Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Review

Why does SARS-CoV-2 hit in different ways? Host genetic factors can influence the acquisition or the course of COVID-19

Maria Monticelli a, Bruno Hay Mele b, Giuseppina Andreotti c,∗, Maria Vittoria Cubellis a, c,∗, Guglielmo Riccio d

a Department of Biology, Università Federico II, 80126, Napoli, Italy
b Integrative Marine Ecology Department, Stazione Zoologica Anton Dohrn, Villa Comunale, 80121, Napoli, Italy
c Istituto di Chimica Biomolecolare –CNR, 80078, Pozzuoli, Italy
d Scuola di Specializzazione in Pediatria, Università degli Studi di Trieste, 34127, Trieste, Italy

ARTICLE INFO

Keywords:
COVID19
Data mining
SNPs
MBL2
TMPRSS2
CD27

ABSTRACT

The identification of high-risk factors for the infection by SARS-CoV-2 and the negative outcome of COVID-19 is crucial. The genetic background of the host might account for individual responses to SARS-CoV-2 infection besides age and comorbidities. A list of candidate polymorphisms is needed to drive targeted screens, given the existence of frequent polymorphisms in the general population.

We carried out text mining in the scientific literature to draw up a list of genes referable to the term “SARS-CoV*”. We looked for frequent mutations that are likely to affect protein function in these genes. Ten genes, mostly involved in innate immunity, and thirteen common variants were identified, for some of these the involvement in COVID-19 is supported by publicly available epidemiological data. We looked for available data on the population distribution of these variants and we demonstrated that the prevalence of five of them, Arg52Cys (rs5030737), Gly54Asp (rs1800450) and Gly57Glu (rs1800451) in MBL2, Ala59Thr (rs25680) in CD27, and Val197Met (rs12329760) in TMPRSS2, correlates with the number of cases and/or deaths of COVID-19 observed in different countries. The association of the TMPRSS2 variant provides epidemiological evidence of the usefulness of transmembrane protease serine 2 inhibitors for the cure of COVID-19. The identified genetic variants represent a basis for the design of a cost-effective assay for population screening of genetic risk factors in the COVID-19 pandemic.

1. Introduction

Many people affected by COVID-19 develop fever, cough, fatigue, shortness of breath, muscular pain, and headache. In the most severe cases medical conditions, such as respiratory failure, occur and eventually lead to death.

Age and comorbidities, such as cardiovascular diseases, diabetes, and chronic lung diseases, are known risk factors for worse outcomes (Fang et al., 2020). Furthermore, altered biomarkers, namely C-reactive Protein, high-sensitivity troponin T, prothrombin time, fibrin degradation products, D-Dimer, and CD4+ count, help differentiate patients who will develop severe symptoms from those who will only be mildly affected (Zheng et al., 2020). These factors alone are not sufficient to identify all subjects at high risk, and on top of that, they tend to overlap.

Genetic factors, on the other hand, also influence risks associated with COVID-19 but are more likely to be independent of age-related comorbidities. Specific studies on the subject are very few and in most cases genes have been selected regardless of the existence of frequent polymorphisms in the general population, lowering their potential as markers for large scale genetic screenings (Hussain et al., 2020).

SARS-CoV-2 and SARS-CoV belong to the Betacoronavirus genus (Chen et al., 2020) and share the same cellular receptor, the angiotensin-converting enzyme 2 (ACE2), and very similar nucleotide sequences (Yan et al., 2020).

In this paper, we carried out a bioinformatic analysis to predict polymorphisms that could either influence the acquisition or the course

https://doi.org/10.1016/j.ejmg.2021.104227
Received 5 November 2020; Received in revised form 14 March 2021; Accepted 12 April 2021
Available online 16 April 2021
1769-7212/© 2021 Published by Elsevier Masson SAS.
of COVID-19. Text mining was carried out to extract a list of human genes that have been associated with the term “SARS-CoV*” in the literature. We highlighted the polymorphisms of these genes that are very frequent in the general population and are likely to affect the function of the mutated protein. We hope that our prediction can guide and speed up experimental tests and help to find those high-risk subjects that have not been recognized by age, co-morbidities, or biomarkers so far.

