Clinical and molecular characterization of COVID-19 hospitalized patients

Elisa Benetti[3], Annarita Giliberti[1], Arianna Emiliozzi[3,5], Floriana Valentino[1], Laura Bergantini[6], Chiara Fallerini[1], Federico Anedda[7], Sara Amitrano[2], Edoardo Conticini[9], Rossella Tita[2], Miriana d'Alessandro[6], Francesca Fava[1,2], Simona Marcantonio[7], Margherita Baldassarri[1], Mirella Bruttini[1,2], Maria Antonietta Mazzei[8], Francesca Montagnani[3,5], Marco Mandalà[4], Elena Bargagli[6], Simone Furini[3], COVID-19 MULTICENTER STUDY, Alessandra Renieri[1,2*] and Francesca Mari[1,2]

1) Medical Genetics, University of Siena, Italy
2) Genetica Medica, Azienda Ospedaliera Universitaria Senese, Italy
3) Department of Medical Biotechnologies, University of Siena, Italy
4) Otolaryngology Unit, University of Siena, Italy
5) Department of Specialized and Internal Medicine, Tropical and Infectious Diseases Unit, Azienda Ospedaliera Universitaria Senese, Italy
6) Unit of Respiratory Diseases and Lung Transplantation, Department of Internal and Specialist Medicine, University of Siena
7) Department of Emergency and Urgency, Medicine, Surgery and Neurosciences, Unit of Intensive Care Medicine, Siena University Hospital, Italy
8) Department of Medical, Surgical and Neuro Sciences and Radiological Sciences, Unit of Diagnostic Imaging, University of Siena, Azienda Ospedaliera Universitaria Senese, Italy
9) Rheumatology Unit, Department of Medicine, Surgery and Neurosciences, University of Siena, Policlinico Le Scotte, Italy

*** Corresponding author:**

Professor Alessandra Renieri
Medical Genetics Unit
University of Siena
Policlinico "Santa Maria alle Scotte"
Viale Bracci, 2 -53100 Siena, Italy
Phone: 39 0577 233303 - FAX 39 0577 233325
Email: alessandra.renieri@unisi.it

ABSTRACT

Clinical and molecular characterization by Whole Exome Sequencing (WES) is reported in 35 COVID-19 patients attending the University Hospital in Siena, Italy, from April 7 to May 7, 2020. Eighty percent of patients required respiratory assistance, half of them being on

mechanical ventilation. Fiftyone percent had hepatic involvement and hyposmia was ascertained in 3 patients. Searching for common genes by collapsing methods against 150 WES of controls of the Italian population failed to give straightforward statistically significant results with the exception of two genes. This result is not unexpected since we are facing the most challenging common disorder triggered by environmental factors with a strong underlying heritability (50%). The lesson learned from Autism-Spectrum-Disorders prompted us to re-analyse the cohort treating each patient as an independent case, following a Mendelian-like model. We identified for each patient an average of 2.5 pathogenic mutations involved in virus infection susceptibility and pinpointing to one or more rare disorder(s). To our knowledge, this is the first report on WES and COVID-19. Our results suggest a combined model for COVID-19 susceptibility with a number of common susceptibility genes which represent the favorite background in which additional host private mutations may determine disease progression.

MAIN TEXT

INTRODUCTION

Italy has been the first European Country experiencing the epidemic wave of SARS-CoV-2 infection, with an apparently more severe clinical picture, compared to other countries. Indeed, the case fatality rate has peaked to 14% in Italy, while it remains stable around 5% in China. (https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200512-covid-19-sitrep-113.pdf?sfvrsn=feac3b6d_2). Currently, SARS-CoV-2 positive subjects in Italy have reached the threshold of 200.000 cases, and we are now experiencing the long awaited descent of viral spread. Since the beginning of the epidemic wave, one of the first observations has been a highly heterogeneous phenotypic response to SARS-CoV-2 infection among individuals. Indeed, while most affected subjects show mild symptoms, a subset of patients develops severe pneumonia requiring mechanical ventilation with a 20% of cases requiring hospitalization; 5% of cases admitted to the Intensive Care Unit (ICU), and ~2.5% requiring intensive support with ventilators or extracorporeal oxygenation (ECMO) machines [1]. Although patients undergoing ventilatory assistance are often older and are affected by other diseases, like diabetes [2], the existing comorbidities alone do not fully explain the differences in clinical severity. A reasonable hypothesis is that at the basis of these different outcomes there are host predisposing genetic factors leading to different immunogenicity/cytokine responses as well as specific receptor permissiveness to virus and antiviral defence [3-6]. Similarly, during the study of host genetics in influenza disease, a pattern of genetic markers has been identified which underlies increased susceptibility to a more severe clinical outcome (as reviewed in [7]). This hypothesis is also supported by a recent work reporting 50% heritability of COVID-19 symptoms [8].

The identification of host genetic variants associated with disease severity is of

utmost importance to develop both effective treatments, based on a personalized approach, and novel diagnostics. Also, it is expected to be of high relevance in providing guidance for the health care systems and societal organizations. However, nowadays, little is known about the impact of host genome variability on COVID-19 susceptibility and severity.

On March 16th, 2020 the University Hospital in Siena launched a study named GEN-COVID with the aim to collect the genomic DNA of 2,000 COVID-19 patients for host genetic analysis. More than 30 different hospitals and community centers throughout Italy joined the study and are providing samples and clinical detailed information of COVID-19 patients. This study is aimed to identify common and rare genetic variants of SARS-CoV-2 infected individuals, using a whole exome sequencing (WES) analysis approach, in order to establish an association between host genetic variants and COVID-19 severity and prognosis.

RESULTS

Clinical data

The cohort consists of 35 COVID-19 patients (33 unrelated and 2 sisters) admitted to the University Hospital in Siena, Italy, from April 7 to May 7, 2020. All patients are of Caucasian ethnicity, except for one North African and one Hispanic. The mean and median age is 64 years (range 31-98): 11 females (median age 66 years) and 24 males (median age 62 years)**.**

The population is clustered into four qualitative severity groups depending on the respiratory impairment and the need for ventilation: high care intensity group (those requiring invasive ventilation), intermediate care intensity group (those requiring non invasive ventilation i.e. CPAP and BiPAP, and high-flows oxygen therapy), low care intensity group (those requiring conventional oxygen therapy) and very low care intensity group (those not requiring oxygen therapy) (groups 1-4 in **Table 1** and different colors in

Fig. 1). In the two most severe groups (groups 1 and 2, including 13 patients) there are 11 males and 2 females, while in the two mildest groups (groups 3 and 4 including 22 patients) males are 13 while females are 9.

Patients were also assigned a lung imaging grading according to X-Rays and CT scans. The mean value is 13 for high care intensity group, 12 for intermediate care intensity group, 8 for low care intensity group and 5 for very low care intensity group.

Regarding immunological findings, a decrease in the total number of peripheral $CD4^+$ T cells were identified in 13 subjects, while NK cells' count was impaired in 10 patients. Six patients showed a reduction of both parameters. IL-6 serum level was elevated in 13 patients. Based on blood groups, the cohort is divided into 15 patients of group 0, 16 patients of group A, 4 patients of group B and none of group AB.

Hyposmia was present in 3 out of 34 evaluated cases (8.8%), and hypogeusia was present in the same subjects plus another case. These four cases belong to the first three severity groups. Liver involvement was present in 7 cases (20%), while pancreas involvement in 4 cases (11%); 10 patients presented both (29%). Heart involvement was detected in 13 cases (37%). 9 patients (25%) showed kidney involvement. Fibrinogen values below 200 mg/dL were identified in 2 cases (6%), between 200 and 400 mg/dL in 7 cases (20%), and above 400mg/dL in 22 cases (63%). D-dimer value below 500 ng/mL was present in 1 case (3%), between 500 and 5000 ng/mL in 26 cases (74%), and in 7 cases (20%) was 10 times higher than the normal value (>5000 ng/mL) (**Table 1**).

Table 1. Clinical characteristics COVID19 patients admitted to the University Hospital of Siena (Italy)

Subject characteristics		**Group 1**	**Group 2**	**Group 3**	**Group 4**
No. of subjects		6	7	15	7
Median age (range) [a]		65 (55-70)	58 (31-74)	66 (49-98)	58 (31-74)
Gender					
	Male	5	6	7	6

Female	1	1	8	1
Blood Group				
A	2	3	6	5
B	2	1	1	0
0	2	3	8	2
PaO_2/FiO_2				
≤ 100	4	0	0	0
100-200	2	5	3	0
200-300	0	1	7	0
>300	0	0	4	5
Unknown	0	1	1	2
Lung imaging grading (CXR score)				
0-10 (mild)	1	2	12	7
11-20 (moderate)	5	5	3	0
21-28 (severe)	0	0	0	0
Laboratory findings				
$CD4^+$ T cells count				
<400 (cell/ul)	4	2	5	2
400-1500 (cell/ul)	2	5	8	5
Unknown	0	0	2	0
NK cells count				
<90 (cell/ul)	3	4	3	1
90-590 (cell/ul)	3	3	10	6
Unknown	0	0	2	0
IL-6 value				
<27 (pg/ml)	*0*	1	3	2
>27 (pg/ml)	*6*	4	2	1
Unknown	*0*	2	10	4
Fibrinogen				
<200 (mg/dL)	*2*	0	0	0
200-400 (mg/dL) normal value		2	3	1
>400 (mg/dL) hyperfibrinogenemia	*3*	3	12	4
Unknown	*0*	2	0	2
CRP				
<0,5 (mg/dl)	*0*	0	2	1
>0,5 (mg/dl)	*6*	7	13	6
LDH				
Normal value M: 135–225 (Ul/l); F: 135-214 (Ul/l)	1	1	2	2
High value	5	6	13	5
D-Dimer				
Less than 10 times higher than normal value	3	5	13	5
normal	0	0	0	1
More than 10 times higher than normal	3	2	2	0

value				
Unknown	0	0	0	1
Hyposmia (VAS score)				
<2 (normal)	4	6	14	7
2-5 (intermediate)	1	0	0	0
>5 (severe)	0	1	1	0
Hypogeusia (VAS score)				
No	4	6	13	7
Yes	1	1	2	0
Heart involvement				
Yes	4	3	6	0
T=T-Troponin >15 (ng/L); B= pro-BNP M > 88 (pg/ml); F > 153 (pg/ml); A=arrhythmia)	T/B 2 B 2	T/B 1 T 1 A 1	T/B 2 T/A 1 B/A 1 A 1 B 1	
No	2	4	9	7
Unknown	0	0	0	0
Hepatic (H)/Pancreatic involvement (P)				
H and P	2	5	6	1
H only	3	0	1	1
P only	0	0	1	1
None	1	2	7	4
Kidney involvement				
Yes	0	3	5	1
No	6	4	10	6
Co-morbidities				
Cardiovascular disease	1	2	3	
Hypertension	2	2	8	
Tumor	2	1	2	1
Diabetes			4	
Pulmonary disease			1	1

[a] Age in years. NA, not applicable.

