

Genes associated with liver damage signalling pathways may impact the severity of COVID-19 symptoms in Spanish and Italian populations

Leire Moya¹, Samaneh Farashi^{1,2}, Prashanth N Suravajhala³, Panchadsaram Janaththani^{1,2}, Jyotsna Batra^{1,2*}

¹Cancer Program, Institute of Health and Biomedical Innovation and School of Biomedical Sciences, Queensland University of Technology, Brisbane, Queensland, Australia

²Australian Prostate Cancer Research Centre, Brisbane, Queensland, Australia

³Department of Biotechnology and Bioinformatics, Birla Institute of Scientific Research, Jaipur, India

*Correspondence should be addressed to:

A/Prof Jyotsna Batra, Australian Prostate Cancer Centre- Queensland, Translational Research Institute, Queensland University of Technology, Woolloongabba, Queensland, Australia 4102. Phone: +61 7 3443 7336, Fax: +61 7 3138 6030, email: jyotsna.batra@qut.edu.au.

Keywords: COVID-19, SARS-CoV-2, coronavirus, genetic variations, population studies, hepatic fibrosis and hepatic cell stellate pathway

Abstract

Aim: The novel SARS-CoV-2 virus, which causes the COVID-19 disease, has infected more than 10 million people and caused 500K deaths worldwide. In Europe, over 2 million confirmed cases have been reported, while nearly 200K people have died from the disease. Despite strict containment measures in Spain and Italy after the first reported COVID-19 patient, these two countries have remained in the top five European nations with the highest mortality rate for over two months. We hypothesised that a genetic mechanism could partially explain the poor survival outcome observed in these two countries.

Methods: An extensive literature search to identify human candidate genes linked to SARS-CoV infection, host immune evasion and disease aggressiveness was carried out. Pathway analysis (IPA) was performed to select the most significantly associated canonical signalling pathways with the genes of interest. The genetic variants' at these genes with $\pm 1\text{Mb}$ flanking region was extracted (GRCh37/hg19 built). Over 80 million single nucleotide polymorphisms (SNPs) were analysed in genome-wide data of 2,504 individuals (1000 genomes, phase III, <https://www.internationalgenome.org/>). Principal component (PC) analysis was performed, ancestry by the whole genome was inferred and subsets of the regions of interest were extracted (PLINK v1.9b, <http://pngu.mgh.harvard.edu/purcell/plink/>). PC1 to PC20 values from five European ancestries, including the Spanish and Italian populations, were used for PC analysis. Gene function predictions were run with our genes of interest as a query to the GeneMANIA Cytoscape plugin (<https://genemania.org/>).

Results: A total of 437 candidate genes associated with SARS were identified, including 21 correlated with COVID-19 aggressiveness. The two most significant pathways associated with all 437 genes (*Caveolar-mediated Endocytosis* and *MSP-RON Signalling*) did not show any segregation at the population level. However, the most significant canonical pathway associated with genes linked to COVID-19 aggressiveness, the *Hepatic Fibrosis and Hepatic Stellate Cell Activation*, showed population-specific segregation. Both the Spanish and Italian populations clustered together from the rest of Europe. This was also observed for the Finnish population but in the opposite direction. These results suggest some of the severe COVID-19 cases reported in Spain and Italy could be partially explained by a pre-existing liver condition (especially liver cancer) and/or may lead to further COVID-19 related liver complications.

Introduction

The novel coronavirus disease 19 (COVID-19), is an infectious disease caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). It was firstly identified in December 2019 [1] and since then, nearly 500K people have died worldwide global deaths (27th June 2020 [2]). SARS-CoV-2 is closely related to SARS-CoV-1, which was firstly identified in November 2002 in China [3]. Despite the virus spreading to 26 countries [4], only five localised hotspots worldwide remained [5], and by July 2003 the virus outbreak was declared officially contained by the World Health Organisation (WHO) [6]. A total of 774 people died [3] and no new cases have been reported since 2004 [7]. On that year, R_0 data from the three main hotspots (Singapore, Toronto and Hong Kong) showed SARS-CoV-1 had an $R_0 = 0.9 - 1.8$ [8]. The R_0 value is and indicative of how an epidemic can spread in a completely susceptible population where no specific control measures are implemented. The higher the value, the more difficult is for the population to contain the spread. Data from February and April this year, report a mean value of $R_0=2.6$ [9, 10] for SARS-CoV-2. This data explains the broad success of the containment measures implemented during the first SARS pandemic [4], where the localisation of the virus lead to its eradication.

At the end of February 2020 (29th), Italy recorded the highest number of community transmitted COVID-19 related deaths (21 total deaths [11]) in Europe. Two weeks later, Spain was the second European county with the highest COVID-19 related deaths (136 deaths, [11]) and less than a month later, both countries had 29.3 and 28.3 deaths per 100,000 respectively, while the rest of Europe reported 0.45 – 11.3 deaths (7th of April, [11]). Currently (4th of June 2020), Belgium and the UK have reported the highest COVID-19 mortality rate while Spain and Italy remain the third and fourth countries [11]. The reported COVID-19 mortality rates are most likely the consequences of a combination of factors such as the Public Health's response, social behaviour and population comorbidities.

As such, we searched what the differences in terms of the responses to the crisis were for these four countries. Both in Spain and Italy, strong containment measures were implemented four and three weeks after the first confirmed case in their countries [12]. Belgium started its most strict rules five weeks after the first case while it took six weeks for the UK to take action [12]. Similarly, several COVID-19 comorbidities have been identified such as aging, diabetes, heart and liver diseases [13-16]. But little is known about which genes and/or pathways are associated with the increased risk of developing severe symptoms.

At the peak of the pandemic in Italy and Spain, we hypothesised that a genetic component in these populations may contribute to the high mortality rate of COVID-19 patients observed. To study this, we performed pathway analysis of 437 candidate genes associated with SARS-like viral intake, immune evasion and clinical phenotype, which were shortlisted from the extensive literature study. Principal Component (PC) analysis of the single nucleotide polymorphisms (SNPs) and their allele frequencies in these candidate gene loci with a 1Mb window was performed. Our analysis revealed both Italian and Spanish populations segregate differently in the pathways involved in liver injury and damage, with a potential role of the hepatic disease being involved in a more aggressive response to the COVID-19 disease.

Material and methods

Mortality rate frequency calculation

The *cause-specific death rate* was measured using the below formula provided by the Centers for Disease Control and Prevention [17] and the data is presented per 100,000 individuals. The total *COVID-19 associated deaths* were defined by the International Guidelines defined by WHO (based on ICD) [18] and the population size was sourced from the Eurostat Data Browser [19] as of 1st of January 2020.

$$\frac{\text{Number of COVID-19 associated deaths during start pandemic}}{\text{Size of the population among which the deaths occurred}} \times 10^5$$

Public health data from the European populations studied here was searched regarding the most common COVID-19 comorbidities reported. Aging has been broadly identified as the main one [13, 20, 21], followed by cardiovascular disease (CVD), diabetes, chronic obstructive pulmonary disease (COPD) and cancer history [13, 15]. Liver dysfunction, smoking status, chronic kidney diseases and immunodeficiency have also been reported [14-16].

Identification of candidate genes for analysis

PubMed was accessed between the 31st of March and 25th of May 2020. The aim was to consolidate all empirically and predicted human genes reported to have a role during SARS-CoV-2 infection to date. This included those human genes that assisted with the viral entry, evasion of the host's immune system and the SARS-CoV-2 – human interactome (Figure 1). Those studies which results did not include human genes interacting direct or indirectly with SARS-CoV-1 or SARS-CoV-2 or being affected by the disease they cause were excluded.

After completing the literature research, all the genes identified were curated for the subsequent pathway and PC analyses. For example, when the literature reference did not specify the gene isoform (e.g. NF- κ B) all those available (e.g. NFKB1 and NFKB2) were included in the analysis. Pathway signalling analysis was performed with IPA software (QIAGEN Inc., [22]) as follows. All candidate genes' symbols, or subtracted genes as described below, were entered for the Core Analysis Expression. As a reference, the *Ingenuity Knowledge Base* was applied and both *Direct and Indirect Analysis Analyses* were selected. The genes involved in the two most significant canonical pathways affected by our candidate genes' list input were selected and their locations in the genome were identified after accessing the *Table Browser* tool from UCSC (<https://genome.ucsc.edu/index.html>) (GRCh37/hg19 built). Duplicates were removed while unique transcript variants for a gene were included. The flanking regions of \pm 1Mb were added and the allele frequencies in these loci were analysed for further PC analysis amongst the populations of interest.

Also, selected genes associated with COVID-19 severe symptoms were entered in the GeneMANIA Cytoscape plugin to predict physical protein-protein interactions [23]. This plugin includes a large dataset of over 800 functional associated networks for human amongst other organisms. The data utilised come from reported studies and public large databases such as BIOGRID (Breitkreutz et al., 2008), GEO (Barrett et al., 2009), I2D (Brown and Jurisica, 2005) and Pathway Commons (<http://www.pathwaycommons.org>). The regular and automated update of the networks makes this tool and up to date resource for the category analysed.

PC Analysis

The genome-wide germline data of 2,506 individuals were obtained from 1000 genomes phase III v.5b project, the latest release of the data (May 2013) (<https://www.internationalgenome.org/>). Starting from the genotype file vcf.gz formats provided by 1000 genomes projects, we performed standard quality control methods using PLINK v1.9b (<http://pngu.mgh.harvard.edu/purcell/plink/>) [24] by removing individuals with more than 3% missing genotypes, SNPs with a call rate <97%.

We performed PC analysis for the samples to account and adjust for population stratification. PCs were computed using PLINK software v1.9b for all samples of 1000G (1KG) project, phase 3 version 5b, to infer the ancestry of the samples based on the whole genome data. We used PLINK [24] to extract subsets of regions for 1Mb window of genes of our interest. PC1 and PC20 values >6 standard deviations from five European ancestries: Great Britain (England and Scotland), CEU (Utah Residents with Northern and Western European Ancestry), Tuscany (Italy), Iberian (Spanish population) and Finland participants were used in the analysis.

Results

Mortality rate and common COVID-19 comorbidities data in European populations

Calculations of the mortality rates reported in Europe were done for the 27 countries that form the European Union [25] as shown in the Methods section. This was first calculated at the beginning of the study (7th of April) and updated later to reflect any changes (4th of June Table 1). Despite the recent separation of the UK from the EU, we included it for its relevancy towards this study. Since April, Spain, Italy and Belgium have remained in the top five countries with the highest mortality rate (Table 1).

Our search from public health databases showed both Italy and Spain countries rank similarly than other European countries for many of the most common COVID-19 comorbidities. The Italian total population is the world's second oldest nation while Spain is the 20th, with other 16 European countries ahead of it [26], suggesting ageing alone would not explain the similarities observed in these countries. Similarly, the burden of CVD is comparable between Northern Europe (384.3 deaths per 100,000 population) and Spain/Italy (average of 342.5 deaths per 100,000 population [27]). We next check the European diabetes status and today, both type I and type II are higher in North Europe than South Europe [28, 29]. The reported smoking population in Italy and Spain (18.9% and 31.8% respectively) is again similar to other European countries such as Ireland (19.6%), Germany (23.7%) or Portugal (36.8%) [30]. Lastly, data regarding several contributors that lead to liver disease-related deaths from 31 European countries showed the main cause to die from liver disease in Spain and Italy is liver cancer, with the highest hepatocellular carcinoma incidence being observed in the later [31]. Finland's main cause of liver-disease-related death is alcohol consumption and this could be relevant for the findings explained below. In the same report, it is shown that the prevalence of cirrhosis and other liver disease's aetiology in Italy is largely due to Hepatitis B and C infections with the highest incidence of hepatitis infections in Europe (5.9%).

Candidate genes identified and data curation

An extensive literature search was carried out and 291 studies were identified after entering the search terms shown in Figure 1. After removing duplicates, 95 abstracts were screened and a total of 45 publications were included in further pathway analysis due to falling in the inclusion criteria as described previously.

