1	2	0.25
STAC	3	0.23
ITK	5	0.22
ZNF106	15	0.21
DNAH2	17	0.20
CTR9	11	0.19
NRROS	3	0.19
RBM47	4	0.18
RAB3C	5	0.17
GFRA1	10	0.15
KANSL1	17	0.14
C9orf92	9	0.14
IKZF2	2	0.13
XKR6	8	0.11
EIF3E	8	0.11
MAST4	5	0.10
PIGX	3	0.10
GAREM1	18	0.10
IQCM	4	0.09
RAP1GAP2	17	0.09
PRKCB	16	0.09
AACSP1	5	0.09
ATXN1	6	0.08
VTI1A	10	0.08
C4orf50	4	0.07
SEMA5A	5	0.07
SLC9A9	3	0.07
SPPL2C	17	0.07

medRxiv preprint doi: https://doi.org/10.1101/2020.06.17.20134015; this version posted June 19, 2020. 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. It is made available under a CC-BY-NC 4.0 International license .

PTPRK	6	0.06
ATRNL1	10	0.06
LHFPL2	5	0.05
PRUNE2	9	0.05
CNTN5	11	0.05
SLIT3	5	0.05
SORCS2	4	0.04
KCNIP4	4	0.03
MACROD2	20	0.02
CNTNAP2	7	0.02

Table 10. Table showing drug repurposing candidates for 10 target genes identified as being COVID-19 related

GENE	COMPOUND CHEMBL ID	DRUG	PHASE	MOLECULE TYPE
CDC7	CHEMBL3544943	bms-863233	2	Small molecule
CDC7	CHEMBL3545090	rxdx-103	1	Small molecule
CDC7	CHEMBL3545321	nms-1116354	1	Small molecule
PSMC1	CHEMBL325041	bortezomib	4	Small molecule
PSMC1	CHEMBL451887	carfilzomib	4	Protein
PSMC1	CHEMBL2103884	oprozomib	1	Small molecule
PSMC1	CHEMBL3545432	ixazomib citrate	4	Small molecule
SRD5A1	CHEMBL254328	abiraterone	4	Small molecule
SRD5A1	CHEMBL1200969	dutasteride	4	Small molecule
RRM2	CHEMBL467	hydroxyurea	4	Small molecule
RRM2	CHEMBL1637	gemcitabine hydrochloride	4	Small molecule
RRM2	CHEMBL1750	clofarabine	4	Small molecule
RRM2	CHEMBL1096882	fludarabine phosphate	4	Small molecule
RRM2	CHEMBL1200983	gallium nitrate	4	Small molecule
RRM2	CHEMBL3544910	motexafin gadolinium	3	Small molecule
RRM2	CHEMBL3989496	tezacitabine	2	Small molecule
ITK	CHEMBL1201733	pazopanib hydrochloride	4	Small molecule
ITK	CHEMBL4085457	pf-06651600	3	Small molecule
ITK	CHEMBL1873475	ibrutinib	4	Small molecule
CIDEA	CHEMBL121	rosiglitazone	4	Small molecule
PLIN4	CHEMBL121	rosiglitazone	4	Small molecule
MLKL	CHEMBL3220918	necrosulfanomide	preclinical	Small molecule
GFRA1	CHEMBL2108380	liatermin	1	Small molecule
PRKCB	CHEMBL300138	enzastaurin	3	Small molecule
PRKCB	CHEMBL91829	ruboxistaurin	3	Small molecule
PRKCB	CHEMBL494089	gsk-690693	1	Small molecule
PRKCB	CHEMBL574737	ucn-01	2	Small molecule
PRKCB	CHEMBL565612	sotrastaurin	2	Small molecule
PRKCB	CHEMBL608533	midostaurin	4	Small molecule
PRKCB	CHEMBL3545332	cep-2563	1	Small molecule

Table 11. Table showing 8 target genes identified as being COVID-19 related that have active compounds in ChEMBL[56]

GENE	NO. OF ACTIVE COMPOUNDS (ChEMBL)
MACROD2	3
MAST4	6
MLKL	11
SLC16A10	70
MEP1B	90
KRAS	120
L3MBTL3	122
MAPT	9,589

medRxiv preprint doi: https://doi.org/10.1101/2020.06.17.20134015; this version posted June 19, 2020. 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. It is made available under a CC-BY-NC 4.0 International license .

Table 12. Table showing all the UK Biobank COVID-19 datasets and numbers of cases:controls (post-QC) used in this study

UK Biobank Data Release	18 May 2020	26 May 2020	6 June 2020
Cases	779 (442 males, 337 females)	877 (492 males, 385 females)	929 (524 males, 405 females)
Controls	1,553	5,438	5,563