Unbiased collapsing gene analysis

At first, we tested the hypothesis that susceptibility could be due to one or more common factor(s) in the cohort of patients compared to controls. According to this idea, damaging variants of that/those gene(s) should be either over- or under- represented in

patients vs controls. We used, as controls, individuals of the Italian population assuming that the majority of them, if infected, would have shown no severe symptoms. WES data of 35 patients were compared with those of 150 controls (corresponding to a sub-group of the Siena part of the Network of Italian Genomes NIG http://www.nig.cineca.it) using a gene burden test which compares the rate of disrupting mutations per gene. The variants were collapsed on a gene-by-gene basis, in order to identify genes with mutational burden statistically different between COVID-19 samples and controls. The analysis identified genes harboring deleterious mutations (according to the DANN score) with a statistically significant higher frequency in controls than in COVID-19 patients such as the olfactory receptor gene *OR4C5* (adjusted p-value of 1.5E-10) and the zinc finger transcription factor *ZNF717,* which regulates IL6 (adjusted p-value 9.2E-06, respectively) (**Fig. 2** and **Supplementary Table S1**) and *FAM104B* and *NDUFAF7,* although to a lesser extent (**Fig. 2** and **Supplementary Table S1**). For all these genes, the susceptibility factor is represented by the functioning (or more functioning) gene. We also identified two additional genes, *PRKRA* and *LAPTM4B*, for which the probability of observing a deleterious variant was computed higher in the COVID-19 samples compared to controls **(Fig. 2** and **Supplementary Table S2).** In these latter cases, the functioning gene represents indeed a protective factor.

Gene analysis using the Mendelian-like model

We then tested the hypothesis that COVID-19 susceptibility is due to different variants in different individuals. A recently acquired knowledge on the genetic bases of Autism Spectrum Disorders suggests that a common disorder could be the sum of many different rare disorders and this genetic landscape can appear indistinguishable at the clinical

level [9]. Therefore, we analyzed our cohort treating each patient as an independent case, following a Mendelian-like model. According to the "pathogenic" definition in ClinVar database (https://www.ncbi.nlm.nih.gov/clinvar/), for each patient, we identified an average of 1 mutated gene involved in viral infection susceptibility and pinpointing to one or more rare disorder(s) or a carrier status of rare disorders **(Fig. 1)**. Following the pipeline used in routine clinical practice for WES analysis in rare disorders we then moved forward checking for rare variants "predicted" to be relevant for infection by the means of common annotation tools. We thus identified an average of additional 1-5 variants per patient which summed up to the previous identified pathogenic variants **(Fig. 1**, **Supplementary Table S3)**.

Known common susceptibility/protective variants analysis

We then checked the cohort for known non rare variants classified as either "pathogenic" or "protective" in ClinVar database and related to viral infection. Variants in six different genes matched the term of "viral infection" and "pathogenic" according to ClinVar **(Fig. 1)**. Overall, a mean of 3 genes with "pathogenic" common variants involved in viral infection susceptibility were present **(Fig. 1)**.

Among the common protective variants, we list as example three variants which confer protection to Human Immunodeficiency Virus (HIV), the first two, and leprosy, the third one: a *CCR2* variant (rs1799864) identified in 8 patients, a *CCR5* (rs1800940) in one patient and a *TLR1* variant (rs5743618) in 26 patients (not shown). A *IL4R* variant (rs1805015) associated with HIV slow progression was present in 8 patients (not shown).

Candidate gene overview

Although not identified by unbiased collapsing gene analysis a number of obvious candidate genes were specifically analyzed. First of all, we noticed that SARS-CoV-2 receptor, ACE2 protein is preserved in the cohort, only a silent mutation V749V being

present in 2 males and 2 heterozygous females. This is in line with our previous suggestion that either rare variants or polymorphisms may impact infectivity [10] The IFITM3 polymorphism (rs12252) was found in heterozygosity in 4 patients as expected by frequency. Eight patients had heterozygous missense mutations in CFTR gene reported as VUS/mild variants, 7 / 8 being among the more severely affected patients.

DISCUSSION

In this study, we present a cohort of 35 COVID-19 patients admitted between April and May 2020 to the University Hospital of Siena who were clinically characterized by a team of 29 MDs belonging to 7 different specialties. As expected, the majority of hospitalized patients are males, confirming previously published data reporting a predominance of males among the most severe COVID-19 affected patients [11]. The distribution of blood types in our patients is not statistically different from that of the general population according to a chi-square test with alpha equal to 0.05 (data not shown). Lung imaging involvement, evaluated through a modified lung imaging grading system [12], did not completely correlate with respiratory impairment since among the 13 patients who required mechanical ventilation (group 1 and 2), grading was either moderate (10) or mild (3). In line with our previous data, lymphocyte subset immunophenotyping revealed a decrease in the total number of CD4 and NK cells count, especially in the most severe patients [13]. Laboratory tests revealed a multiple-organ involvement, confirming that COVID-19 is a systemic disease rather than just a lung disorder (**Fig. 1**). We thus propose that only a detailed clinical characterization can allow to disentangle the complex relationship between genes and signs/symptoms.

In order to test the hypothesis that the COVID-19 susceptibility is due to one or more genes in common among patients, we used the gene burden test to compare the rate of disrupting mutations per gene. This test has already been successfully applied to discover

susceptibility genes for Respiratory Syncytial Virus infection [14]. We identified 2 pretty robust genes whose damage represents a protective factor: *OR4C5* and *ZNF717*.

OR4C5 is a "resurrected" pseudogene, known to be non functioning in half of the European population. In our genome, in addition to 413 Olfactory Receptor (OR) loci, there are 244 segregating pseudogenes, 26 of which are "resurrected" from a pseudogene status. The *OR4C5* locus belongs to this subgroup and it has the highest intrapopulation variability, with a frequency of inactive allele of 0.62 in Asians, 0.48 in Europeans and 0.16 in Africans [15]. The locus has 68 different haplotypes and it presents an extensive range of copy-numbers across individuals [15,16]. Our data demonstrate an inverse correlation (p-value 2.13911369444695E-14 and Adj p-value 1.53930621452403E-10) between the *OR4C5* mutation burden and COVID-19 status suggesting that "resurrected" functional alleles of *OR4C5* are susceptibility alleles for COVID-19 infection.

Each sensory cell expresses a single allele of a single OR locus, thus transmitting a molecularly defined signal to the brain. It is known that the olfactory bulb is a critical immunosensory effector organ that effectively clears viruses. After the infection, the neuroepithelium triggers nitric oxide (NO) and major histocompatibility antigens (MHC) activation which starts innate immunity [17]. The entry of the virus into neuroepithelial cells induces the activation of microglia/macrophages which are involved in the phagocytic process, clinically corresponding to Hypo(an)osmia. This response is essential during the initial stages of infection and blocks the virus access to the CNS [18].

Expression of the "resurrected" pseudogene *OR4C5* may help in triggering the natural immunity leading to virus and cell death. It is interesting to note that protein atlas shows OR4C5 protein expression in the liver without the corresponding mRNA expression (www.proteinatlas.org). Usually, dissociation between protein and mRNA means that the protein is produced elsewhere and then transported into the organ. This is the case, for instance, for neurotransmitters that are synthetized in neuronal cell bodies located in

abdominal ganglia and are subsequently transported to the liver through the axons innervating the organ [19]. Thus, we may speculate that OR4C5 reaches the liver through nerve terminals. If this is the case, those individuals expressing the resurrected *OR4C5* gene may have more triggers of innate immunity and subsequently higher organ damage. It is noteworthy that in our cohort the putative expression of *OR4C5* (white boxes) is identified in patients with liver damage (**Fig. 1**).

Previous studies reported a prevalence of olfactory disorders in COVID-19 population ranging from 5% to 98%. A recent meta-analysis of 10 studies demonstrated a 52.73% prevalence for smell dysfunction in COVID-19 subjects [20]. In our population, only 3/35 (8.6%) subjects reported olfactory disorders. Both the limited sample size and the characteristic of the population (severely affected hospitalized subjects) could explain this result. However, a report focusing on smell dysfunction in severely affected hospitalized subjects reported a prevalence of 23.7% among 59 patients [21].

Kruppel-associated box zinc-finger protein 717 (*ZNF717*) belongs to a large group of transcriptional regulators playing important roles in different cellular processes, including cell proliferation, differentiation and apoptosis, and in the regulation of viral replication and transcription. *ZNF717* variations were detected in more than 10% of the WGS samples in Hepatitis B Virus (HBV)-related hepatocellular carcinoma (HCC) [22] where it has been identified as a potential driver gene with high frequency mutations at both single-cell and population levels. Moreover, this gene is one of the most recurrent somatically mutated genes in gastric cancer, together with *TP53* [23]. From a functional standpoint it acts through the regulation of the IL-6/STAT3 pathway. *ZNF717* knockdown in HCC cell lines results in increased levels of IL-6 and upregulation of STAT3 and its target genes [24]. According to our results those people who have this gene more damaged are more protected from SARS-CoV-2 infection, likely because they have a smarter innate immunity.