Overall, we identified 63 genes to have a role during a SARS-CoV infection (Supplementary Table 3), with 21 of those being associated with disease aggressiveness and patients' survival (Supplementary Table 4). In addition, 332 genes have recently been identified in an exclusive study of SARS-CoV-2 – human interactome [32]. These 405 genes were curated as described previously and a total of 437 genes were included in the pathway analysis as a group and as a subgroup, based on the severity of

the disease. Empirically and predicted human pathways affected during a SARS infection reported in the literature were identified as well (Supplementary Table 2).

Some of the genes identified, act as ligands or proteases of the viral Spike (S) receptor, assisting in the viral host recognition and entry of the viral genome into the host cell. Genes in this category reported in at least two independent studies have been included in Table 2, while those only identified in one study are in Supplementary Table 1. To mention some, the human transmembrane receptor angiotensin converting enzyme II (ACE2) has been widely reported to be a high-affinity receptor of the SARS-CoV S protein [33-36]. Interestingly, SARS-CoV-2 S protein has shown to have a higher binding affinity to this receptor than SARS-CoV-1 [37]. Several additional human transmembrane receptors and co-activators have also been identified *in vitro* and *in silico* [35, 38-50], indicating SARS-CoV-2 viral RNA entry into the host cell can occur through a different molecular mechanism.

Additionally, several candidate human genes were associated with evasion and tuning of the innate and acquired immune system *in vitro* and are summarised on Table 3 and Supplementary Table 1. The genes categorised in this group in at least two independent studies have been included in Table 3, while those only identified in one are on Supplementary Table 1. The main mechanisms of SARS-like virus to escape immune detection and/or processing were by interfering with the hosts' protein translation [51, 52] (including those associated with antiviral response) and hijacking the immune response [53-64]. They also sequester different host proteins to make precise rearrangements of the endothelium reticulum membrane [34, 63] to build the replication and transcription complex (RTC), effectively creating a protective barrier to replicate, assemble and release viral particles [36, 65, 66]. The most common immune signalling pathways affected were those regulated by IFN and NF- κ B [55-58].

Candidate genes associated with pathogenicity

Clinical data from COVID-19 patients' blood samples have been consolidated in several reviews and retrospective studies, reporting an aberrant chemokine balance in COVID-19 patients (Table 4). While *in vitro* assays suggest this dysregulated production might occur as soon as the S viral protein binds to the host membrane cell receptor [33] there is a correlation between this "cytokine storm" and the severity of the disease [54, 67, 68]. Recent clinical studies report lymphopenia, which it is diagnosed when blood lymphocyte counts are lower than the threshold, to occur early in COVID-19 patients that develop a more severe form of the disease, suggesting its potential application as an early prognostic biomarker [68-72]. Those genes associated with disease aggressiveness too were analysed independently from the rest in the pathway analysis. Additionally, Protein-Protein Interaction map of

all the 21 candidate genes associated with COVID-19 aggressiveness, highlighted the physical interactions with lactate dehydrogenase A (LDHA) and LDHB were vital players.

Signalling pathway analysis

After running the IPA analysis with all the 437 candidate genes, the two most significant canonical signalling pathways were *Caveolar-mediated Endocytosis* ($P = 2 \times 10^{-19}$) and *MSP-RON Signalling pathways* ($P = 6.1 \times 10^{-19}$). *Caveolar-mediated Endocytosis* signalling pathway controls different cellular processes such as endocytosis, cellular signalling and lipid recycling [73]. It regulates the internalisation of different particles, including virus and bacteria, after generating flask-shaped invaginations of the plasma membrane. The *MSP-RON Signalling* pathway contributes to the macrophage induced immune response to assist the host in the viral recognition via the Macrophage Stimulating Protein (MSP) and the transmembrane receptor kinase RON Protein Tyrosine Kinase/Receptor d [74].

On the other hand, when running only those genes associated with COVID-19 survival, the two most significant signalling pathways were *Hepatic Fibrosis and Hepatic Stellate Cell (HSC) Activation* ($P = 2.5 \times 10^{-18}$) and the *Communication between Innate and Adaptive Immune Cells* ($P = 1.6.1 \times 10^{-14}$). The *Hepatic Fibrosis / HSC Activation* occurs in patients with chronic liver damage after accumulating an excess of extracellular matrix (ECM) proteins [75]. This process initiates a series of pro-inflammatory events which downstream effect is the constitutive activation of HSC. When activated, HSCs cells secrete cytokines and their constant activation creates another source of extra ECM, which leaves patients developing tissue fibrosis and liver cirrhosis [75]. Finally, the *Communication between Innate and Adaptive Immune Cells* is the process in which both the immune and adaptive responses interact with each other to defend the host from infection. The initial innate response is crucial as it will determine how strong and specific the adaptive response will be [76].

The two full reports are presented in Supplementary Table 5 and Supplementary Table 6 respectively. The genes associated in these four signalling pathways and their flanking regions can be found in Supplementary Table 7, 8, 9 and 10 and they were further studied in the subsequent PC analysis.

PC analysis

A total of 81,271,745 variants and 2,504 people passed filters and QC and were therefore included to account and adjust for population stratification for the regions of interest within the genome. The PC analysis results are depicted in Figure 3. Out of four analyses, only the *Hepatic Fibrosis and HSC Activation* (Figure 3c) represented a genetic differentiation for some of the populations tested, with the Spanish and Italian populations clustering closely together and the Finnish population segregating

independently in this pathway. The specificity of the Finnish segregation may suggest a protective role of this pathway in this population (5.8 deaths per 100,000 as per 4th of June [11]) or could be a reflection of its well-known unique genetic background [77, 78].

Discussion

We aimed to identify possible genetic variations associated with SARS-CoV viruses' infections, where allele frequencies are different in the Spanish and Italian populations when compared to other European countries. Despite the strict imposed social restrictions [12] in these two countries, high-mortality rates were maintained for many weeks in the aftermath [11]. Our search from available Public Health data, confirmed the most common COVID-19 comorbidities [13-16, 20, 21] did not fully explain the high mortality rates reported in these two countries, thus we hypothesised that a shared genetic set of traits could partially explain the poor survival outcome. Indeed, genetic studies have shown that, although both populations are closely related to other European populations, they share a higher number of genomic loci [79] while both present certain grades of genomic uniqueness like the diverse haplotypic structure in the Spanish population [80] and the genetic variation inside the Italian population when comparing regions in the country [81]. Also, there are already some well-known genetic variations, in the Italian and Spanish populations associated with disease and disorders. For instance, SNPs in the alcohol dehydrogenase 4 (ADH4) gene are known to play a role in alcohol dependence in the Italian population [82]. Similarly, a Spanish-specific SNP in the ADH1B (rs1229984) and ADH6 (rs3857224) genes are associated with a heavy intake of alcohol amongst this population [83]. In this work, we performed an extensive literature search to identify candidate genes associated with SARS-like viruses' infection, evasion and hijacking of the immune responses and disease aggressiveness as genetic variations in any of these genes and potential regulatory regions around could have an effect of the pathogeny *in vivo*. For instance, Chan et al. identified a polymorphism in the gene coding for the extracellular region of the transmembrane receptor L-SIGN (liver/lymph node-specific intracellular adhesion molecules-3 grabbing non-integrin, encoded by *CLEC4M*) that binds to S viral protein to be associated with a lower risk of infections in SARS patients [45]. Similarly, the furin-like cleavage site-specific to the SARS-CoV-2 S protein has been associated with higher pathogenicity *in vitro* [68, 84]. Several host transmembrane receptors have been identified to bind to the S protein [33-36, 41-46, 50, 85] or to prime it [34, 35, 38, 86], with ACE2 and the transmembrane protease serine 2 (TMPRSS2) the two most commonly reported S protein binding receptor and activator respectively. This protease expresses in a larger set of tissues [87] than ACE2 [87-89], suggesting it is the latter which narrows down the infectious cell tropism of the virus. As a note, co-expression studies of ACE2 and TMPRSS2 found it is highest in the nasal epithelium, supporting the observation of a higher viral

load detected in nasal than throat swabs in COVID-19 patients and easy droplets viral transmission which could in part explain the high infectious rate [87]. After the S protein is activated, the viral genome enters the host. There are reports of the less well-known membrane trafficking pathway regulator, the two-pore channel subtype 2 (TPC2), which expression is affected by the levels of viral uptake *in vitro* [50], suggesting a role in the pathogenicity *in vivo* similarly to the ssRNA Ebola virus [90]. Our pathway analysis with those genes associated with COVID-19 severity revealed the *Hepatic Fibrosis and HSC Activation* as the most significant canonical pathway. The PC analysis that followed, showed population-specific segregation in the Italian, Spanish and Finnish populations from the rest of the countries analysed in this pathway only. Both the Spanish and Italian populations clustered closely together while the Finnish population segregated independently. These findings, alongside high mortality rate, make this network an area of interest to further investigate treatments for Spanish and Italian COVID-19 patients with current medications for liver disease as an option for those who develop liver malfunction during the infection while those with a pre-existing liver condition may need to go through a more thorough follow up after the recovery period. As mentioned previously, Finland has a unique genetic background originated from geographic and cultural isolation and its population has been the focus of heritable diseases studies and precision medicine treatments [91]. Interestingly, as mentioned above liver cancer is the main cause of liver disease deaths in Spain and Italy while it is alcohol consumption in Finland. This, coupled with the results observed in here, may reflect a correlation between SARS-CoV-2 severe infections in Spanish and Italian patients who suffer not only from liver disease in general but liver cancer in particular while liver damage associated with alcohol consumption does not increase mortality rate as observed in the low mortality rate reported from Finland. It is also possible that the independent opposite segregation of the Finnish population in the *Hepatic Fibrosis / HSC activation signalling* might be a reflection of the unique genetic background previously described. Therefore, we propose that some of the aggressive outcomes observed in COVID-19 patients from Spain and Italy could be partially associated with a pre-existing liver condition and/or may lead to further COVID-19 associated liver complications. Indeed, liver injury in severe COVID-19 patients has been widely reported (review by Zhang C et al. [92]) and SARS-CoV-2 infected patients with symptomatic abnormal liver function has previously accounted for up to 39.4% of the severe cases in a COVID-19 cohort (N = 1,099) [16]. Additionally, although the HSC activation can occur via several molecular and signal pathways [93] its detection could be of special interest in Spain and Italy's COVID-19 patients. Three biomarkers have recently been identified: lactic dehydrogenase (LDH), lymphocyte and high-sensitivity C-reactive protein (hs-CRP) and when combined they can predict COVID-19 mortality with > 95% accuracy [94]. This could be relevant as LDH can activate HSC [93] and an increase of this biomarker alone (which can predict survival with 92% accuracy on its own)

could be an early biomarker of severe cases in these two. On the other hand, there is growing evidence of long non coding RNAs (lncRNAs) associated with SARS-CoV-2 hijacking of mitochondrial DNA and possible lncRNA as key modulators in this [95]. To check this, we asked how many lncRNAs are associated with LDHC which is one of those key proteins regulating liver diseases. We found that NONHSAT15889.1 and NONHSAT018286.2 are the key lncRNAs that are highly expressed in testes from Reads Per Kilobase of the transcripts, per Million mapped reads (RPKM) as available from noncode.org. This agrees with the fact that many men are likely to be near more fatal to death than women, which overlies the liver diseases. As a note, co-expression analysis of ACE2 and TMPRSS2 revealed they co-express in many tissues but not in liver [87], suggesting a different S protein membrane receptors and co-activator mechanism in those severe COVID-19 cases that develop liver dysfunction during the infection. Other efforts to identify several genetic variation in the Spanish and Italian populations associated with severe COVID-19 symptoms have also been reported [96]. However, in this study they focused on potential genes associated with respiratory failure only and these two populations were not analysed in the broader European context.

In terms of future directions, genomic data of other European countries (especially Belgium) and a broader representation of the Italian population should be included in the PC analysis. Validation of the results with a second cohort, preferably with COVID-19 patients, is uncompromising to confirm the results. However, we argue that the recently established *COVID-19 Host Genetics Initiative* will allow us to answer this and many more questions [97].

Conclusion

In summary, 437 candidate human genes reported to have a role during a SARS virus infection have been identified. Pathway analysis of 21 reported genes associated with severe COVID-19 cases identified *Hepatic Fibrosis and HSC Activation* as the most significant canonical pathway. It was this pathway only where PC analysis showed specific segregation of one or more populations. The Spanish and Italian populations clustered from the rest of the countries while Finnish segregated independently. We propose COVID-19 patients hospitalised in Spain and Italy should be actively surveyed if presenting pre-existing liver conditions, especially those who suffer from liver cancer. Early prognostic biomarkers of liver malfunction in patients with a non-pre-existing condition may play an important role in reducing the death burden observed in these countries.

Conflict of Interest Statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Authors Contributions Statement

LM conducted the literature search. LM and PJ searched for candidate genes positions and run pathway analysis. LM prepared the tables and figures. SF carried out PC analysis. PS studied plausible lncRNAs associated with LDHC and conducted protein-protein interaction analysis. LM drafted the manuscript. SF, PS, PJ and JP reviewed and edited the article before submission. JB conceived the initial research hypothesis, provided vital feedback and supervised the work.

Funding

This work has partially been funded by the National Health Medical Research Council, Cancer Australia and Cancer Cancer Queensland.

Acknowledgements

We would like to thank you all the team lead by A/Prof Jyotsna Batra for their support during this study and the Cancer Council Australia.

References

1. Wang, C.; Horby, P. W.; Hayden, F. G.; Gao, G. F., A novel coronavirus outbreak of global health concern. *The Lancet* **2020**, 395, (10223), 470-473.
2. Dong, E.; Du, H.; Gardner, L., An interactive web-based dashboard to track COVID-19 in real time. *The Lancet Infectious Diseases* **2020**, 20, (5), 533-534.
3. Prevention, C. f. D. C. a. FAQ: Severe Acute Respiratory Syndrome (SARS). <https://www.cdc.gov/sars/about/fag.html> (07/04/2020),
4. Wilder-Smith, A.; Chiew, C. J.; Lee, V. J., Can we contain the COVID-19 outbreak with the same measures as for SARS? *The Lancet Infectious Diseases*.
5. Cherry, J. D., The chronology of the 2002-2003 SARS mini pandemic. *Paediatric respiratory reviews* **2004**, 5, (4), 262-9.
6. World Health Organisation, W. SARS outbreak contained worldwide. <https://www.who.int/mediacentre/news/releases/2003/pr56/en/> (07/04/2020),
7. World Health Organisation, W., International Travel and Health: SARS.
8. Chowell, G.; Castillo-Chavez, C.; Fenimore, P. W.; Kribs-Zaleta, C. M.; Arriola, L.; Hyman, J. M., Model parameters and outbreak control for SARS. *Emerging infectious diseases* **2004**, 10, (7), 1258-1263.
9. Bi, Q.; Wu, Y.; Mei, S.; Ye, C.; Zou, X.; Zhang, Z.; Liu, X.; Wei, L.; Truelove, S. A.; Zhang, T.; Gao, W.; Cheng, C.; Tang, X.; Wu, X.; Wu, Y.; Sun, B.; Huang, S.; Sun, Y.; Zhang, J.; Ma, T.; Lessler, J.; Feng, T., Epidemiology and transmission of COVID-19 in 391 cases and 1286 of their close contacts in Shenzhen, China: a retrospective cohort study. *The Lancet Infectious Diseases*.
10. Liu, Y.; Gayle, A. A.; Wilder-Smith, A.; Rocklöv, J., The reproductive number of COVID-19 is higher compared to SARS coronavirus. *Journal of Travel Medicine* **2020**, 27, (2).
11. World Health Organisation, W. Coronavirus disease (COVID-2019) situation reports. <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports>
12. Univeristy of Oxford, U.; Blavatnik School of Government, B., CORONAVIRUS GOVERNMENT RESPONSE TRACKER. In Blavatnik School of Government, 2020.
13. Guan, W.-j.; Liang, W.-h.; Zhao, Y.; Liang, H.-r.; Chen, Z.-s.; Li, Y.-m.; Liu, X.-q.; Chen, R.-c.; Tang, C.-l.; Wang, T.; Ou, C.-q.; Li, L.; Chen, P.-y.; Sang, L.; Wang, W.; Li, J.-f.; Li, C.-c.; Ou, L.-m.; Cheng, B.; Xiong, S.; Ni, Z.-y.; Xiang, J.; Hu, Y.; Liu, L.; Shan, H.; Lei, C.-l.; Peng, Y.-x.; Wei, L.; Liu, Y.; Hu, Y.-h.; Peng, P.; Wang, J.-m.; Liu, J.-y.; Chen, Z.; Li, G.; Zheng, Z.-j.; Qiu, S.-q.; Luo, J.; Ye, C.-j.; Zhu, S.-y.; Cheng, L.-l.; Ye, F.; Li, S.-y.; Zheng, J.-p.; Zhang, N.-f.; Zhong, N.-s.; He, J.-x., Comorbidity and its impact on 1590 patients with Covid-19 in China: A Nationwide Analysis. **2020**, 2000547.
14. Brake, S. J.; Barnsley, K.; Lu, W.; McAlinden, K. D.; Eapen, M. S.; Sohal, S. S., Smoking Upregulates Angiotensin-Converting Enzyme-2 Receptor: A Potential Adhesion Site for Novel Coronavirus SARS-CoV-2 (Covid-19). *Journal of clinical medicine* **2020**, 9, (3).
15. Worldometer Age, Sex, Existing Conditions of COVID-19 Cases and Deaths. <https://www.worldometers.info/coronavirus/coronavirus-age-sex-demographics/> (20/04/2020),
16. Guan, W.-j.; Ni, Z.-y.; Hu, Y.; Liang, W.-h.; Ou, C.-q.; He, J.-x.; Liu, L.; Shan, H.; Lei, C.-l.; Hui, D. S. C.; Du, B.; Li, L.-j.; Zeng, G.; Yuen, K.-Y.; Chen, R.-c.; Tang, C.-l.; Wang, T.; Chen, P.-y.; Xiang, J.; Li, S.-y.; Wang, J.-l.; Liang, Z.-j.; Peng, Y.-x.; Wei, L.; Liu, Y.; Hu, Y.-h.; Peng, P.; Wang, J.-m.; Liu, J.-y.; Chen, Z.; Li, G.; Zheng, Z.-j.; Qiu, S.-q.; Luo, J.; Ye, C.-j.; Zhu, S.-y.; Zhong, N.-s., Clinical Characteristics of Coronavirus Disease 2019 in China. *New England Journal of Medicine* **2020**, 382, (18), 1708-1720.
17. Prevention, C. f. D. C. a. Mortality Frequency Measures. <https://www.cdc.gov/csels/dsepd/ss1978/lesson3/section3.html> (04/06/2020),

18. World Health Organisation, W. *INTERNATIONAL GUIDELINES FOR CERTIFICATION AND CLASSIFICATION (CODING) OF COVID-19 AS CAUSE OF DEATH*; WHO: 20/04/2020, 2020; p 14.
19. Population 1st January 2020. In Eurostat Data Browser.
20. Caramelo, F.; Ferreira, N.; Oliveiros, B., Estimation of risk factors for COVID-19 mortality - preliminary results. **2020**, 2020.02.24.20027268.
21. Shahid, Z.; Kalayanamitra, R.; McClafferty, B.; Kepko, D.; Ramgobin, D.; Patel, R.; Aggarwal, C. S.; Vunnam, R. R.; Sahu, N.; Bhatt, D.; Jones, K.; Golamari, R.; Jain, R., COVID-19 And Older Adults: What We Know. *Journal of the American Geriatrics Society* **2020**.
22. Krämer, A.; Green, J.; Pollard, J., Jr.; Tugendreich, S., Causal analysis approaches in Ingenuity Pathway Analysis. *Bioinformatics (Oxford, England)* **2014**, 30, (4), 523-30.
23. Montojo, J.; Zuberi, K.; Rodriguez, H.; Kazi, F.; Wright, G.; Donaldson, S. L.; Morris, Q.; Bader, G. D., GeneMANIA Cytoscape plugin: fast gene function predictions on the desktop. *Bioinformatics (Oxford, England)* **2010**, 26, (22), 2927-8.
24. Purcell, S.; Neale, B.; Todd-Brown, K.; Thomas, L.; Ferreira, M. A.; Bender, D.; Maller, J.; Sklar, P.; de Bakker, P. I.; Daly, M. J.; Sham, P. C., PLINK: a tool set for whole-genome association and population-based linkage analyses. *American journal of human genetics* **2007**, 81, (3), 559-75.
25. European Union. In Europa.eu.
26. Population Reference Bureau, P. Countries With the Oldest Populations in the World. <https://www.prb.org/countries-with-the-oldest-populations/>
27. Townsend, N.; Wilson, L.; Bhatnagar, P.; Wickramasinghe, K.; Rayner, M.; Nichols, M., Cardiovascular disease in Europe: epidemiological update 2016. *European heart journal* **2016**, 37, (42), 3232-3245.
28. Patterson, C. C.; Harjutsalo, V.; Rosenbauer, J.; Neu, A.; Cinek, O.; Skrivarhaug, T.; Rami-Merhar, B.; Soltesz, G.; Svensson, J.; Parslow, R. C.; Castell, C.; Schoenle, E. J.; Bingley, P. J.; Dahlquist, G.; Jarosz-Chobot, P. K.; Marčiulionytė, D.; Roche, E. F.; Rothe, U.; Bratina, N.; Ionescu-Tirgoviste, C.; Weets, I.; Kocova, M.; Cherubini, V.; Rojnic Putarek, N.; deBeaufort, C. E.; Samardzic, M.; Green, A., Trends and cyclical variation in the incidence of childhood type 1 diabetes in 26 European centres in the 25 year period 1989–2013: a multicentre prospective registration study. *Diabetologia* **2019**, 62, (3), 408-417.
29. Khan, M. A. B.; Hashim, M. J.; King, J. K.; Govender, R. D.; Mustafa, H.; Al Kaabi, J., Epidemiology of Type 2 Diabetes - Global Burden of Disease and Forecasted Trends. *Journal of epidemiology and global health* **2020**, 10, (1), 107-111.
30. Gallus, S.; Lugo, A.; Liu, X.; Behrakis, P.; Boffi, R.; Bosetti, C.; Carreras, G.; Chatenoud, L.; Clancy, L.; Continente, X.; Dobson, R.; Effertz, T.; Filippidis, F. T.; Fu, M.; Geshanova, G.; Gorini, G.; Keogan, S.; Ivanov, H.; Lopez, M. J.; Lopez-Nicolas, A.; Precioso, J.; Przewozniak, K.; Radu-Loghin, C.; Ruprecht, A.; Semple, S.; Soriano, J. B.; Starchenko, P.; Traperro-Bertran, M.; Tigova, O.; Tzortzi, A. S.; Vardavas, C.; Vyzikidou, V. K.; Colombo, P.; Fernandez, E., Who smokes in Europe? Data from 12 European countries in the TackSHS survey (2017-2018). *Journal of epidemiology* **2020**.
31. Pimpin, L.; Cortez-Pinto, H.; Negro, F.; Corbould, E.; Lazarus, J. V.; Webber, L.; Sheron, N., Burden of liver disease in Europe: Epidemiology and analysis of risk factors to identify prevention policies. *Journal of Hepatology* **2018**, 69, (3), 718-735.
32. Gordon, D. E.; Jang, G. M.; Bouhaddou, M.; Xu, J.; Obernier, K.; White, K. M.; O'Meara, M. J.; Rezelj, V. V.; Guo, J. Z.; Swaney, D. L.; Tummino, T. A.; Huettenhain, R.; Kaake, R. M.; Richards, A. L.; Tutuncuoglu, B.; Foussard, H.; Batra, J.; Haas, K.; Modak, M.; Kim, M.; Haas, P.; Polacco, B. J.; Braberg, H.; Fabius, J. M.; Eckhardt, M.; Soucheray, M.; Bennett, M. J.; Cakir, M.; McGregor, M. J.; Li, Q.; Meyer, B.; Roesch, F.; Vallet, T.; Mac Kain, A.; Miorin, L.; Moreno, E.; Naing, Z. Z. C.; Zhou, Y.; Peng, S.; Shi, Y.; Zhang, Z.; Shen, W.; Kirby, I. T.; Melnyk, J. E.; Chorbha, J. S.; Lou, K.; Dai, S. A.; Barrio-Hernandez, I.; Memon, D.; Hernandez-Armenta, C.; Lyu, J.; Mathy, C. J. P.; Perica, T.; Pilla, K. B.; Ganesan, S. J.; Saltzberg, D. J.; Rakesh, R.; Liu,