PRKRA (protein kinase activator A, also known as PACT; OMIM# *603424) is a

protein kinase activated by double-stranded RNA which mediates the effects of interferon in response to viral infection, resulting in antiviral activity [25]. Multiple transcript variants have been identified, including a polymorphism removing the canonical ATG. Mutations in the *PRKRA* gene have been associated with autosomal recessive dystonia [26]. The innate immune response is activated by the detection of viral structures, potentially via dsRNA-binding partners such as PRKRA [27]. In line with PRKRA antiviral activity, in our cohort we found a higher burden of potentially deleterious variants in COVID-19 patients, suggesting that these variants might reduce PRKRA functionality, potentially impairing IFN-mediated immune response.

Fourteen percent of patients bear highly disrupting mutations in *LAPTM4B* (Lysosomal Protein Transmembrane 4 Beta) gene, which was selected using burden gene test (P-val 4.04625E-06 and adj P-val 0.02911682). LAPTM4B protein is involved in the endosomal network, which eventually enables productive viral infection [28]. In particular, in HBV infection, endocytosis of EGF Receptor (EGFR) drives the translocation of HBV particles from the cell surface to the endosomal network, enabling productive viral infection [28]. In this process, LAPTM4B down regulates the formation of late lysosomes, suppressing EGFR lysosomal degradation and thus leading to a prolonged permanence of EGFR on the cell surface [29]. Accordingly, *LAPTM4B* knockdown significantly promotes HBV infection [28]. This perfectly fits with our results, indicating a significantly increased probability of deleterious changes in COVID-19 patients compared to controls.

Being affected by a rare disorder and/or being a carrier of rare disorders may represent a susceptibility factor to infections (**Fig. 1**). Having this in mind and driven by the lesson learned from the studies on the genetics bases of Autism Spectrum Disorders, we explored the possibility that each patient could have one or a unique (personalized approach) combination of rare pathogenic or highly relevant variants related for different reasons to infection susceptibility [9]. For instance, one male patient is affected by Glucose-6-

phosphate dehydrogenase (G6PD)-deficiency (rs137852318), the most common enzymopathy in humans, affecting 400 million people worldwide. G6PD-deficient cells are more susceptible to several viruses including coronavirus and in G6PD-deficient cells, innate immunity is down regulated, in line with the observed very low levels of IL-6 in this patient (**Fig. 1**) [30]. The same patient also presents a *ZEB1*-linked corneal dystrophy. *ZEB1* gene is known to function in immune cells, playing an important role in establishing both the effector response and future immunity in response to pathogens [31].

In addition, two sisters have *TGFBI* mutations, associated with corneal dystrophy, several patients are carriers of pseudoxanthoma elasticum bearing *ABCC6* gene mutations, others have likely hypomorphic mutations in *CHD7* or *COL5A1/2* variants. All these genes play a role as modulators of immune cells activity and/or response to infections [32-39].

Other rare variants were identified in the following interesting genes: *ADAR*, involved in viral RNA editing; *CLEC4M*, an alternative receptor for SARS-CoV [40]; *HCRTR1/2*, receptors of Hypocretin, important in the regulation of fatigue during infections [41]; *FURIN*, a serine protease that cleaves the SARS-Cov-2 minor capsid protein important for ACE2 contact and viral entry into the host cells [42,43].

Additional interesting variants have been identified in *NOS3 and OPRM1*. COVID-19 aggravates NitricOxide (NO) production deficit in patients with *NOS3* polymorphisms. Management of eNOS/iNOS ratio (endothelial/inducible NO synthase) and NO level can prevent development of severe acute respiratory distress syndrome [44]. From an immunological point of view, NO is mainly produced through the iNOS, which can be selectively expressed both in epithelial and white blood cells. It is postulated that NO plays a crucial role in innate and specific host defense, particularly against protozoa and bacteria. However, its role in viral infection is debated, as conflicting results have been reported in literature. Even though iNOS is generally overexpressed in patients with active viral infection, experiments conducted in murine models show that the inhibition of iNOS leads

to a significant improvement of HSV-viral pneumonia, despite the impairment of viral clearance [45**]**. Moreover, NO production may facilitate virus mutations and selection of more resistant strains. However, it seems that NO may have different effects according to causative agents. Notably, few studies demonstrated that NO was able to significantly reduce viral infection and replication of SARS-CoV through two distinct mechanisms: impairment of the fusion between the spike protein and its receptor ACE2, and reduction of viral RNA production [46]. Very few data are currently available on NO specific effects in COVID-19. The promising results reported in SARS-CoV infection may suggest a similar effectiveness also in COVID-19, considering that SARS-CoV and SARS-CoV2 share more than 70% of RNA sequence and have shown similar mechanisms of infection and viral replication. Clinical trials are ongoing to evaluate the effectiveness of inhaled NO in COVID-19 patients: although inhaled NO is conceived as a vasodilator therapy and therefore is indicated for the optimization of ventilation/perfusion mismatch in intubated patients, the results may be helpful to elucidate its potential antiviral properties [47,48**].**

Opioid ligands may regulate the expression of chemokines and chemokine receptors [49]. Due to immunomodulatory effect of morphine, *OPRM1* has been supposed to be involved in immune response and in HIV expansion [49,50]. Proudnikov et al. suggested that Opioid receptor *OPRM1* variants may affect the pathophysiology of HIV infection in the response to HIV treatment. They found an association with clinical improvement in subjects bearing alleles IVS1+1959A, IVS1+14123A, and IVS2+31A. Association of the same variants with HIV status was also reported [50].

Several rare variants in Interleukins (*ILs*) and Interleukins receptors (*ILRs*) are found. Interleukins are crucial in modulating immune response against all types of infective agents. The variants reported in this study include different interleukins that are not specifically involved in the defense against virus but are critical in balancing both innate and specific adaptive immune response. IL-2 and IL-15 are essential for proliferation and

activation of CD8+ effector T-cells, while IL-13 is fundamental, together with IL-4, in inducing and maintaining a type-2 inflammation in airways, causing mucus hypersecretion and favouring potential fibrogenetic processes of the bronchial wall. On the other hand, IL-12 promotes Th1 polarization of inflammatory processes in airways and has an important role in the defense against intracellular pathogens and mycobacteria. Evidence concerning the role of IL-12 against viruses are scarce and conflicting: however, IL-12 is reported also to enhance NK cell cytotoxic properties and therefore should be helpful in innate response against viral infection. IL16 is a pleiotropic cytokine chemotactic for T cells and able to modulate lymphocyte activation. IL-16 in HIV infection seems to inhibit virus replication [51]. On the other hand, in mice infected with influenza A virus, an IL-16 deficiency was associated with increased Th1 and cytotoxic lymphocyte responses.

We also identified common "pathogenic" variants in genes known to be linked to viral infection, such as *MBL2*, *IRGM* and *SAA1*, and/or specific organ damage as *PRSS1*.

PRSS1 encodes for a serine protease secreted from the pancreas. Mutations in this gene, including those identified in our cohort, are associated with autosomal dominant hereditary pancreatitis (OMIM#167800) [52]. *MBL2* encodes a mannose-binding lectin (MBL) secreted by the liver as part of the acute-phase response and involved in innate immune defense. Its deficiency causes increased susceptibility to infections, possibly due to a negative impact on the ability to mount an immune response [53,54]. MBL2 deficiency has also been suggested to be a susceptibility factor for vascular disease [55]. *IRGM* plays a role in autophagy and control of intracellular mycobacteria [56]. Autophagy plays relevant roles in many processes, including innate and adaptive immunity and antimicrobial defense. Accordingly, alterations in *IRGM* regulation affect the efficacy of autophagy and are considered relevant contributing factors in chronic inflammatory diseases, including Inflammatory Bowel Disease (IBD; OMIM#612278) [57]. *SAA1*, encoding the serum amyloid A (SAA) protein, is an apolipoprotein reactant, mainly produced by hepatocytes

and regulated from inflammatory cytokines. In patients with chronic inflammatory diseases, the SAA cleavage product, Amyloid protein A (AA), is deposited systemically in vital organs including liver, spleen and kidneys, causing amyloidosis [58].

For the last above reported genes and pathogenic variants or predicted variants relevant for infection, a statistically significant difference in variant's frequency was not found between cases and controls looking at either the single variant or the single gene, as a burden effect of variants. However, as depicted in the overall **Fig. 1,** we could hypothesize a combined model in which common susceptibility genes will sum to less common or private susceptibility variants. A specific combination of these 2 categories may determine type (organotropism) and severity of the disease.

Our observations related to the huge amount of data, both on phenome and genome sides, and represented in Figure 1, could also lay the bases for association rule mining approaches. Artificial intelligence techniques based on pattern recognition may discover an intelligible picture which appears blurred at present.

Further analyses in larger cohort of cases are mandatory in order to test this hypothesis of a combined model for COVID-19 susceptibility with a number of common susceptibility genes which represent the fertile background in which additional private, rare or low frequency mutations confer to the host the most favourable environment for virus growth and organ damage.

METHODS

Patients clinical data and Samples collection.

Thirty-five patients admitted to the University Hospital in Siena, Italy, from April 7 to May 7, 2020 were recruited. The study was consistent with Institutional guidelines and approved by the University Hospital (Azienda Ospedaliera Universitaria Senese) Ethical Review Board, Siena, Italy (Prot n. 16929, dated March 16, 2020). Written informed consent was

obtained from all patients. Peripheral blood samples in EDTA-containing tubes and detailed clinical data were collected. All these data were inserted in a section dedicated to COVID-19 of the established and certified Biobank and Registry of the Medical Genetics Unit of the Hospital. An example of the Clinical questionnaire is illustrated in **Supplementary Fig. S1**.
Each patient was assigned a continuous quantitative respiratory score, the PaO2/FiO2 ratio (normal values >300) (P/F), as the worst value during the hospitalization.
Patients were also assigned a lung imaging grading according to X-Rays and CT scans. In particular, lung involvement was scored through imaging at the time of admission and during hospitalization (worst score), annotating the chest X-Ray (CXR) score (in 34 patients) and CT score in 1 patient for whom X-Rays were not available. To obtain the score (from 0 to 28) each CXR was divided in four quadrant (right upper, right lower, left upper and left lower) and for each quadrant the presence of consolidation (0= no consolidation; 1 <50%, 2>50%), ground glass opacities (GGOs: 0= no GGOs, 1<50%, 2 >50%), reticulation (0= no GGOs, 1<50%, 2 >50%) and pleural effusion on left or right side (0= no, 1= minimal; 2= large) were recorded. The same score was applied for CT (1 patient).
For each patient, the presence of hyposmia and hypogeusia was also investigated through otolaryngology examination, Burghart sniffin' sticks [59] and a visual analog scale (VAS). Whenever the sign was present, a score ranging from 0 to 10 was assigned to each patient using VAS where 0 means the best sense of smell and 10 represents the absence of smell sensation [60].
The presence of hepatic involvement was defined on the basis of a clear hepatic enzymes elevation as glutamic pyruvic transaminase (ALT) and glutamic oxaloacetic transaminase (AST) both higher than 40 UI/L. Pancreatic involvement was considered on the basis of an increase of pancreatic enzymes as pancreatic amylase higher than 53 UI/l and lipase higher than 60UI/l. Heart involvement was defined on the basis of one or more of the following abnormal data: Troponin T (indicative of ischemic disorder), NT-proBNP (indicative of heart failure) and arrhythmias (indicative of elettric disorder). Kidney involvement was defined in the presence of a creatinine value higher than 1,20 mg/dl in males and higher than 1,10 mg/dl in females.

Whole Exome Sequencing analysis

Genomic DNA was extracted from peripheral blood using the MagCore®Genomic DNA Whole Blood kit (RBC Biosciences) according to manufacturer's protocol. Whole exome sequencing analysis was performed on Illumina NovaSeq 6000 system (Illumina, San

Diego, CA, USA). DNA fragments were hybridized and captured by Illumina Exome Panel (Illumina) according to manufacturer's protocol. The libraries were tested for enrichment by qPCR, and the size distribution and concentration were determined using an Agilent Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA, USA). The Novaseq 6000 platform (Illumina), along with 150 bp paired-end reads, was used for sequencing of DNA.

Genetic data analysis

Reads were mapped to the hg19 reference genome by the Burrow-Wheeler aligner BWA [61]. Variants calling was performed according to the GATK4 best practice guidelines [62]. Namely, duplicates were first removed by *MarkDuplicates*, and base qualities were recalibrated using *BaseRecalibration* and *ApplyBQSR*. *HaplotypeCaller* was used to calculate Genomic VCF files for each sample, which were then used for multi-sample calling by *GenomicDBImport* and *GenotypeGVCF*. In order to improve the specificity-sensitivity balance, variants quality scores were calculated by *VariantRecalibrator* and *ApplyVQSR*. Variants were annotated by ANNOVAR [63], and with the number of articles answering the query “gene_name AND viral infection” in Pubmed, where gene_name is the name of the gene affected by the variant.

In order to identify candidate genes according to the Mendelian-like model, rare variants were filtered by a prioritization approach. We used the ExAC database (http://exac.broadinstitute.org/), in particular the ExAC_NFE reported frequency to filter variants according to a minor allele frequency < 0.01. Synonymous, intronic and non-coding variants were excluded from the analysis. Mutation disease database ClinVar (ncbi.nlm.nih.gov/clinvar/) was used to identify previous pathogenicity classifications and variants reported as likely benign/benign were discarded. Filtering and prioritization of variants was completed using the CADD_Phred pathogenicity prediction tool. Finally, we selected genes involved in infection susceptibility using the term “viral infection” as Pubmed database search.

In order to identify genes with a different prevalence of functionally relevant variants between COVID-19 patients and control samples, the following score was calculated:

$$\square_{\square} = \sum_{\square=1}^{\square} \square_{\square}\square_{\square,\square} \qquad (1),$$

Where $\square_{\square}$ is a weight associated with the *i-th* variant; and $\square_{\square,\square}$ is equal to 0 if the variant is not present in sample *j*, 1 if sample *j* has the variant in heterozygous state, and 2 if sample *j* has the variant is homozygous state. The weight $\square_{\square}$ was assumed equal to the DANN score of the variant [64], which provides an estimate of the likelihood that the variant has deleterious functional effects (i.e. variants more likely to have a functional effect contribute more to the score). The sum in equation (1) was performed over all the variants in the gene where the DANN score was available. Genes with less than 5 annotated variants were discarded from the analysis. The scores calculated by equation (1) were ranked for all the samples, and the sum of the ranking for the COVID-19 samples, named $\square_{\square\square\square\square\square}$, was calculated. Then, sample labels were permuted 10.000 times, and these permutations were used to estimate the average value and the standard deviation of $\square_{\square\square\square\square\square}$ under the null-hypothesis. The p-value was calculated assuming a normal distribution for the sum of the ranking [65].

DATA AVAILABILITY

Data about the gene-based analyses and variants are available as Supplementary Material. The results of variant calling are available as aggregated data in the Network for Italian Genomes database (http://www.nig.cineca.it). The datasets generated during the current study are available from the corresponding author on reasonable request.

REFERENCES

[1] Guan, W. J., *et al.* Clinical Characteristics of Coronavirus Disease 2019 in China, *N Engl J Med*. **382**, 1708-1720 (2020).

[2] Wu, Z. & McGoogan, J. M. Characteristics of and Important Lessons From the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72 314 Cases From the Chinese Center for Disease Control and Prevention, *JAMA*. **323**, 1239-1242 (2020).

[3] Liu, R., *et al*. Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection. *Cell*. **86**, 367–377 (1996).

[4] Blumenthal, K. G., *et al*. Risk of meticillin resistant Staphylococcus aureus and Clostridium difficile in patients with a documented penicillin allergy: population based matched cohort study. *BMJ*. **361**, k2400 (2018).

[5] Woziwodzka A., *et al*. TNF-α polymorphisms affect persistence and progression of HBV infection. *Mol. Genet. Genomic Med.* **7**, e00935 (2019).

[6] Tian, T., *et al*. CD40 polymorphisms were associated with HCV infection susceptibility among Chinese population. *BMC Infect. Dis*. **19**, 840 (2019).

[7] Nogales, A., & L DeDiego, M. Host Single Nucleotide Polymorphisms Modulating Influenza A Virus Disease in Humans. *Pathogens*. **8**, E168 (2019).

[8] Williams, F.M.K., *et al*. Self-reported symptoms of covid-19 including symptoms most predictive of SARS-CoV-2 infection, are heritable. *medRxiv* DOI**:** https://doi.org/10.1101/2020.04.22.20072124 (2020).

[9] Satterstrom, F.K., *et al*. Large-scale exome sequencing study implicates both developmental and functional changes in the neurobiology of autism. *Cell*, **180**.3: 568-584. e23 (2020).

[10] Benetti, E. *et al*. ACE2 gene variants underlie interindividual variability and susceptibility to COVID-19 in Italian population. *MedRxiv* https://doi.org/10.1101/2020.04.03.20047977 (2020)

[11] Cai, H. Sex difference and smoking predisposition in patients with COVID-19. *The Lancet Respiratory Medicine*,**8**.4: e20 (2020).

[12] Borghesi, A. & Maroldi, R. COVID-19 outbreak in Italy: Experimental chest x-ray scoring system for quantifying and monitoring disease progression. *La Radiologia Medica*, **125** (5):509□513. doi:10.1007/s11547-020-01200-3 (2020).

[13] D'alessandro M.,*et al*. Peripheral lymphocyte subset monitoring in COVID19 patients: a prospective Italian real-life case series. *Minerva Med*.; DOI: 10.23736/S0026-4806.20.06638-0 (2020)

[14] Salas, A., *et al*. Whole exome sequencing reveals new candidate genes in host genomic susceptibility to respiratory syncytial virus disease. *Scientific reports*, **7**.1: 1-13 (2017).

[15] Olender, T., *et al*. Personal receptor repertoires: olfaction as a model. *BMC Genomics*. **13**, 414 (2012).

[16] Waszak, S. M., *et al*. Systematic inference of copy-number genotypes from personal genome sequencing data reveals extensive olfactory receptor gene content diversity. *PLoS Comput Biol*. **6**, e1000988 (2010).

[17] Durrant, D. M., Soumitra, G., & Robyn S.K. The Olfactory Bulb: An Immunosensory Effector Organ during Neurotropic Viral Infections. *ACS Chem Neurosci*. **7**, 464-9 (2016).

[18] Mori, I., *et al*. Olfactory receptor neurons prevent dissemination of neurovirulent influenza A virus into the brain by undergoing virus-induced apoptosis. *J Gen Virol*. **9**, 2109-2116 (2002)

[19] Streba, L.A. M., Vere, C.C., Lonescu, A.G., Streba, C.T. & Rogoveanu, I. Role of intrahepatic innervation in regulating the activity of liver cells. *World journal of hepatology*, **6**.3: 137 (2014).

[20] Tong, J.Y., Wong, A., Zhu, D., Fastenberg, J.H. & Tham, T. The Prevalence of Olfactory and Gustatory Dysfunction in COVID-19 Patients: A Systematic Review and Meta-analysis. *Otolaryngology–Head and Neck Surgery*, DOI: 0194599820926473 (2020).

[21] Giacomelli, A., *et al*. Self-reported olfactory and taste disorders in patients with severe acute respiratory coronavirus 2 infection: a cross-sectional study. *Clinical Infectious Diseases*, https://doi.org/10.1093/cid/ciaa330 (2020).

[22] Chen, Y., Wan, L., Xu, H., Liu, X. & Zhao,Y. Exome capture sequencing reveals new insights into hepatitis B virus-induced hepatocellular carcinoma at the early stage of tumorigenesis. *Oncology reports*, **30**.4: 1906-1912 (2013).

[23] Cui, J., *et al*. Comprehensive characterization of the genomic alterations in human gastric cancer. *International journal of cancer*, **137**.1: 86-95 (2015).

[24] Duan, M., *et al*. Diverse modes of clonal evolution in HBV-related hepatocellular carcinoma revealed by single-cell genome sequencing. *Cell research*, **28**.3: 359-373 (2018)

[25] Chan, C.P., *et al*. Antiviral activity of double-stranded RNA–binding protein PACT against influenza A virus mediated via suppression of viral RNA polymerase. *The FASEB Journal*, **32**.8: 4380-4393 (2018).

[26] Dos Santos, C.O., *et al*. The prevalence of PRKRA mutations in idiopathic dystonia. *Parkinsonism & related disorders*, **48**: 93-96 (2018).