- X.; Rosenthal, S. B.; Calviello, L.; Venkataramanan, S.; Liboy-Lugo, J.; Lin, Y.; Huang, X.-P.; Liu, Y.; Wankowicz, S. A.; Bohn, M.; Safari, M.; Ugur, F. S.; Koh, C.; Savar, N. S.; Tran, Q. D.; Shengjuler, D.; Fletcher, S. J.; O'Neal, M. C.; Cai, Y.; Chang, J. C. J.; Broadhurst, D. J.; Klippsten, S.; Sharp, P. P.; Wenzell, N. A.; Kuzuoglu, D.; Wang, H.-Y.; Trenker, R.; Young, J. M.; Cavero, D. A.; Hiatt, J.; Roth, T. L.; Rathore, U.; Subramanian, A.; Noack, J.; Hubert, M.; Stroud, R. M.; Frankel, A. D.; Rosenberg, O. S.; Verba, K. A.; Agard, D. A.; Ott, M.; Emerman, M.; Jura, N.; von Zastrow, M.; Verdin, E.; Ashworth, A.; Schwartz, O.; d'Enfert, C.; Mukherjee, S.; Jacobson, M.; Malik, H. S.; Fujimori, D. G.; Ideker, T.; Craik, C. S.; Floor, S. N.; Fraser, J. S.; Gross, J. D.; Sali, A.; Roth, B. L.; Ruggero, D.; Taunton, J.; Kortemme, T.; Beltrao, P.; Vignuzzi, M.; García-Sastre, A.; Shokat, K. M.; Shoichet, B. K.; Krogan, N. J., A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature* **2020**.
33. Haga, S.; Yamamoto, N.; Nakai-Murakami, C.; Osawa, Y.; Tokunaga, K.; Sata, T.; Yamamoto, N.; Sasazuki, T.; Ishizaka, Y., Modulation of TNF- α -converting enzyme by the spike protein of SARS-CoV and ACE2 induces TNF- α production and facilitates viral entry. **2008**, 105, (22), 7809-7814.
34. Heurich, A.; Hofmann-Winkler, H.; Gierer, S.; Liepold, T.; Jahn, O.; Pöhlmann, S., TMPRSS2 and ADAM17 Cleave ACE2 Differentially and Only Proteolysis by TMPRSS2 Augments Entry Driven by the Severe Acute Respiratory Syndrome Coronavirus Spike Protein. **2014**, 88, (2), 1293-1307.
35. Hoffmann, M.; Kleine-Weber, H.; Schroeder, S.; Kruger, N.; Herrler, T.; Erichsen, S.; Schiergens, T. S.; Herrler, G.; Wu, N. H.; Nitsche, A.; Muller, M. A.; Drosten, C.; Pöhlmann, S., SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* **2020**.
36. Miller, S.; Krijnse-Locker, J., Modification of intracellular membrane structures for virus replication. *Nature Reviews Microbiology* **2008**, 6, (5), 363-374.
37. Coutard, B.; Valle, C.; de Lamballerie, X.; Canard, B.; Seidah, N. G.; Decroly, E., The spike glycoprotein of the new coronavirus 2019-nCoV contains a furin-like cleavage site absent in CoV of the same clade. *Antiviral Research* **2020**, 176, 104742.
38. Qi, F.; Qian, S.; Zhang, S.; Zhang, Z., Single cell RNA sequencing of 13 human tissues identify cell types and receptors of human coronaviruses. *Biochemical and Biophysical Research Communications* **2020**.
39. Yeager, C. L.; Ashmun, R. A.; Williams, R. K.; Cardellicchio, C. B.; Shapiro, L. H.; Look, A. T.; Holmes, K. V., Human aminopeptidase N is a receptor for human coronavirus 229E. *Nature* **1992**, 357, (6377), 420-2.
40. Raj, V. S.; Mou, H.; Smits, S. L.; Dekkers, D. H.; Muller, M. A.; Dijkman, R.; Muth, D.; Demmers, J. A.; Zaki, A.; Fouchier, R. A.; Thiel, V.; Drosten, C.; Rottier, P. J.; Osterhaus, A. D.; Bosch, B. J.; Haagmans, B. L., Dipeptidyl peptidase 4 is a functional receptor for the emerging human coronavirus-EMC. *Nature* **2013**, 495, (7440), 251-4.
41. Muramatsu, T., Basigin (CD147), a multifunctional transmembrane glycoprotein with various binding partners. *J Biochem* **2016**, 159, (5), 481-490.
42. Wang, K.; Chen, W.; Zhou, Y.-S.; Lian, J.-Q.; Zhang, Z.; Du, P.; Gong, L.; Zhang, Y.; Cui, H.-Y.; Geng, J.-J.; Wang, B.; Sun, X.-X.; Wang, C.-F.; Yang, X.; Lin, P.; Deng, Y.-Q.; Wei, D.; Yang, X.-M.; Zhu, Y.-M.; Zhang, K.; Zheng, Z.-H.; Miao, J.-L.; Guo, T.; Shi, Y.; Zhang, J.; Fu, L.; Wang, Q.-Y.; Bian, H.; Zhu, P.; Chen, Z.-N., SARS-CoV-2 invades host cells via a novel route: CD147-spike protein. **2020**, 2020.03.14.988345.
43. Jeffers, S. A.; Tusell, S. M.; Gillim-Ross, L.; Hemmila, E. M.; Achenbach, J. E.; Babcock, G. J.; Thomas, W. D.; Thackray, L. B.; Young, M. D.; Mason, R. J.; Ambrosino, D. M.; Wentworth, D. E.; DeMartini, J. C.; Holmes, K. V., CD209L (L-SIGN) is a receptor for severe acute respiratory syndrome coronavirus. **2004**, 101, (44), 15748-15753.
44. Lozach, P.-Y.; Burleigh, L.; Staropoli, I.; Amara, A., The C type lectins DC-SIGN and L-SIGN: receptors for viral glycoproteins. *Methods Mol Biol* **2007**, 379, 51-68.

45. Chan, V. S. F.; Chan, K. Y. K.; Chen, Y.; Poon, L. L. M.; Cheung, A. N. Y.; Zheng, B.; Chan, K.-H.; Mak, W.; Ngan, H. Y. S.; Xu, X.; Screaton, G.; Tam, P. K. H.; Austyn, J. M.; Chan, L.-C.; Yip, S.-P.; Peiris, M.; Khoo, U.-S.; Lin, C.-L. S., Homozygous L-SIGN (CLEC4M) plays a protective role in SARS coronavirus infection. *Nature Genetics* **2006**, *38*, (1), 38-46.
46. Alvarez, C. P.; Lasala, F.; Carrillo, J.; Muñiz, O.; Corbí, A. L.; Delgado, R., C-Type Lectins DC-SIGN and L-SIGN Mediate Cellular Entry by Ebola Virus in *cis* and in *trans*. **2002**, *76*, (13), 6841-6844.
47. Vankadari, N.; Wilce, J. A., Emerging COVID-19 coronavirus: glycan shield and structure prediction of spike glycoprotein and its interaction with human CD26. *Emerging Microbes & Infections* **2020**, *9*, (1), 601-604.
48. Kleine-Weber, H.; Schroeder, S.; Krüger, N.; Prokscha, A.; Naim, H. Y.; Müller, M. A.; Drosten, C.; Pöhlmann, S.; Hoffmann, M., Polymorphisms in dipeptidyl peptidase 4 reduce host cell entry of Middle East respiratory syndrome coronavirus. *Emerging Microbes & Infections* **2020**, *9*, (1), 155-168.
49. Ibrahim, I. M.; Abdelmalek, D. H.; Elshahat, M. E.; Elfiky, A. A., COVID-19 spike-host cell receptor GRP78 binding site prediction. *The Journal of Infection* **2020**.
50. Ou, X.; Liu, Y.; Lei, X.; Li, P.; Mi, D.; Ren, L.; Guo, L.; Guo, R.; Chen, T.; Hu, J.; Xiang, Z.; Mu, Z.; Chen, X.; Chen, J.; Hu, K.; Jin, Q.; Wang, J.; Qian, Z., Characterization of spike glycoprotein of SARS-CoV-2 on virus entry and its immune cross-reactivity with SARS-CoV. *Nature communications* **2020**, *11*, (1), 1620.
51. Kamitani, W.; Huang, C.; Narayanan, K.; Lokugamage, K. G.; Makino, S., A two-pronged strategy to suppress host protein synthesis by SARS coronavirus Nsp1 protein. *Nature structural & molecular biology* **2009**, *16*, (11), 1134-40.
52. Tanaka, T.; Kamitani, W.; DeDiego, M. L.; Enjuanes, L.; Matsuura, Y., Severe Acute Respiratory Syndrome Coronavirus nsp1 Facilitates Efficient Propagation in Cells through a Specific Translational Shutoff of Host mRNA. **2012**, *86*, (20), 11128-11137.
53. Law, A. H. Y.; Lee, D. C. W.; Cheung, B. K. W.; Yim, H. C. H.; Lau, A. S. Y., Role for Nonstructural Protein 1 of Severe Acute Respiratory Syndrome Coronavirus in Chemokine Dysregulation. **2007**, *81*, (1), 416-422.
54. Chu, H.; Chan, J. F.; Wang, Y.; Yuen, T. T.; Chai, Y.; Hou, Y.; Shuai, H.; Yang, D.; Hu, B.; Huang, X.; Zhang, X.; Cai, J. P.; Zhou, J.; Yuan, S.; Kok, K. H.; To, K. K.; Chan, I. H.; Zhang, A. J.; Sit, K. Y.; Au, W. K.; Yuen, K. Y., Comparative replication and immune activation profiles of SARS-CoV-2 and SARS-CoV in human lungs: an ex vivo study with implications for the pathogenesis of COVID-19. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* **2020**.
55. Sun, L.; Xing, Y.; Chen, X.; Zheng, Y.; Yang, Y.; Nichols, D. B.; Clementz, M. A.; Banach, B. S.; Li, K.; Baker, S. C.; Chen, Z., Coronavirus papain-like proteases negatively regulate antiviral innate immune response through disruption of STING-mediated signaling. *PLoS One* **2012**, *7*, (2), e30802.
56. Chen, X.; Yang, X.; Zheng, Y.; Yang, Y.; Xing, Y.; Chen, Z., SARS coronavirus papain-like protease inhibits the type I interferon signaling pathway through interaction with the STING-TRAF3-TBK1 complex. *Protein & Cell* **2014**, *5*, (5), 369-81.
57. Chang, Y. J.; Liu, C. Y.; Chiang, B. L.; Chao, Y. C.; Chen, C. C., Induction of IL-8 release in lung cells via activator protein-1 by recombinant baculovirus displaying severe acute respiratory syndrome-coronavirus spike proteins: identification of two functional regions. *Journal of immunology (Baltimore, Md. : 1950)* **2004**, *173*, (12), 7602-14.
58. Nieto-Torres, J. L.; Verdiá-Báguena, C.; Jimenez-Guardeño, J. M.; Regla-Nava, J. A.; Castaño-Rodríguez, C.; Fernandez-Delgado, R.; Torres, J.; Aguilella, V. M.; Enjuanes, L., Severe acute respiratory syndrome coronavirus E protein transports calcium ions and activates the NLRP3 inflammasome. *Virology* **2015**, *485*, 330-339.