[27] Miyamoto, M. & Komuro, A. PACT is required for MDA5-mediated immunoresponses triggered by Cardiovirus infection via interaction with LGP2. *Biochemical and biophysical research communications*, **494**.1-2: 227-233 (2017).

[28] Iwamoto, M., *et al*. Epidermal growth factor receptor is a host-entry cofactor triggering hepatitis B virus internalization. *Proceedings of the National Academy of Sciences*, **116**.17: 8487-8492 (2019).

[29] Tan, X., *et al*. LAPTM4B is a PtdIns (4, 5) P2 effector that regulates EGFR signaling, lysosomal sorting, and degradation. *The EMBO journal*, **34**.4: 475-490 (2015).

[30] Wu, Y. H., *et al*. Glucose-6-phosphate dehydrogenase deficiency enhances human coronavirus 229E infection. *The Journal of infectious diseases*, **197**.6: 812-816 (2008).

[31] Guan, T. *et al*. ZEB1, ZEB2, and the miR-200 family form a counterregulatory network to regulate CD8+ T cell fates. *Journal of Experimental Medicine*. **215**.4: 1153-1168 (2018).

[32] Klamer, S. E., *et al*. TGFBI Expressed by Bone Marrow Niche Cells and Hematopoietic Stem and Progenitor Cells Regulates Hematopoiesis. *Stem cells and development.* **27**.21: 1494-1506 (2018).

[33] Ebersole, J.L.; Peyyala, R. & Gonzalez, O.A. Biofilm□induced profiles of immune response gene expression by oral epithelial cells. *Molecular oral microbiology*. **34** (1): 0.1111/omi.12251. DOI: 10.1111/omi.12251 (2019).

[34] Marton, J. *et al*. Cyclosporine a treatment inhibits Abcc6-dependent cardiac necrosis and calcification following coxsackievirus B3 infection in mice. *PloS one*. **10**.9 (2015).

[35] Janssen, N., *et al*. Mutation update on the CHD7 gene involved in CHARGE syndrome. *Hum Mutat*. **33**(8):1149□1160 (2012).

[36] Theodoropoulos, D.S. & Theodoropoulos, G.A. Immune deficiency and hearing loss in CHARGE Association. *Pediatrics*. **111**.3: 711-712 (2003).

[37] Gennery, A.R., et al. Mutations in CHD7 in patients with CHARGE syndrome cause T–B+ natural killer cell+ severe combined immune deficiency and may cause Omenn□like syndrome. *Clinical & Experimental Immunology*. **153**.1:75-80 (2008).

[38] Randall, V. *et al*. Great vessel development requires biallelic expression of Chd7 and Tbx1 in pharyngeal ectoderm in mice. *The Journal of clinical investigation.* **119**.11: 3301-3310 (2009).

[39] Zhetkenev, S. *et al*. Association of rs12722 COL5A1 with Pulmonary Tuberculosis infection: a preliminary case-control study in a Kazakhstani population. *medRxiv*. 19008995 DOI: 10.1101/19008995 (2019).

[40] Chan, V.S.F., *et al*. Homozygous L-SIGN (CLEC4M) plays a protective role in SARS coronavirus infection. *Nature genetics*, **38**.1: 38-46 (2006).

[41] Zhan, S., *et al*. Tumor necrosis factor-alpha regulates the Hypocretin system via mRNA degradation and ubiquitination. *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease*, **1812**.4: 565-571 (2011)

[42] Braun, E., *et al*. Guanylate-binding proteins 2 and 5 exert broad antiviral activity by inhibiting furin-mediated processing of viral envelope proteins. *Cell reports*, **27**.7: 2092-2104. e10 (2019).

[43] Shang, J. *et al*. Cell entry mechanisms of SARS-CoV-2. Proc Natl Acad Sci U S A. 202003138. DOI: 10.1073/pnas.2003138117 (2020)

[44] Buinitskaya, Y., *et al*. Highlights of COVID-19 pathogenesis. Insights into Oxidative Damage. DOI: 10.6084/m9.figshare.12121575 (2020).

[45] Adler, H., *et al*.Suppression of herpes simplex virus type 1 (HSV-1)-induced pneumonia in mice by inhibition of inducible nitric oxide synthase (iNOS, NOS2). *J Exp Med.* **185**(9):1533-40 (1997).

[46] Åkerström, S., Gunalan, V., Keng, C.T., Yan, Y.J. & Mirazimi, A. Dual effect of nitric oxide on SARS-CoV replication: viral RNA production and palmitoylation of the S protein are affected. *Virology*, **395**.1: 1-9 (2009).

[47] Akerström S., *et al*. Nitric oxide inhibits the replication cycle of severe acute respiratory syndrome coronavirus. *J Virol*.**79**(3):1966-9 (2005)

[48] Zamanian R.T., *et al*.Outpatient Inhaled Nitric Oxide in a Patient with Vasoreactive IPAH and COVID-19 Infection. *Am J Respir Crit Care Med* DOI:10.1164/rccm.202004-0937LE (2020)

[49] Finley, M. J., Happel, C.M., Kaminsky, D.E. & Rogers, T.J. Opioid and nociceptin receptors regulate cytokine and cytokine receptor expression. *Cellular immunology*, **252**.1-2: 146-154 (2008).

[50] Sacerdote, P. Opioid-induced immunosuppression. *Curr Opin Support Palliat Care*. **2**, 14–8 (2008).

[51] Steele, A.D., Henderson, E.E. & Rogers, T.J. Mu-opioid modulation of HIV-1 coreceptor expression and HIV-1 replication. *Virology*. **309**, 99–107 (2003).

[52] Teich, N., *et al*. Gene conversion cetween functional trypsinogen genes *PRSS1* and *PRSS2* associated with chronic pancreatitis in a six□year□old girl. *Human Mutation*. **25**: 343-347 (2005).

[53] Thio, C. L., *et al*. Mannose binding lectin genotypes influence recovery from hepatitis B virus infection. *Journal of Virology*. **79**: 9192-9196 (2005).

[54] Dean, M. M., *et al*. Mannose-binding lectin deficiency influences innate and antigen-presenting functions of blood myeloid dendritic cells. *Immunology*. **132**(2):296□305 (2011).

[55] Madsen, H. O., Videm, V., Svejgaard, A., Svennevig, J. L. & Garred, P. Association of mannose-binding-lectin deficiency with severe atherosclerosis. *The Lancet* **352**: 959-960 (1998).

[56] Singh, S. B., Davis, A. S., Taylor, G. A. & Deretic, V. Human IRGM induces autophagy to eliminate intracellular mycobacteria. *Science* **313**: 1438-1441 (2006).

[57] Rufini, S., *et al* Autophagy and inflammatory bowel disease: Association between variants of the autophagy-related IRGM gene and susceptibility to Crohn's disease. *Digestive and Liver Disease*. **47**(9):744-50 (2015).

[58] Zhang, Y., Zhang, J., Sheng, H., Li, H. & Wang. R. Acute phase reactant serum amyloid A in inflammation and other diseases. *Adv Clin Chem*. **90**:25□80 (2019).

[59] Oleszkiewicz, A.,Schriever,V. A, Croy, I., Hähner, A. & Hummel, T. . Updated Sniffin'Sticks normative data based on an extended sample of 9139 subjects. *European Archives of Oto-Rhino-Laryngology*, **276**.3: 719-728 (2019).

[60] Klimek, L., *et al*. Visual analogue scales (VAS): Measuring instruments for the documentation of symptoms and therapy monitoring in cases of allergic rhinitis in everyday health care. *Allergo journal international*, **26**.1: 16-24 (2017).

[61] LI, H. and Durbin, R. Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics*, **26**.5: 589-595 (2010).

[62] R. Poplin *et al*.. Scaling accurate genetic variant discovery to tens of thousands of samples. *BioRxiv*, p. 201178, (2018).

[63] Wang, K., *et al*.“ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data,” *Nucleic Acids Res.*, vol. **38**, no. 16, pp. e164–e164, (2010).

[64] Quang, D., Chen,Y., & Xie, X. DANN□: a deep learning approach for annotating the pathogenicity of genetic variants. *Bioinformatics*, **31**.5: 761-763 (2015).

[65] Dering, C., Hemmelmann, C., Pugh, E. & Ziegler, A. Statistical Analysis of Rare Sequence Variants□: An Overview of Collapsing Methods. *Genetic epidemiology*. **35**.S1: S12-S17. (2011).

COVID-19 MULTICENTER STUDY (composition at May 22, 2020)

Alessandra Renieri[1,2], Elisa Benetti[3], Francesca Montagnani[3,5], Rossella Tita[2], Chiara Fallerini[1], Sara Amitrano[2], Mirella Bruttini[2], Gabriella Doddato[1], Annarita Giliberti[1], Floriana Valentino[1], Susanna Croci[1], Laura Di Sarno[1], Francesca Fava[1,2], Margherita Baldassarri[1], Andrea Tommasi[1,2], Sergio Daga[1], Maria Palmieri[1], Arianna Emiliozzi[3,5], Massimiliano Fabbiani[5], Barbara Rossetti[5], Giacomo Zanelli[3,5], Laura Bergantini[6], Miriana d’Alessandro[6], Paolo Cameli[6], David Bennett[6], Federico Anedda[7], Simona Marcantonio[7], Sabino Scolletta[7], Federico Franchi[7], Maria Antonietta Mazzei[8], Edoardo Conticini[9], Luca Cantarini[19], Bruno Frediani[20], Danilo Tacconi[10], Chiara Spertilli[10], Marco Feri[11], Alice Donati[11], Raffaele Scala[12], Luca Guidelli[12], Agostino Ognibene[13], Genni Spargi[14], Marta Corridi[14], Cesira Nencioni[15], Leonardo Croci[15], Gian Piero Caldarelli[16], Maurizio Spagnesi[17], Paolo Piacentini[17], Anna Canaccini[18], Agnese Verzuri[18], Valentina Anemoli[18], Massimo Vaghi[23], Antonella D’Arminio Monforte[24], Esther Merlini[24], Mario Umberto Mondelli[25,26], Stefania Mantovani[25], Serena Ludovisi[26], Massimo Girardis[27], Sophie Venturelli[27], Andrea Cossarizza[28], Andrea Antinori[29], Alessandra Vergori[29], Stefano Rusconi[30,31], Matteo Siano[30,31], Arianna Gabrieli[31], Daniela Francisci[32,33], Elisabetta Schiaroli[32], Pier Giorgio Scotton[34], Francesca Andretta[34], Sandro Panese[35], Renzo Scaggiante[36], Saverio Giuseppe Parisi[37], Francesco Castelli[38], Maria Eugenia Quiros Roldan[38],

Paola Magro[38], Cristina Minardi[38], Matteo Della Monica[39], Carmelo Piscopo[39], Mario Capasso[40,41,42], Massimo Carella[43], Marco Castori[43], Giuseppe Merla[43], Filippo Aucella[44], Pamela Raggi[45], Matteo Bassetti[46,47], Antonio Di Biagio[47], Maurizio Sanguinetti[48,49], Luca Masucci[48,49], Chiara Gabbi[21], Serafina Valente[22], Susanna Guerrini[8], Elisa Frullanti[1], Ilaria Meloni[1], Maria Antonietta Mencarelli[2], Caterina Lo Rizzo[2], Anna Maria Pinto[2], Elena Bargagli[6], Marco Mandalà[4], Simone Furini[3], Francesca Mari[1,2].

1) Medical Genetics, University of Siena, Italy
2) Genetica Medica, Azienda Ospedaliera Universitaria Senese, Italy
3) Department of Medical Biotechnologies, University of Siena, Italy
4) Otolaryngology Unit, University of Siena, Italy
5) Department of Specialized and Internal Medicine, Tropical and Infectious Diseases Unit, Azienda Ospedaliera Universitaria Senese, Italy
6) Unit of Respiratory Diseases and Lung Transplantation, Department of Internal and Specialist Medicine, University of Siena
7) Department of Emergency and Urgency, Medicine, Surgery and Neurosciences, Unit of Intensive Care Medicine, Siena University Hospital, Italy
8) Department of Medical, Surgical and Neuro Sciences and Radiological Sciences, Unit of Diagnostic Imaging, University of Siena, Azienda Ospedaliera Universitaria Senese, Italy
9) Rheumatology Unit, Department of Medicine, Surgery and Neurosciences, University of Siena, Policlinico Le Scotte, Italy
10) Department of Specialized and Internal Medicine, Infectious Diseases Unit, San Donato Hospital Arezzo, Italy
11) Department of Emergency, Anesthesia Unit, San Donato Hospital, Arezzo, Italy
12) Department of Specialized and Internal Medicine, Pneumology Unit and UTIP, San Donato Hospital, Arezzo, Italy
13) Clinical Chemical Analysis Laboratory, San Donato Hospital, Arezzo, Italy
14) Department of Emergency, Anesthesia Unit, Misericordia Hospital, Grosseto, Italy
15) Department of Specialized and Internal Medicine, Infectious Diseases Unit, Misericordia Hospital, Grosseto, Italy
16) Clinical Chemical Analysis Laboratory, Misericordia Hospital, Grosseto, Italy
17) Department of Prevention, Azienda USL Toscana Sud Est, Italy
18) Territorial Scientific Technician Department, Azienda USL Toscana Sud Est, Italy
19) Department of Medical Sciences, University of Siena, Italy
20) Research Center of Systemic Autoinflammatory Diseases and Behçet's Disease and Rheumatology-Ophthalmology Collaborative Uveitis Center, Department of Medical Sciences, Surgery and Neurosciences, University of Siena, Italy
21) Independent Scientist, Milan, Italy
22) Department of Cardiovascular Diseases, University of Siena, Italy
23) Chirurgia Vascolare, Ospedale Maggiore di Crema, Italy
24) Department of Health Sciences, Clinic of Infectious Diseases, ASST Santi Paolo e Carlo, University of Milan, Italy
25) Division of Infectious Diseases and Immunology, Department of Medical Sciences and Infectious Diseases, Pavia, Italy.
26) Department of Internal Medicine and Therapeutics, University of Pavia, Italy
27) Department of Anesthesia and Intensive Care, University of Modena and Reggio Emilia, Modena, Italy

28) Department of Medical and Surgical Sciences for Children and Adults, University of Modena and Reggio Emilia, Modena, Italy
29) HIV/AIDS Department, National Institute for Infectious Diseases, IRCCS, Lazzaro Spallanzani, Rome, Italy
30) III Infectious Diseases Unit, ASST-FBF-Sacco, Milan, Italy
31) Department of Biomedical and Clinical Sciences Luigi Sacco, University of Milan, Milan, Italy
32) Infectious Diseases Clinic, Department of Medicine 2, Azienda Ospedaliera di Perugia and University of Perugia, Santa Maria Hospital, Perugia, Italy
33) Infectious Diseases Clinic, "Santa Maria" Hospital, University of Perugia, Perugia, Italy
34) Department of Infectious Diseases, Treviso Hospital, Local Health Unit 2 Marca Trevigiana, Treviso, Italy
35) Infectious Diseases Department, Ospedale Civile "SS. Giovanni e Paolo" , Venice , Italy
36) Infectious Diseases Clinic, ULSS1, Belluno, Italy
37) Department of Molecular Medicine, University of Padova, Italy
38) Department of Infectious and Tropical Diseases, University of Brescia and ASST Spedali Civili Hospital, Brescia , Italy.
39) Medical Genetics and Laboratory of Medical Genetics Unit, A.O.R.N. "Antonio Cardarelli", Naples, Italy.
40) Department of Molecular Medicine and Medical Biotechnology, University of Naples Federico II, Naples, Italy.
41) CEINGE Biotecnologie Avanzate, Naples, Italy
42) IRCCS SDN, Naples, Italy.
43) Division of Medical Genetics, Fondazione IRCCS Casa Sollievo della Sofferenza Hospital, San Giovanni Rotondo, Italy.
44) Department of Nephrology and Dialysis, Fondazione IRCCS Casa Sollievo della Sofferenza Hospital, San Giovanni Rotondo, Italy.
45) Department of Medical Sciences, Fondazione IRCCS Casa Sollievo della Sofferenza Hospital, San Giovanni Rotondo, Italy.
46) Department of Health Sciences, University of Genova, Genova, Italy.
47) Infectious Diseases Clinic, Policlinico San Martino Hospital, IRCCS for Cancer Research Genova, Italy.
48) Microbiology, Fondazione Policlinico Universitario Agostino Gemelli IRCCS, Catholic University of Medicine, Rome, Italy.
49) Department of Laboratory Sciences and Infectious Diseases, Fondazione Policlinico Universitario A. Gemelli IRCCS, Rome, Italy.

ACKNOWLEDGEMENTS

This study is part of GEN-COVID, https://sites.google.com/dbm.unisi.it/gen-covid the Italian multicenter study aimed to identify the COVID-19 host genetic bases The *Genetic and COVID-19 Biobank of Siena*, member of BBMRI-IT, of Telethon Network of Genetic

Biobanks (project no. GTB18001), of EuroBioBank, and of D-Connect, provided us with specimens. We thank the CINECA consortium for providing computational resources and Network for Italian Genomes NIG http://www.nig.cineca.it. We thank private donors' support to A.R. (Department of Medical Biotechnologies, University of Siena) for the COVID-19 host genetics research project (D.L n.18 of March 17th 2020).

ETHICAL APPROVAL

The GEN-COVID study was approved by the University Hospital of Siena Ethical Review Board (Prot n. 16929, dated March 16, 2020).

AUTHOR CONTRIBUTIONS STATEMENT

A.R. and F.M. designed the project and experiments. F.F, M.B, A.E, F.A, E.C., M.dA., S.M, M.A.M., F.M., M.M., E.B., A.R. and F.M. performed clinical evaluations. A.G., F.V., S.A., L.B., and M.B. carried the experiments. E.B. and S.F performed bioinformatic analyses. R.T., C.F., and E.B. carried out statistical analysis and prepared the tables. E.B., A.R. and F.M. wrote the manuscript. E.B. submitted this paper. All authors reviewed the manuscript.

ADDITIONAL INFORMATION

The authors declare no competing interests.

FIGURE LEGENDS

Fig. 1. Clinical characteristic and mutated genes. The population is clustered into four qualitative severity groups indicated with different colors depending on the respiratory impairment and the need for ventilation. Red color is used for high care intensity group (those requiring invasive ventilation), orange for intermediate care intensity group (those requiring non invasive ventilation i.e. CPAP and BiPAP, and high-flows oxygen therapy), pink for low care intensity groups (those requiring conventional oxygen therapy) and light blue for very low care intensity groups (those not requiring oxygen therapy). Patients 13 and 14 are reported in grey because they are siblings. A detailed clinical characterization is provided (i.e blood group, multiple organs involvement, presence of comorbidity, clinical laboratory parameters, etc.) along with the genetic background for each patient. Liver and pancreas involvement is indicated in green color, heart involvement in magenta and kidney involvement in purple. The presence of hyposmia, classified from 0 to 10 using VAS score, is indicated in light blue. A darker shade of color represents a more severe organ damage. In-silico predicted deleterious variants in genes relevant for infection and pathogenic variants (both common and rare) reported in ClinVar Database are described and a further subdivision between genes involved in a mendelian disorder and/or viral infection susceptibility is provided. For all these gene categories, dark grey is used to identify the homozygous status of the variants while light grey for the heterozygous status.

In the end, statistically significant genes obtained after Gene Burden analysis are listed: *PRKRA* and *LAPTM4B* mutational burden resulted to be enriched in the 35 COVID-19 patients compared to the controls, while *OR4C5*, *ZNF717*, *FAM104B* and *NDUFAF7* have proven to have an opposite trend, having a mutational burden more enriched in controls. For this reason, for *NDUFAF7*, *OR4C5*, *ZNF717* and *FAM104B* genes the grey color and the white color are inverted because having less variants and by consequence a more functional

gene represents a susceptibility factor. For this category, white color underlies a higher mutational burden while grey color indicates lower mutational burden.

Fig. 2. Mutational burden at the gene level. The significance threshold including the Bonferroni correction is shown as a red line (□ = 0.05, number of test, m = 7196). P-values for genes below the significance threshold are shown as grey dots. Black circles are used for the 6 genes that were identified as statistically different between COVID-19 samples and controls.