59. Wathelet, M. G.; Orr, M.; Frieman, M. B.; Baric, R. S., Severe Acute Respiratory Syndrome Coronavirus Evades Antiviral Signaling: Role of nsp1 and Rational Design of an Attenuated Strain. **2007**, 81, (21), 11620-11633.
60. Frieman, M.; Ratia, K.; Johnston, R. E.; Mesecar, A. D.; Baric, R. S., Severe Acute Respiratory Syndrome Coronavirus Papain-Like Protease Ubiquitin-Like Domain and Catalytic Domain Regulate Antagonism of IRF3 and NF- κ B Signaling. **2009**, 83, (13), 6689-6705.
61. Siu, K.-L.; Yuen, K.-S.; Castano-Rodriguez, C.; Ye, Z.-W.; Yeung, M.-L.; Fung, S.-Y.; Yuan, S.; Chan, C.-P.; Yuen, K.-Y.; Enjuanes, L.; Jin, D.-Y., Severe acute respiratory syndrome Coronavirus ORF3a protein activates the NLRP3 inflammasome by promoting TRAF3-dependent ubiquitination of ASC. **2019**, 33, (8), 8865-8877.
62. Saikatendu, K. S.; Joseph, J. S.; Subramanian, V.; Clayton, T.; Griffith, M.; Moy, K.; Velasquez, J.; Neuman, B. W.; Buchmeier, M. J.; Stevens, R. C.; Kuhn, P., Structural basis of severe acute respiratory syndrome coronavirus ADP-ribose-1''-phosphate dephosphorylation by a conserved domain of nsP3. *Structure (London, England : 1993)* **2005**, 13, (11), 1665-75.
63. Forni, D.; Cagliani, R.; Mozzi, A.; Pozzoli, U.; Al-Daghri, N.; Clerici, M.; Sironi, M., Extensive Positive Selection Drives the Evolution of Nonstructural Proteins in Lineage C Betacoronaviruses. *J Virol* **2016**, 90, (7), 3627-3639.
64. Abraham, R.; Hauer, D.; McPherson, R. L.; Utt, A.; Kirby, I. T.; Cohen, M. S.; Merits, A.; Leung, A. K. L.; Griffin, D. E., ADP-ribosyl-binding and hydrolase activities of the alphavirus nsP3 macrodomain are critical for initiation of virus replication. **2018**, 115, (44), E10457-E10466.
65. Sakai, Y.; Kawachi, K.; Terada, Y.; Omori, H.; Matsuura, Y.; Kamitani, W., Two-amino acids change in the nsp4 of SARS coronavirus abolishes viral replication. *Virology* **2017**, 510, 165-174.
66. Srinivasan, S.; Cui, H.; Gao, Z.; Liu, M.; Lu, S.; Mkandawire, W.; Narykov, O.; Sun, M.; Korkin, D., Structural Genomics of SARS-CoV-2 Indicates Evolutionary Conserved Functional Regions of Viral Proteins. *Viruses* **2020**, 12, (4).
67. Ye, Q.; Wang, B.; Mao, J., The pathogenesis and treatment of the 'Cytokine Storm' in COVID-19. *Journal of Infection* **2020**.
68. Liu, P. P.; Blet, A.; Smyth, D.; Li, H., The Science Underlying COVID-19: Implications for the Cardiovascular System. **2020**, 0, (0).
69. Tan, L.; Wang, Q.; Zhang, D.; Ding, J.; Huang, Q.; Tang, Y.-Q.; Wang, Q.; Miao, H., Lymphopenia predicts disease severity of COVID-19: a descriptive and predictive study. *Signal Transduction and Targeted Therapy* **2020**, 5, (1), 33.
70. Chen, G.; Wu, D.; Guo, W.; Cao, Y.; Huang, D.; Wang, H.; Wang, T.; Zhang, X.; Chen, H.; Yu, H.; Zhang, X.; Zhang, M.; Wu, S.; Song, J.; Chen, T.; Han, M.; Li, S.; Luo, X.; Zhao, J.; Ning, Q., Clinical and immunological features of severe and moderate coronavirus disease 2019. *The Journal of clinical investigation* **2020**, 130, (5), 2620-2629.
71. Huang, C.; Wang, Y.; Li, X.; Ren, L.; Zhao, J.; Hu, Y.; Zhang, L.; Fan, G.; Xu, J.; Gu, X.; Cheng, Z.; Yu, T.; Xia, J.; Wei, Y.; Wu, W.; Xie, X.; Yin, W.; Li, H.; Liu, M.; Xiao, Y.; Gao, H.; Guo, L.; Xie, J.; Wang, G.; Jiang, R.; Gao, Z.; Jin, Q.; Wang, J.; Cao, B., Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *The Lancet* **2020**, 395, (10223), 497-506.
72. Zhao, Q.; Meng, M.; Kumar, R.; Wu, Y.; Huang, J.; Deng, Y.; Weng, Z.; Yang, L., Lymphopenia is associated with severe coronavirus disease 2019 (COVID-19) infections: A systemic review and meta-analysis. *Int J Infect Dis* **2020**, 96, 131-135.
73. Ingenuity Pathway Analysis, Q., Caveolar-mediated Endocytosis Signaling. In Ingenuity Target Explorer, 2020.
74. Ingenuity Pathway Analysis, Q., MSP-RON Signaling Pathway. In Ingenuity Target Explorer, 2020.
75. Ingenuity Pathway Analysis, Q., Hepatic Fibrosis / Hepatic Stellate Cell Activation. In Ingenuity Target Explorer, 2020.

76. Ingenuity Pathway Analysis, Q., the Communication between Innate and Adaptive Immune Cells In Ingenuity Target Explorer, 2020.
77. Chheda, H.; Palta, P.; Pirinen, M.; McCarthy, S.; Walter, K.; Koskinen, S.; Salomaa, V.; Daly, M.; Durbin, R.; Palotie, A.; Aittokallio, T.; Ripatti, S., Whole-genome view of the consequences of a population bottleneck using 2926 genome sequences from Finland and United Kingdom. *European journal of human genetics : EJHG* **2017**, 25, (4), 477-484.
78. Locke, A. E.; Steinberg, K. M.; Chiang, C. W. K.; Service, S. K.; Havulinna, A. S.; Stell, L.; Pirinen, M.; Abel, H. J.; Chiang, C. C.; Fulton, R. S.; Jackson, A. U.; Kang, C. J.; Kanchi, K. L.; Koboldt, D. C.; Larson, D. E.; Nelson, J.; Nicholas, T. J.; Pietilä, A.; Ramensky, V.; Ray, D.; Scott, L. J.; Stringham, H. M.; Vangipurapu, J.; Welch, R.; Yajnik, P.; Yin, X.; Eriksson, J. G.; Ala-Korpela, M.; Järvelin, M.-R.; Männikkö, M.; Laivuori, H.; Dutcher, S. K.; Stitzel, N. O.; Wilson, R. K.; Hall, I. M.; Sabatti, C.; Palotie, A.; Salomaa, V.; Laakso, M.; Ripatti, S.; Boehnke, M.; Freimer, N. B.; FinnGen, P., Exome sequencing of Finnish isolates enhances rare-variant association power. *Nature* **2019**, 572, (7769), 323-328.
79. Dopazo, J.; Amadoz, A.; Bleda, M.; Garcia-Alonso, L.; Alemán, A.; García-García, F.; Rodriguez, J. A.; Daub, J. T.; Muntané, G.; Rueda, A.; Vela-Boza, A.; López-Domingo, F. J.; Florido, J. P.; Arce, P.; Ruiz-Ferrer, M.; Méndez-Vidal, C.; Arnold, T. E.; Spleiss, O.; Alvarez-Tejado, M.; Navarro, A.; Bhattacharya, S. S.; Borrego, S.; Santoyo-López, J.; Antiñolo, G., 267 Spanish Exomes Reveal Population-Specific Differences in Disease-Related Genetic Variation. *Mol Biol Evol* **2016**, 33, (5), 1205-1218.
80. Gayán, J.; Galan, J. J.; González-Pérez, A.; Sáez, M. E.; Martínez-Larrad, M. T.; Zabena, C.; Rivero, M. C.; Salinas, A.; Ramírez-Lorca, R.; Morón, F. J.; Royo, J. L.; Moreno-Rey, C.; Velasco, J.; Carrasco, J. M.; Molero, E.; Ochoa, C.; Ochoa, M. D.; Gutiérrez, M.; Reina, M.; Pascual, R.; Romo-Astorga, A.; Susillo-González, J. L.; Vázquez, E.; Real, L. M.; Ruiz, A.; Serrano-Ríos, M., Genetic structure of the Spanish population. *BMC Genomics* **2010**, 11, 326-326.
81. Parolo, S.; Lisa, A.; Gentilini, D.; Di Blasio, A. M.; Barlera, S.; Nicolis, E. B.; Boncoraglio, G. B.; Parati, E. A.; Bione, S., Characterization of the biological processes shaping the genetic structure of the Italian population. *BMC Genetics* **2015**, 16, (1), 132.
82. Turchi, C.; Piva, F.; Solito, G.; Principato, G.; Buscemi, L.; Tagliabracchi, A., ADH4 intronic variations are associated with alcohol dependence: results from an Italian case-control association study. *Pharmacogenetics and genomics* **2012**, 22, (2), 79-94.
83. Munoz, X.; Amiano, P.; Celorrio, D.; Dorronsoro, M.; Sanchez, M. J.; Huerta, J. M.; Barricarte, A.; Arriola, L.; Navarro, C.; Molina-Montes, E.; Chirlaque, M. D.; Ardanaz, E.; Rodriguez, L.; Duell, E. J.; Hijona, E.; Herreros-Villanueva, M.; Sala, N.; Bujanda, L., Association of alcohol dehydrogenase polymorphisms and life-style factors with excessive alcohol intake within the Spanish population (EPIC-Spain). *Addiction (Abingdon, England)* **2012**, 107, (12), 2117-27.
84. Braun, E.; Sauter, D., Furin-mediated protein processing in infectious diseases and cancer. *Clin Transl Immunology* **2019**, 8, (8), e1073-e1073.
85. Guzzi, P. H.; Mercatelli, D.; Ceraolo, C.; Giorgi, F. M., Master Regulator Analysis of the SARS-CoV-2/Human Interactome. *Journal of clinical medicine* **2020**, 9, (4).
86. Bertram, S.; Dijkman, R.; Habjan, M.; Heurich, A.; Gierer, S.; Glowacka, I.; Welsch, K.; Winkler, M.; Schneider, H.; Hofmann-Winkler, H.; Thiel, V.; Pöhlmann, S., TMPRSS2 Activates the Human Coronavirus 229E for Cathepsin-Independent Host Cell Entry and Is Expressed in Viral Target Cells in the Respiratory Epithelium. **2013**, 87, (11), 6150-6160.
87. Sungnak, W.; Huang, N.; Bécavin, C.; Berg, M.; Queen, R.; Litvinukova, M.; Talavera-López, C.; Maatz, H.; Reichart, D.; Sampaziotis, F.; Worlock, K. B.; Yoshida, M.; Barnes, J. L.; Banovich, N. E.; Barbry, P.; Brazma, A.; Collin, J.; Desai, T. J.; Duong, T. E.; Eickelberg, O.; Falk, C.; Farzan, M.; Glass, I.; Gupta, R. K.; Haniffa, M.; Horvath, P.; Hubner, N.; Hung, D.; Kaminski, N.; Krasnow, M.; Kropski, J. A.; Kuhnemund, M.; Lako, M.; Lee, H.; Leroy, S.; Linnarson, S.; Lundeberg, J.; Meyer, K. B.; Miao, Z.; Misharin, A. V.; Nawijn, M. C.; Nikolic, M. Z.; Nosedá,