Table 1. Clinical characteristics of Italian COVID19 patients. COVID-19 cohort is grouped in 4 qualitative severity groups depending on the respiratory impairment and the need of ventilation. Group 1 requires invasive ventilation. Group 2 requires CPAP/BiPAP/high-flows oxygen therapy. Group 3 requires conventional oxygen therapy. Group 4 does not require oxygen therapy. Clinical characteristics are listed and the number of patients are indicated for each of them.

Supplementary Information.

Supplementary Table S1. List of genes conferring COVID-19 susceptibility identified with the gene burden test analysis. Genes harboring deleterious mutations with statistically significant higher frequency in control than in COVID-19 patients are ordered based on p-value deriving from gene burden test analysis. The p-value adjusted is provided after Bonferroni correction.

Supplementary Table S2. List of COVID-19 protective genes identified with the gene burden test analysis. Genes harboring deleterious mutations with statistically significant

higher frequency in COVID-19 patients than in control are ordered based on p-value deriving from gene burden test analysis. The p-value adjusted is provided after Bonferroni correction.

Supplementary Table S3. Rare variants identified in patients cohort. Rare variants identified in COVID-19 patients according to the Mendelian-like model are reported (see *Methods section*)

Supplementary Fig. S1: Clinical Data Questionnaire. The Questionnaire includes five different categories of data: Patient personal anamnesis and family history, Diagnostic Information, Laboratory Tests, Therapy and Complications. Clinical data were collected in detail for all COVID-19 patients.

	Sample ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35
CLINICAL CHARACTERIZATION	**Gender**	M	F	M	M	M	M	M	M	M	M	M	M	F	F	F	M	F	M	F	F	M	M	M	F	F	M	M	F	M	M	M	F	M	M	M
	Age	68	70	62	68	55	55	42	60	51	75	81	62	60	62	77	49	98	49	67	66	91	60	65	60	64	74	81	87	55	65	65	58	31	74	31
	Blood Group	A	B	A	B	0	0	A	0	B	A	0	A	0	0	0	0	A	0	0	A	0	A	A	A	B	0	0	A	A	A	A	0	0	A	A
	Respiratory Severity P/F score (worst value)	67	96	80	103	93	149	Unknown	126	200	130	156	162	210	339	293	318	140	285	368	100-200	264	280	Unknown	100-200	200-300	200-300	279	312	>300	304	347	Unknown	>300	400	Unknown
	Lung Imaging Grading (0-28) (worst)	13	13	16	14	11	10	13	11	13	10	15	13	9	5	6	7	14	5	8	11	10	10	3	9	7	14	8	6	4	8	9	9	3	5	0
	Hyposmia VAS Grading (0-10)	N/A	No	0	0	0	5	0	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	0	0	0	0	0	0	0	0	0	0	0
	Hypogeusia (Yes/No)	N/A		No	No	No	Yes	No	Yes	No	No	No	No	No	No	No	No	No	No	Yes	No	No	No	No	Yes	No	No	No	No	No	No	No	No	No	No	No
	Heart involvement (T=T-Troponin >15 ng/L; B= NT-	Yes T/B	No	Yes T/B	Yes B	Yes B	No	No	No	Yes T	Yes A	Yes T/B	No	No	No	No	No	Yes T/B	No	No	No	Yes B/A	Yes A	No	No	No	Yes T/A	Yes T/B	Yes B	No	No	No	No	No	No	No
	Liver involvement (both ALT and AST > 40 U/L)	Yes	Yes	Yes	Yes	No	Yes	Yes	No	Yes	No	Yes	Yes	Yes	No	No	Yes	No	Borderline	No	No	No	Yes	No	No	Yes	No	No	No	Yes	Yes	Yes	No	Yes	No	No
	Pancreas involvement (lipase >60 UI/L and pacreatic amylase >53 UI/L)	No	No	Yes	No	No	Yes	Yes	No	Yes	No	Yes	Yes	Yes	Yes	No	No	Yes	Borderline	No	No	Yes	Yes	No	No	Yes	No	No	Yes	No	No	No	No	Yes	No	No
	Kidney involvement (creatinine M > 0.7-1.20; F > 0.5-1.10 mg/dl)	No	No	No	No	No	No	No	No	No	Yes	Yes	Yes	No	No	No	No	Yes	Yes	No	No	Yes	Yes	Yes	No	No	No	No	No	Yes	No	No	No	No	No	No
	CD4 lymphocytes (cell/ul (mm^3); n.v. 400-1500)(worst value)	286	314	56	532	250	533	618	35	1108	1028	367	536	582	1064	81	-	158	-	1493	458	185	2115	616	758	1705	361	591	194	623	226	485	1116	381	741	736
	NK cells (cell/ul (mm ^3); n.v. 90-590)(worst value)	103	95	25	24	65	94	130	2	267	59	20	73	118	158	57	-	126	-	37	120	128	112	85	147	180	305	98	101	334	40	141	204	291	117	252
	Neutrophils (min value)/Lymphocytes (min value)	16.4	20	13.1	4	12	3	2	1.6	3.4	2.3	28	2.9	0.7	2.3	15	2.8	13.7	5.6	3.6	1.8	10.9	7.6	2.4	4.7	4.6	4.4	15.2	5.4	1.7	6.3	6.9	1.1	1.6	12.1	2.32
	IL6 (pg/ml; n.v. <27)(basal value)	450	73	46.3	88.2	1138	746	41.5	217.5	-	-	1134	13.5	195.6	-	14.9	5.2	-	2.9	-	28.6	36.3	-	-	-	-	-	-	-	-	126.1	19	5.4	-	-	-
	CRP (mg/dl; n.v. <0.5)(max value)	28.3	13.25	32.24.00	26.08.00	25	5.4	5.12	7.16	20.21	2.31	28.46.00	13.24	1.1	0.46	5.13	22.27	18.17	40.08.00	9.6	6.5	4.09	25.3	0.19	8.42	7.08	15.08	3.05	3.34	3.14	5.11	9.06	0.59	2.34	5.56	0.05
	Fibrinogen (mg/dl; n.v. Klauss 200-400)(min value)	139	546	356	575	456	145	638	-	-	606	314	338	518	512	566	701	575	772	680	282	486	859	672	391	471	660	373	531	427	660	546	378	-	623	-
	LDH (UI/l; n.v. M 135-225; F 135-214)(max value)	551	387	238	416	217	367	407	330	371	649	700	454	315	240	593	371	350	255	528	264	250	272	198	264	358	412	220	335	213	309	292	232	294	255	143
	D-dimer (ng/ml; n.v. <500)(max value)	31470	847	17550	6853	3286	847	788	1009	1526	2071	88236	55009	735	3510	6765	903	2190	1306	2934	1749	790	1002	1007	1167	666	912	1055	23894	1148	1286	1509	599	-	621	425
	B=Baricitinib)	T	T	T	T	T	No	T	T, R	T	No	T	R	T	No	No	T	No	No	No	T	B	No	No	No	No	No	No	No	No	No	No	No	No	No	-
	D=Dexamethasone)	No	No	M	M	M	M	D	M	M	M	M	M	M	No	D	No	M	No	D	No	M	M	M	M	M	M	M	M	No	No	No	No	-	M	-
	Enoxaparin (Yes/No)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	Yes	Yes	No	Yes	Yes	No	Yes	no	Yes	Yes	No	Yes	Yes	-	Yes	Yes	Yes	-	Yes	-
	Other anti-COVID19 drugs (HCQ=Hydroxicloroquine; A=Azithromycine; REM=Remdesevir;	HCQ, A, REM	HCQ, A	HCQ	HCQ, A, L/R	HCQ	HCQ, A	HCQ, A, L/R	-	HCQ, L/R	HCQ, A	HCQ, A	C,A	HCQ, A	HCQ, A	-	HCQ	HCQ, A	-	-	HCQ, A	HCQ	HCQ, A	-	A	HCQ	HCQ, A	HCQ, A	HCQ	-	HCQ	HCQ, L/R	-	HCQ, L/R	HCQ, A	-
	Clinical known co-morbidities	Congestive Heart Failure, Hypertension	Breast Cancer	Colon Cancer	Hypothyroidism	Hypertension, Ictus Cerebri	Obesity	None	Leukemia (BPDCN)	None	FA, Aortic Aneurism, Hypertension, DM	Congestive Heart Failure, Kidney Failure	Hypothyroidism, Hypertension	Congenital Rickets	Chronic obstructive pulmonary disease (COPD)	Hypertension, Diabetes mellitus (DM)	Hypothyroidism	Breast Cancer, Hypertension, Hypothyroidism	G6PD Deficiency, Hypertension	Obesity, Psichiatric Disorder	None	Cutaneous squamous cell carcinoma of the left ear, Hypertension, Atrial Fibrillation, Congestive Heart Failure	Hypertension	Diabetes mellitus, Previous HBV infection, Hypertension, Acute myocardial infarction	Breast Cancer, Diabetes Mellitus, Hypothyroidism, Obesity, Dyslipidemia	Hypertension	Atrial Fibrillation, Diabetes Mellitus	Hypertension	Hypertension	Diverticolosis	Bronchiectasis	None	Depression, Thyroid Nodules	-	Prostate Cancer, Melanoma, Basal cell carcinoma	-
RARE VARIANTS REPORTED AS PATHOGENIC IN ClinVar	**MENDELIAN DISORDER** - NO VIRAL INFECTION SUSCEPTIBILITY	*PRLR* (I170L)					*F11* (A430T)							*FLG* (S3247*)	*FLG* (S3247*)						*PPP2R1B* (G90D)									*PRLR* (I170L)				*PRLR* (I170L)		*PPOX (P256R)*
	MENDELIAN DISORDER - VIRAL INFECTION SUSCEPTIBILITY													*TGFBI comp*	*TGFBI het*				*G6PD/ ZEB1*						*PMS2*											
	CARRIER of MEDELIAN DIS - VIRAL INFECTION SUSCEPTIBILITY		*GALC*			*MPO/ BCHE*			*GJB2/ AGXT*			*LAMA2*		*MPO*				*LPL*				*FCN3*				*MPO*	*FCN3*	*PMM2* (R141H)			*ALG8* (R364*)	MPO		*SLC36A2* (G87V)		*ATP7B / OCA2*
	CARRIER of MEDELIAN DIS - VIRAL INFECTION SUSCEPTIBILITY - *ABCC6*	Intronic			R1304R								Intronic				UTR3'									Intronic		Intronic				R1304R	V400I			
RARE VARIANTS RELEVANT FOR INFECTION	*ADAR*																							A935V												
	AIRE																						H415R									R473T				
	CASP8																									Y310*								I315V		
	CHD7													N2891S	N2891S					R947Q			G2786R													
	CLEC4M						Q10K																V742M					Q10K								
	COL5A1/2	*COL5A1* (R630W)			*COL5A1* (P1529H)														*COL5A1* (P149L; E269K)											*COL5A2* (I1279V)	*COL5A1* (T915M)					
	FURIN			T413I																				R483W												
	HCRTR1/2			R405*	C193S				L407F														V93I												
	ILs and ILRs													*IL2RB* (N43S)	*IL2RB* (N43S)				*IL12RB1* (G441S)			*IL13* (T41I)						*IL16* (T439M)							*IL13* (L72P)	
	Immunodeficiency genes		*PTPRC (V1224I)*							*ZAP70* (S313R)		*PRKDC* (R2627C)											*DCLRE1C* (S520fs5*)		*CORO1A* (V174M); *JAK3* (R840C)		*FOXN1* (P475L); (S570L)	DOCK8 (H541Q)							*CARD11* (I544L)	
	NOS3					P1084S												V1073A													G173R					P1084S
	OPRM1				S42T													S437N	N386D									R260H								
COMMON VARIANTS REPORTED AS PATHOGENIC IN ClinVar	*MBL2*	G54D	G54D		R52C	G54D	G54D	G54D		R52C		G54D	G54D					G54D		G54D	G54D				G54D		G54D				G54D			G54D	G54D	
	IRGM (L105L)																																			
	PTPRJ (Q276P)																																			
	CST3 (A25T)																																			
	SAA1 (A70V)																																			
	PRSS1	N29I/ N54S	N54S	N29I/ N54S		N29I/ N54S	N29I/ N54S	N29I/ N54S	N29I	N54S	N54S	N54S	N29I/ N54S	N54S	N54S	N29I/ N54S	N29I/ N54S	N29I/ N54S	N29I/ N54S	N54S	N54S	N29I/ N54S	N29I	N54S	N29I	N29I/ N54S	N29I/ N54S	N29I/ N54S	N29I/ N54S	N29I/ N54S	N29I/ N54S	N54S	N29I/ N54S	N29I/ N54S	N54S	N29I/ N54S
STATISTICALLY SIGNIFICANT GENES by collapsing gene analysis	*PRKRA* (I226N)																																			
	LAPTM4B					P219L					P220L	I109F								P50T													P219L			
	OR4C5 (mutations more common in controls)																																			
	ZNF717 (mutations more common in controls)																																			
	FAM104B (mutations more common in controls)																																			
	NDUFAF7 (mutations more common in controls)																					E56D			R17H											