- M.; Ordovas-Montanes, J.; Oudit, G. Y.; Pe'er, D.; Powell, J.; Quake, S.; Rajagopal, J.; Tata, P. R.; Rawlins, E. L.; Regev, A.; Reyfman, P. A.; Rozenblatt-Rosen, O.; Saeb-Parsy, K.; Samakovlis, C.; Schiller, H. B.; Schultze, J. L.; Seibold, M. A.; Seidman, C. E.; Seidman, J. G.; Shalek, A. K.; Shepherd, D.; Spence, J.; Spira, A.; Sun, X.; Teichmann, S. A.; Theis, F. J.; Tsankov, A. M.; Vallier, L.; van den Berge, M.; Whitsett, J.; Xavier, R.; Xu, Y.; Zaragosi, L.-E.; Zerti, D.; Zhang, H.; Zhang, K.; Rojas, M.; Figueiredo, F.; Network, H. C. A. L. B., SARS-CoV-2 entry factors are highly expressed in nasal epithelial cells together with innate immune genes. *Nature Medicine* **2020**.
88. To, K. F.; Lo, A. W., Exploring the pathogenesis of severe acute respiratory syndrome (SARS): the tissue distribution of the coronavirus (SARS-CoV) and its putative receptor, angiotensin-converting enzyme 2 (ACE2). *The Journal of pathology* **2004**, 203, (3), 740-3.
89. Baig, A. M.; Khaleeq, A.; Ali, U.; Syeda, H., Evidence of the COVID-19 Virus Targeting the CNS: Tissue Distribution, Host-Virus Interaction, and Proposed Neurotropic Mechanisms. *ACS chemical neuroscience* **2020**, 11, (7), 995-998.
90. Falasca, L.; Agrati, C.; Petrosillo, N.; Di Caro, A.; Capobianchi, M. R.; Ippolito, G.; Piacentini, M., Molecular mechanisms of Ebola virus pathogenesis: focus on cell death. *Cell Death Differ* **2015**, 22, (8), 1250-1259.
91. Kääriäinen, H.; Muilu, J.; Perola, M.; Kristiansson, K., Genetics in an isolated population like Finland: a different basis for genomic medicine? *J Community Genet* **2017**, 8, (4), 319-326.
92. Zhang, C.; Shi, L.; Wang, F. S., Liver injury in COVID-19: management and challenges. *The lancet. Gastroenterology & hepatology* **2020**, 5, (5), 428-430.
93. Hou, W.; Syn, W.-K., Role of Metabolism in Hepatic Stellate Cell Activation and Fibrogenesis. *Front Cell Dev Biol* **2018**, 6, 150-150.
94. Yan, L.; Zhang, H.-T.; Goncalves, J.; Xiao, Y.; Wang, M.; Guo, Y.; Sun, C.; Tang, X.; Jing, L.; Zhang, M.; Huang, X.; Xiao, Y.; Cao, H.; Chen, Y.; Ren, T.; Wang, F.; Xiao, Y.; Huang, S.; Tan, X.; Huang, N.; Jiao, B.; Cheng, C.; Zhang, Y.; Luo, A.; Mombaerts, L.; Jin, J.; Cao, Z.; Li, S.; Xu, H.; Yuan, Y., An interpretable mortality prediction model for COVID-19 patients. *Nature Machine Intelligence* **2020**, 2, (5), 283-288.
95. Singh, K. K.; Chaubey, G.; Chen, J. Y.; Suravajhala, P., Decoding SARS-CoV-2 Hijacking of Host Mitochondria in Pathogenesis of COVID-19. *American journal of physiology. Cell physiology* **2020**.
96. Ellinghaus, D.; Degenhardt, F.; Bujanda, L.; Buti, M.; Albillos, A.; Invernizzi, P.; Fernandez, J.; Prati, D.; Baselli, G.; Asselta, R.; Grimsrud, M. M.; Milani, C.; Aziz, F.; Kassens, J.; May, S.; Wendorff, M.; Wienbrandt, L.; Uellendahl-Werth, F.; Zheng, T.; Yi, X.; de Pablo, R.; Chercoles, A. G.; Palom, A.; Garcia-Fernandez, A.-E.; Rodriguez-Frias, F.; Zanella, A.; Bandera, A.; Protti, A.; Aghemo, A.; Lleo de Nalda, A.; Biondi, A.; Caballero-Garralda, A.; Gori, A.; Tanck, A.; Latiano, A.; Fracanzani, A. L.; Peschuck, A.; Julia, A.; Pesenti, A.; Voza, A.; Jimenez, D.; Mateos, B.; Jimenez, B. N.; Quereda, C.; Angelini, C.; Cea, C.; Solier, A.; Pestana, D.; Sandoval, E.; Paraboschi, E. M.; Navas, E.; Ceriotti, F.; Martinelli-Boneschi, F.; Peyvandi, F.; Blasi, F.; Tellez, L.; Blanco-Grau, A.; Grasselli, G.; Costantino, G.; Cardamone, G.; Foti, G.; Aneli, S.; Kurihara, H.; ElAbd, H.; My, I.; Martin, J.; Erdmann, J.; Ferrusquia-Acosta, J.; Garcia-Etxebarria, K.; Izquierdo-Sanchez, L.; Bettini, L. R.; Terranova, L.; Moreira, L.; Santoro, L.; Scudeller, L.; Mesonero, F.; Roade, L.; Schaefer, M.; Carrabba, M.; Riveiro Barciela, M. d. M.; Basso, M. E. F.; Valsecchi, M. G.; Hernandez-Tejero, M.; Acosta-Herrera, M.; Angio, M.; Baldini, M.; Cazzaniga, M.; Schulzky, M.; Cecconi, M.; Wittig, M.; Ciccarelli, M.; Rodriguez-Gandia, M.; Boccione, M.; Miozzo, M.; Braun, N.; Martinez, N.; Palmieri, O.; Faverio, P.; Preatoni, P.; Bonfanti, P.; Omodei, P.; Tentorio, P.; Castro, P.; Rodrigues, P. M.; Blandino Ortiz, A.; Roca, R. F.; Gualtierotti, R.; Nieto, R.; Badalamenti, S.; Marsal, S.; Matullo, G.; Pelusi, S.; Monzani, V.; Wesse, T.; Pumarola, T.; Rimoldi, V.; Bosari, S.; Albrecht, W.; Peter, W.; Gomez, M. R.; Amato, M.; Duga, S.; Banales, J. M.; Hov, J. R.; Folseraas, T.; Valenti, L.; Franke, A.; Karlsen, T. H., The ABO blood group locus and a chromosome 3 gene cluster

- associate with SARS-CoV-2 respiratory failure in an Italian-Spanish genome-wide association analysis. *medRxiv* **2020**, 2020.05.31.20114991.
97. The, C.-H. G. I., The COVID-19 Host Genetics Initiative, a global initiative to elucidate the role of host genetic factors in susceptibility and severity of the SARS-CoV-2 virus pandemic. *European Journal of Human Genetics* **2020**, *28*, (6), 715-718.
 98. Cava, C.; Bertoli, G.; Castiglioni, I., In Silico Discovery of Candidate Drugs against Covid-19. *Viruses* **2020**, *12*, (4).
 99. Khan, R. J.; Jha, R. K.; Amera, G.; Jain, M.; Singh, E.; Pathak, A.; Singh, R. P.; Muthukumar, J.; Singh, A. K., Targeting SARS-CoV-2: A Systematic Drug Repurposing Approach to Identify Promising Inhibitors Against 3C-like Proteinase and 2'-O-RiboseMethyltransferase. *Journal of biomolecular structure & dynamics* **2020**, 1-40.
 100. Pfefferle, S.; Schöpf, J.; Kögl, M.; Friedel, C. C.; Müller, M. A.; Carbajo-Lozoya, J.; Stellberger, T.; von Dall'Armi, E.; Herzog, P.; Kallies, S.; Niemeyer, D.; Ditt, V.; Kuri, T.; Züst, R.; Pumpor, K.; Hilgenfeld, R.; Schwarz, F.; Zimmer, R.; Steffen, I.; Weber, F.; Thiel, V.; Herrler, G.; Thiel, H.-J.; Schwegmann-Weßels, C.; Pöhlmann, S.; Haas, J.; Drosten, C.; von Brunn, A., The SARS-Coronavirus-Host Interactome: Identification of Cyclophilins as Target for Pan-Coronavirus Inhibitors. *PLOS Pathogens* **2011**, *7*, (10), e1002331.
 101. Taylor, J. K.; Coleman, C. M.; Postel, S.; Sisk, J. M.; Bernbaum, J. G.; Venkataraman, T.; Sundberg, E. J.; Frieman, M. B., Severe Acute Respiratory Syndrome Coronavirus ORF7a Inhibits Bone Marrow Stromal Antigen 2 Virion Tethering through a Novel Mechanism of Glycosylation Interference. *J Virol* **2015**, *89*, (23), 11820-11833.
 102. Zhou, Y.; Hou, Y.; Shen, J.; Huang, Y.; Martin, W.; Cheng, F., Network-based drug repurposing for novel coronavirus 2019-nCoV/SARS-CoV-2. *Cell Discovery* **2020**, *6*, (1), 14.
 103. Cottam, E. M.; Whelband, M. C.; Wileman, T., Coronavirus NSP6 restricts autophagosome expansion. *Autophagy* **2014**, *10*, (8), 1426-1441.

Table 1: COVID-19 mortality rate per 100,000 people in the European Union.

	Country	Deaths in 7th April	Country	Deaths in 4th June
1	Spain	27.9	Belgium	83.1
2	Italy	27.2	UK	59.9
3	Belgium	14.6	Spain	59.5
4	France	13.6	Italy	55.7
5	Netherlands	10.7	Sweden	44.4
6	UK	7.9	France	43.2
7	Luxembourg	6.8	Netherlands	34.6
8	Sweden	4.7	Ireland	33.8
9	Ireland	3.5	Romania	24.2
10	Denmark	3.2	Luxembourg	17.9

Total country population as of 1st of January 2020 (<https://ec.europa.eu/eurostat/databrowser/>) [19] and COVID-19 total deaths from WHO daily reports [11].

Table 2: Host cell candidate genes associated with viral entry.

Candidate genes	Virus	Host cell response/viral mechanism	Assay methods	Ref.
ACE2	SARS-CoV-1, SARS-CoV-2	High S binding affinity. Facilitating host cell recognition.	Multiple <i>in vitro</i> and <i>in silico</i> analysis.	[33-36, 50, 85]
TMPRSS2	SARS-CoV-1, SARS-CoV-2	S protein activator, leading to viral membrane conformational change and facilitating SARS virus.	Multiple <i>in vitro</i> and <i>in silico</i> analysis.	[34, 35, 38, 86]
BSG	SARS-CoV-2, malaria, HIV, HepB and HHV	<i>Basigin</i> genes, which encode CD147 transmembrane glycoprotein recognised by several pathogens. CD147 directly binds to SARS-CoV-2 S protein affecting viral replication.	Review and <i>in vitro</i> . SARS-CoV-2 strain isolated from COVID-19 patients. Direct in vitro infection, Co-IP and ELISA.	[41, 42]
HAT	SARS-CoV-1, HCoV (229E)	<i>Histone acetyltransferases</i> family, encoding for a family of cell nuclear enzymes. They contribute to SARS-CoV-1 entry, but it is not essential for S protein activation.	<i>In vitro</i> . Both studies: gene cloning, lentiviral expression system, protein expression and cell-cell fusion analysis.	[34, 86]
CLEC4M	SARS-CoV-1, EVD, Dengue, HCV, CMV, Sindbis, HIV	<i>C-type lectin domain family 4 member M</i> genes encode for L-SIGN membrane receptor, recognised by the S protein. Homozygous L-SIGN associated with SARS disease protective role.	<i>In vitro</i> . Infection of SARS-CoV-1 human cells, gene expression, cDNA library, IHC assays. Genetic risk association from SARS patients and controls.	[43-46]
ANPEP, ENPEP, DPP4 (or CD26)	ACE2 studies, HCoV-22944, MERS-CoV45	Closest co-expression of these three peptidases (R>0.8) with ACE2 in different human tissues. HCoV-22944 binds to ENPEP while MERS to DPP4.	Single cell <i>in silico</i> ligand-receptor affinity assays. Data sourced from GEO, Human Cell Atlas, Viral Receptor and Membranome databases.	[38-40]
Cathepsin-B-L	SARS-CoV-1, SARS-CoV-2, MERS-CoV	Facilitates SARS-CoV-2 cell entry by virus-cell membrane fusion mechanisms but its inhibition does not disable virus entry.	<i>In vitro</i> . SARS-CoV-2 S protein pseudovirus system in human lung cell model.	[35, 50]

ACE2: angiotensin I converting enzyme 2, TMPRSS2: transmembrane Serine Protease 2, ANPEP: alanyl aminopeptidase membrane, ENPEP: glutamyl aminopeptidase, DPP4: dipeptidyl peptidase 4, SARS: Severe acute respiratory syndrome, HIV: human immunodeficiency virus, HepB: hepatitis B virus, HHV: human Herpesviridae, HCoV: human coronavirus, EVD: Ebola virus disease, HCV: hepatitis C virus, CMV: Cytomegalovirus, MERS: middle East respiratory syndrome, S: spike viral protein, CD147 (cluster of differentiation 147), L-SIGN: liver/lymph node-specific intracellular adhesion molecules-3 grabbing non-integrin, Co-IP: co-immunoprecipitation, IHC: immunohistochemistry

Table 3: Host cell candidate genes associated with viral immune system evasion.

Candidate genes	Virus	Host cell response/viral mechanism	Research assay	Ref.
40s	Nsp1 studies	Encoding for ribosomal protein S3, interacts with viral Nsp1, inhibiting the host's protein translation by capping the 5'mRNA.	<i>In vitro</i> . Reporter gene assays followed by transcriptomics, RNA immunoprecipitation and proteomics assays.	[51, 52]
CCL5, CCL3, CXCL10	SARS-CoV-1,	These genes encode for IP-10* protein. Increased levels in lung epithelial cells after Nsp1-direct activation of the NF-kB pathway. IP-10 showed specific up-regulation in COVID-19 lung model (when compared to SARS patients).	<i>In vitro</i> . Gene cloning, mRNA and protein expression analysis. <i>Ex vivo</i> . Lung tissue transfected with COVID-19 and gene expression analysis.	[53, 54]
STING1, TRAF3, TBK1, IKKε	SARS-CoV-1, HCoV (NL63)	The SARS-CoV-1 PLP transmembrane protein interacts with STING, TRAF3, TBK1 and IKKε disrupting the STING/TBK1/IKKε complex formation and suppressing production of IFN-α and IFN-β, vital for initial innate immune response. PLP protein highly conserved in both SARS-CoV viruses, highlighting the use of potential agonists for this protein as treatments.	<i>In vitro</i> . SARS-CoV-1 propagation, and plasmids expressing genes of interest's co-transduction. Co-IP and ubiquitination signalling detection. <i>In silico</i> . Homology alignments of both SARS viruses, approved compounds database screening and homology models predictions.	[55, 56]
ADP-ribose	ssRNA	After binding to Nsp3, post-translational modification of PARP15, PARP14 and PARP10 associated with anti-viral response.	<i>In vitro</i> . Cloning, gene expression, mutagenesis, protein purification and crystallization. <i>In silico</i> , sequence alignments, glycosylation sites' predictions and 3D mapping.	[62-64]

40s: 40 subunit, CCL5: C-C motif chemokine ligand 5, CCL3: C-C motif chemokine ligand 3, CXCL10: C-X-C Motif Chemokine Ligand 10, STING1: stimulator of interferon response cGAMP interactor 1, TRAF3: TNF Receptor Associated Factor 3, IKKε: inhibitor of nuclear factor kappa-B kinase subunit epsilon, Nsp1: non-structural protein 1, SARS: Severe acute respiratory syndrome, HCoV: human coronavirus, ssRNA: single strand RNA, *previously known as IP-10, now CXCL10: C-X-C Motif Chemokine Ligand 10, NF-kB: nuclear factor-kB, PLP: papain-like protein, COVID-19: coronavirus disease 2019, PARP genes: Poly(ADP-Ribose) Polymerase

Table 4: Candidate genes expressed in SARS and COVID-19 patients and disease severity.

Candidate genes	Disease	Reported observations	Ref.
<i>IL-1B, IFN-γ, CXCL10, and MCP-1</i>	SARS and COVID-19	Increased in SARS and COVID-19 patients. Associated with Th1 cell immunity aberrant response and ARDS	[67]
<i>IL-4 and IL-10</i>	COVID-19	Increased in COVID-19 patients. Associated with Th2 cell immunity response, facilitating further ARDS	[67]
<i>IL-2R and IL-6</i>	COVID-19	High expression levels positively correlated with the severity of the disease	[54, 67, 68]
<i>CXCL10, MCP-1 and TNF-α.</i>	COVID-19	COVID-19 ICU patients have increased serum levels of these genes when compared to non-ICU COVID-19 patients	[54, 71]
<i>IL-1β, IL-6, cRP</i>	COVID-19	High levels IL-1β of associated with poor prognosis. IL-6 and cRP potential early risk biomarkers	[68]
<i>Serpine1*</i>	COVID-19	High levels associated with vascular inflammation and higher risk of thrombosis.	[68]
<i>LDH-hsCRP-lymphocyte</i>	COVID-19	Very accurate (>90%) predictive mortality biomarker signature	[94]

IL: interleukin, IFN: interferon, MCP-1: monocyte chemoattractant protein 1, CXCL10: C-X-C Motif Chemokine Ligand 10, hsCRP: high-sensitivity C-reactive protein, LDH: lactic dehydrogenase, SARS: Severe acute respiratory syndrome, COVID-19: coronavirus disease 2019, Th: T helper type, ARDS: Acute respiratory distress syndrome, ICU: intensive unit care, *encodes PAI-I protein

PubMed searched terms	Number publications	
COVID-19	+ genes	28
	+ protein interaction	40
	+ protein receptor	85
SARS-CoV-2	+ gene	66
	+ interactome	4
	+ protein interaction	31
SARS-CoV	+ interactome	7
	+ nsp protein	30
Total		291

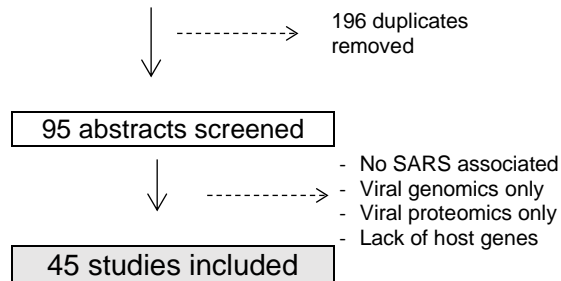


Figure 1: PubMed search between the 31st of March and 25th of May 2020. Several searching terms were entered in the database to identify all reported SARS-CoV host interactome reported genes.

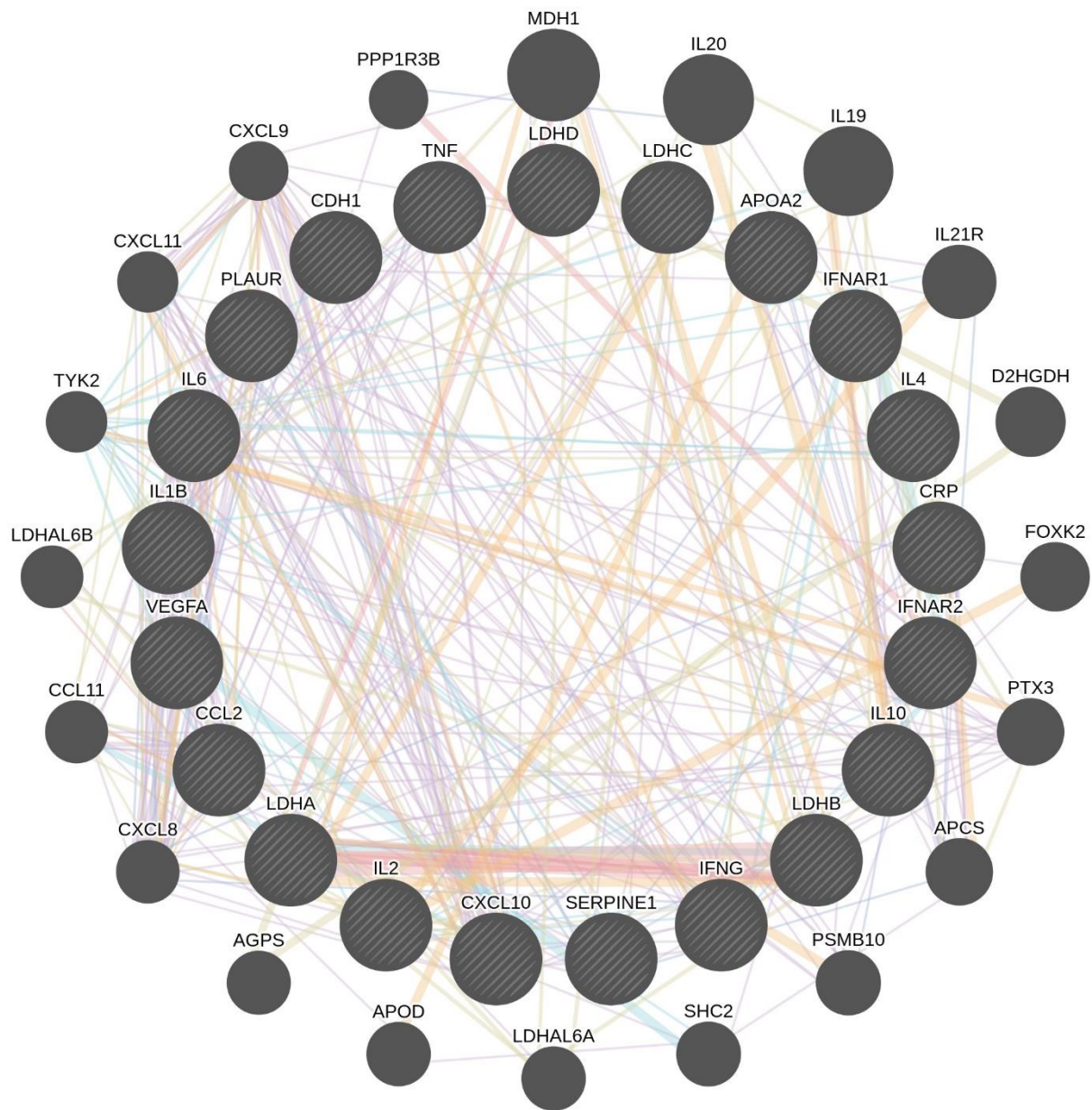


Figure 2: Protein-Protein Interaction map of all the 21 candidate genes associated with COVID-19 aggressiveness from our pilot study. The pin edge syndicate the amount of physical interactions that are well known from public literature with LDHA and LDHB as key players. Analysis run with GeneMANIA (<https://genemania.org/>). LDH: lactate dehydrogenase