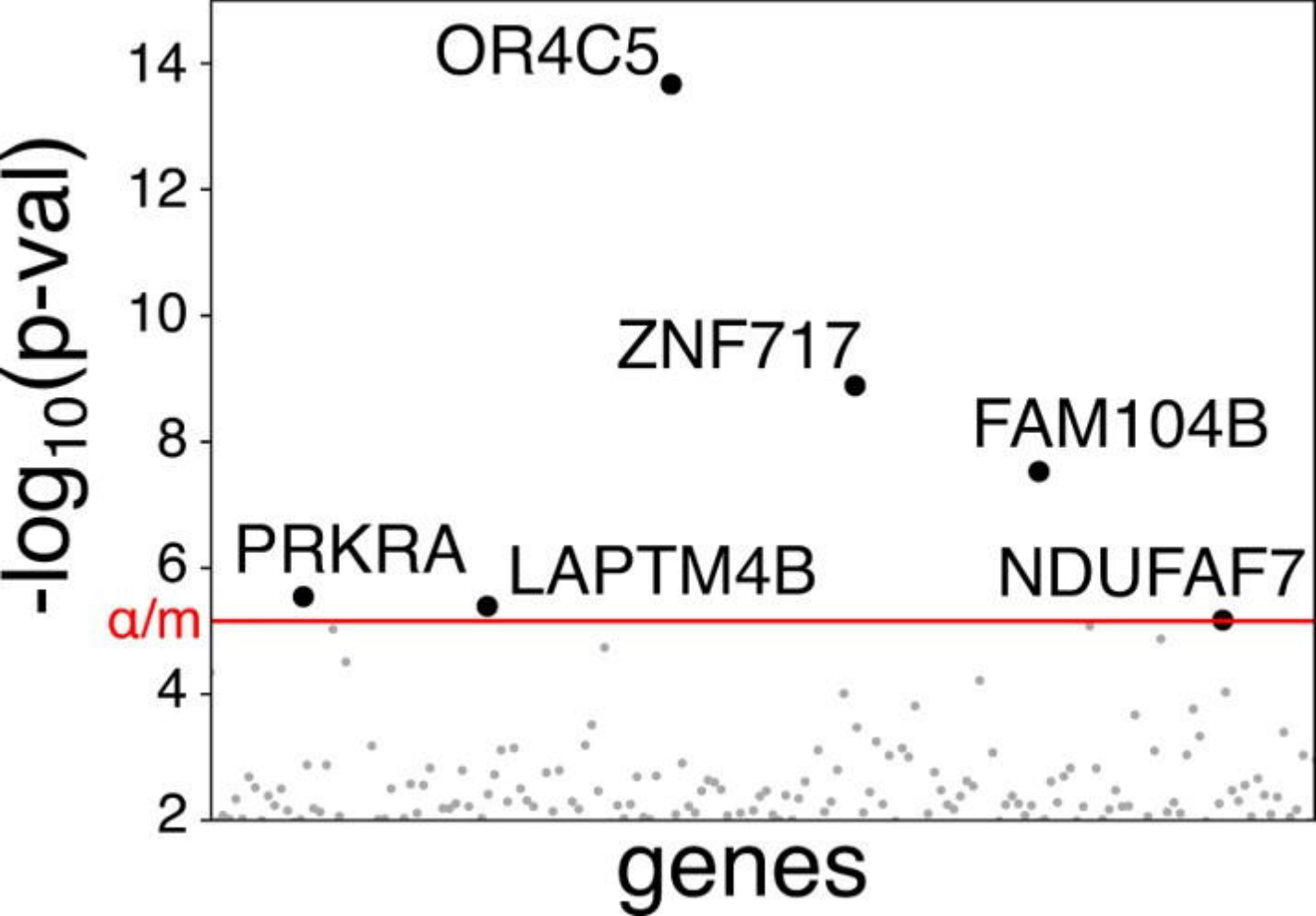
OR4C5
ZNF717
FAM104B
PRKRA
LAPTM4B
NDUFAF7
α/m
14
12
10
8
6
4
2
-log10(p-val)
genes

Table 1. Clinical characteristics COVID19 patients admitted to the University Hospital of Siena (Italy)

Subject characteristics	**Group 1**	**Group 2**	**Group 3**	**Group 4**
No. of subjects	6	7	15	7
Median age (range) [a]	65 (55-70)	58 (31-74)	66 (49-98)	58 (31-74)
Gender				
Male	5	6	7	6
Female	1	1	8	1
Blood Group				
A	2	3	6	5
B	2	1	1	0
0	2	3	8	2
PaO_2/FiO_2				
≤ 100	4	0	0	0
100-200	2	5	3	0
200-300	0	1	7	0
>300	0	0	4	5
Unknown	0	1	1	2
Lung imaging grading (CXR score)				
0-10 (mild)	1	2	12	7
11-20 (moderate)	5	5	3	0
21-28 (severe)	0	0	0	0
Laboratory findings				
$CD4^+$ T cells count				
<400 (cell/ul)	4	2	5	2
400-1500 (cell/ul)	2	5	8	5
Unknown	0	0	2	0
NK cells count				
<90 (cell/ul)	3	4	3	1
90-590 (cell/ul)	3	3	10	6
Unknown	0	0	2	0
IL-6 value				
<27 (pg/ml)	*0*	1	3	2
>27 (pg/ml)	*6*	4	2	1
Unknown	*0*	2	10	4
Fibrinogen				
<200 (mg/dL)	*2*	0	0	0
200-400 (mg/dL) normal value		2	3	1
>400 (mg/dL) hyperfibrinogenemia	*3*	3	12	4
Unknown	*0*	2	0	2
CRP				
<0,5 (mg/dl)	*0*	0	2	1
>0,5 (mg/dl)	*6*	7	13	6
LDH				
Normal value M: 135–225 (Ul/l); F: 135-214 (Ul/l)	1	1	2	2
High value	5	6	13	5

D-Dimer				
Less than 10 times higher than normal value	3	5	13	5
normal	0	0	0	1
More than 10 times higher than normal value	3	2	2	0
Unknown	0	0	0	1
Hyposmia (VAS score)				
<2 (normal)	4	6	14	7
2-5 (intermediate)	1	0	0	0
>5 (severe)	0	1	1	0
Hypogeusia (VAS score)				
No	4	6	13	7
Yes	1	1	2	0
Heart involvement				
Yes	4	3	6	0
T=T-Troponin >15 (ng/L); B= pro-BNP M > 88 (pg/ml); F > 153 (pg/ml); A=arrhythmia)	T/B 2 B 2	T/B 1 T 1 A 1	T/B 2 T/A 1 B/A 1 A 1 B 1	
No	2	4	9	7
Unknown	0	0	0	0
Hepatic (H)/Pancreatic involvement (P)				
H and P	2	5	6	1
H only	3	0	1	1
P only	0	0	1	1
None	1	2	7	4
Kidney involvement				
Yes	0	3	5	1
No	6	4	10	6
Co-morbidities				
Cardiovascular disease	1	2	3	
Hypertension	2	2	8	
Tumor	2	1	2	1
Diabetes			4	
Pulmonary disease			1	1

[a] Age in years. NA, not applicable.