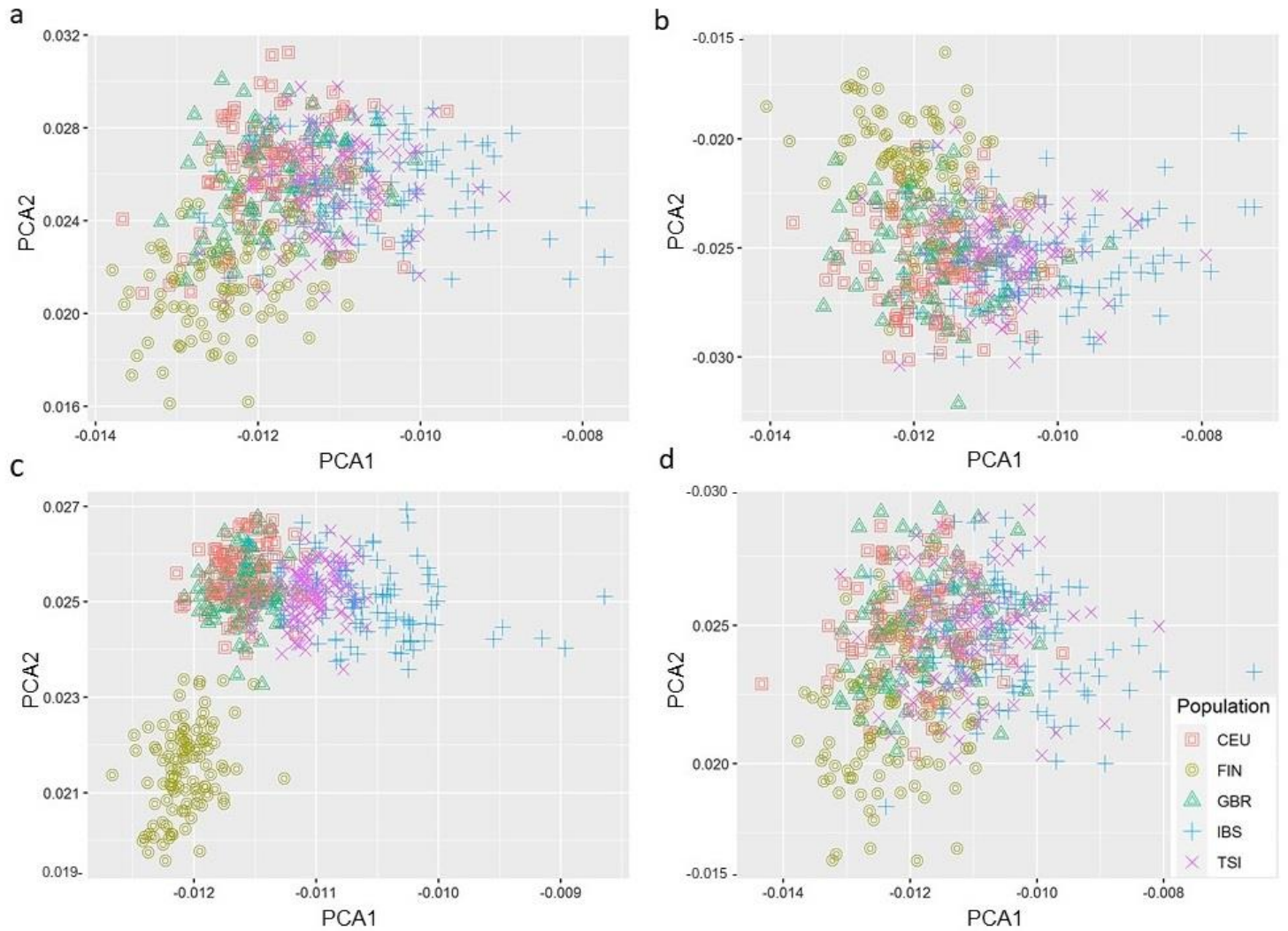


Figure 3: Principal component analysis from five European ancestries for the following canonical signalling pathways: a, Caveolar-mediated Endocytosis **b,** MSP-RON Signalling **c,** Hepatic Fibrosis and Hepatic Stellate Cell Activation **d,** Communication between Innate and Adaptive Immune. Population include, CEU: Utah residents with Northern and Western European Ancestry, FIN: Finnish population, GBR: Great Britain (England and Scotland), IBS: Iberian (Iberian Population in Spain), TSI: Tuscany (Tuscany in Italy). Genes obtained from pathway analysis (IPA, Qiagen) and genes locations (GRCh37/hg19 built) included a ± 1 Mb window. Genome data outsourced from 1,000G phase III v.5b. PC analysis. PC analysis performed using PLINK software v1.9b. Isoforms locations were flanked by ± 1 Mb.

Supplementary table 1: Candidate genes associated with viral entry and immune evasion reported in one study.

Candidate genes	Virus	Host cell response/viral mechanism	Research assay	Ref.
ADAM17	SARS-CoV-1	Competes with TMPRSS2 to cleave ACE2.	<i>In vitro</i> . Reporter gene assays, recombinant virus transduction, gene and protein expression analysis.	[34]
CD26 (or DPP4)	SARS-CoV-2	Interaction with S1 domain glycoprotein. May play a role in immunoregulation and virulence.	<i>In silico</i> structure and glycoprotein binding models.	[47]
DPP4 (or CD26)	MERS-CoV	After S interaction, it is activated by TMPRSS2 or cathepsin-L. Some isoforms associated with lower viral intake.	<i>In vitro</i> . MERS-CoV S cloned in human cell model. Co-IP, flow cytometry and viral titer quantification.	[48]
TNF-α	SARS-CoV-1	SARS S protein induces TACE- dependent cleavage of ACE2. Associated with early stage infections.	<i>In vitro</i> . Gene cloning and human cell viral infection, gene expression and ACE2 activity assays.	[33]
Furin	SARS-CoV-2 specific	Furin-like cleave site in the S2 domain associated with increased virulence and viral replication efficiency.	<i>In vitro</i> . SARS S protein genomic and proteomics analysis.	[37]
LRRK2, ACSL5, HSD17B4, EPHX1, MCCC2, GSTA4, ACACA, HGD, ROS1, CRIP2	ACE2 studies	Highest positive expression correlation score with ACE2 in lung ($P \leq 10^{-6}$). <i>CRIP2</i> , lowest negative correlation expression ($P = 5 \times 10^{-8}$).	<i>In silico</i> assay using RNA-seq lung normal tissue (TCGA), GEO and GO.	[98]
PIKfyve	SARS-CoV-2	Main enzyme to synthesize early/late endosome regulators. Assists SARS-CoV-2 pseudovirions <i>in vitro</i> entry.	<i>In vitro</i> . SARS-CoV-2 S protein pseudovirions expression system in human lung cell model.	[50]
TPC2	SARS-CoV-2	Downstream effector of main endosome formation regulator PI(3,5)P2. Assists SARS-CoV-2 pseudovirions <i>in vitro</i> entry.	<i>In vitro</i> . SARS-CoV-2 S protein pseudovirus expression system in human lung cell model.	[50]
GRP78	SARS-CoV-2	Encoded by <i>Hsp5a</i> and targeted by PeP42 in cancer cells. S COVID-19 protein model predicts similar binding site.	<i>In silico</i> structure and binding predictions of S proteins. Data sourced from HCoV (NCBI), EMBL-EBI.	[49]
CMTR1 (2'-O-MTase)	SARS-CoV	Predicted to methylate the viral mRNA and, escaping antiviral recognition.	<i>In silico</i> . Bioinformatics binding prediction studies.	[99]
IL-8	SARS-CoV-1	Increased production after S protein activation.	<i>In vitro</i> . Reporter gene assays, recombinant virus transduction in lung epithelial model. EMSA, gene and protein expression assays.	[57]
ubiquitin-aldehyde	SARS-CoV-2	Interaction with Nsp3 PLP binding site domain. This highly conserved binding site suggests a vital role in SARS virus.	<i>In silico</i> : NCBI Virus repository, MODELLER software, Protein Data Bank, Y2H and literature mining.	[66]
NLRP3	SARS-CoV-1	Activated after SARS-CoV-1 E protein calcium ion channel formation. It stimulates inflammasome by overproduction of cytokines.	<i>In vitro</i> . SARS-CoV-1 E protein synthetic peptide analysis.	[58]

60s	Nsp1 studies	After interacting with Nsp1, 60s is suppressed from joining 40s and the 80s translationally competent ribosome cannot translate.	<i>In vitro</i> . Expression, translation and rRNA immunoprecipitation and polysome assays. Ribosome binding assay.	[51]
STAT1	SARS-CoV-1	SARS-CoV-1 Nsp1 suppresses STAT1 phosphorylation, affecting IFN production.	<i>In vitro</i> . Reporter gene assays followed by immunoblot and northern blot analyses.	[59]
PPIA, PPIB, PPIH, PPIG, FKBP1A, FKBP1B	SARS-CoV-1	Human cyclophilins/immunophilin that modulate the Calcineurin/NFAT pathway after binding Nsp1. Possibly associated with viral pathogenesis.	<i>In vitro</i> . Reporter gene assays followed by SARS-CoV transduction and inhibition assays.	[100]
IRF3	SARS-CoV-1	SARS PLP proteases indirectly inhibits IRF3 phosphorylation and its nuclear translocation, suppressing IRF3-dependent IFN production and its pathway.	<i>In vitro</i> . Gene cloning, mRNA and protein expression analysis followed by immunoprecipitation.	[60]
TRAF3, IL-1B	SARS-CoV-1	After viral ORF3a is translated, it interacts with host's TRAF3, activating the NF-kB pathway and promoting IL-1B transcription.	<i>In vitro</i> . Gene cloning, cell transfection, reporter gene and gene expression analysis.	[61]
EDM1, OS9	SARS-CoV-1	Nsp3 and Nsp4 interactions initiate RTC synthesis.	<i>In vitro</i> . Gene cloning, recombinant BAC-SARS-Nsp3 expression and IP.	[65]
BTS-2	SARS-CoV-1	SARS ORF7a inhibits BTS-2 glycosylation, allowing for new virions to be released from host cell.	<i>In vitro</i> . Gene cloning, human cells transfection, IP and CD analysis.	[101]

Supplementary Table 2: Host cell pathways and protein complexes associated to viral entry and immune system evasion.

Human proteins/pathways	Virus	Host cell response	Detection method	Ref.
EEF1A1, RNF128*, MIF4GD, EIF4B, DCTN2, NDUFA10, ACE2	SARS-CoV-2	Most significantly up-regulated proteins	<i>In silico</i> analysis using COVID-19 patients' transcriptome data	[85]
MCL1	SARS-CoV-2	Most significantly down-regulated proteins	<i>In silico</i> analysis using COVID-19 patients' transcriptome data	[85]
CD74, MIF	ssRNA viruses	Most active ACE2 binding membrane receptors in 13 human tissues (present in macrophage)	Single cell <i>in silico</i> ligand-receptor affinity assays. Data sourced from GEO, Human Cell Atlas, Viral Receptor and Membranome databases°	[38]
Calcineurin/NFAT pathway	SARS-CoV-1, SARS-CoV-2	Activation after Nsp1-host proteins interactions	<i>In silico</i> . Virus-host proteins resourced from Y2H and literature data mining. <i>In vitro</i> . Reporter gene assays followed by SARS-CoV transduction and inhibition assays.	[66, 100]
Acetyl-CoA and leucine conversion/degradation and cholesterol biosynthesis	N/A	Enriched pathways after ACE2-activated networks in lungs, bronchial epithelium and oral mucosa tissues	<i>In silico</i> assay using RNA-seq lung (TCGA, controls), GEO and GO data	[98]
TRAF6-NF-kB	Coxsackievirus	Monocyte migration and potential myocarditis	<i>In vivo</i> assays in heart models after coronavirus inoculation	[68]
JUN, XPO1, NPM1, and HNRNPA1	SARS-CoV-1, MERS-CoV, HCoV-229E, and HCoV-NL63	Protein clusters with highest number of pathways activated	<i>In silico</i> analysis. Virus-host proteins resourced from Y2H, and <i>in vitro</i> Co-IP and RNA knock down experiment	[102]
MTOR	IBV and M-CoV	Autophagosome inactivation and viral degradation halt	<i>In vitro</i> human coronavirus models.	[103]
MAPKs and AP-1 pathways	SARS-CoV-1	Activated in lung cell models after IL-8 S viral protein activation	<i>In vitro</i> . Reporter gene assays, recombinant virus transduction followed by gene, protein and EMSA assays	[57]
STING-TRAF3-TBK1 And STING-MAVS-TBK1/IKKε complexes	SARS-CoV-1	SARS-CoV PLP disrupts both complexes, suppressing IRF3 activation and IFN antiviral signalling	<i>In vitro</i> . SARS-CoV gene cloning and lung cell model. Protein and Co-IP analysis and dimer detection	[55, 56]
NLRP3, ASC and caspase 1 complex	SARS-CoV-1	Inflammasome complex recruited after viral ORF3a and host TRAF3 interaction	<i>In vitro</i> . Gene cloning, BAC-ORF3a expression system, CRISPR, gene expression	[61]