2. Question: which host genetic factors are likely to influence sensitivity to covid-19?

High variability of symptoms and outcomes characterizes SARS-CoV-2 infection. The phenotypic spectrum of the disease ranges from asymptomatic individuals to patients requiring ventilation due to severe pneumonia (Garcia, 2020). Different hypotheses have been put forward, the role of the host’s genetic background is currently emerging (Godri Pollitt et al., 2020; Latini et al., 2020; Strafella et al., 2020; Zhang, S.-Y. et al., 2020). Data from a genomewide association (GWAS) study involving 1980 patients with COVID-19 and 2381 healthy control participants in Spain and Italy have been recently published (Group, 2020). Two loci were strongly associated to susceptibility to SARS-CoV-2, one on chromosome 9 covering the ABO blood group and another one on chromosome 3 covering six genes, SLC6A20, LZTFL1, CCR9, FYCO1, CXCR6, and XCR1.

We would like to contribute to gene-specific candidate-driven studies by suggesting missense variants that can cause severe symptoms or facilitate the infection and expand the list of identified by GWAS studies. We evaluated missense mutations that are frequent in the general population and occur in genes that do not cause haploinsufficiency. We selected those that are deleterious for the protein product and we excluded those associated to blood groups. A deleterious mutation on a single allele might not be enough to cause a life-threatening disease and hence it could be relatively frequent but might confer special sensitivity or resistance to viral infection. We excluded variants occurring in promoters, introns, or untranslated regions because their deleteriousness is less predictable. Hence our study will provide a subset of the genetic variants that are likely to affect the sensitivity to infection or the course of COVID-19.

2.1. Procedure to identify candidate variants

We obtained from the Exome Aggregation Consortium (ExAC) a list of human genes (14995) with a pLI < 0.9. pLI is a value that measures the probability of being loss-of-function intolerant (pLI). Genes with a pLI < 0.9 have a low probability of causing haploinsufficiency when affected by the severe mutation. We developed a simple pipeline to use such a list of genes and corresponding proteins as input and look within the Scopus database for papers containing any of these protein names and the term “SARS-CoV*” in the article’s title, abstract, or keywords. The script, which is written in R language and exploits the scopus package, is provided in Supplementary File S1. We excluded the genes covered by the GWAS analysis (Group, 2020) obtained a “SARS-CoV* referable genes” list (106 genes; Supplementary File S2).

We searched within GnomAD for frequent variants, missense or loss-of-function, in “SARS-CoV* referable genes”, setting the threshold to allele frequency at 1%. Lastly, we used Polyphen2 (http://genetics.bwh.harvard.edu/pph2/), to exclude benign mutations. Frequent missense variants were found in 56 genes and deleterious frequent deleterious variants in ten genes. The scheme of the search is provided in Fig. 1.

The list of identified variants with their frequencies is reported in Table 1. We observed that the candidate variants were mostly in genes involved in innate immune defense. References to databases and programs utilized in this analysis are provide as a supplementary file S3.

2.2. Correlation between allele frequencies and COVID-19 cases and deaths per population

Data on the prevalence in different populations were available for six of our candidate variants. The datasets relative to rs12329760 (Val197Met, TMPRSS2), rs25680 (Ala59Thr, MBL2), and rs1800451 (Gly57Glu) were reported by Verdu et al. Data on the population distribution of the remaining six variants were not available. Notwithstanding the limitations associated with publicly available data, we observed interesting results for five variants, namely Arg52Cys (rs1800450), and Gly57Glu (rs1800451) in MBL2, Val197Met (rs12329760) in TMPRSS2 and Ala59Thr (rs25680) in CD27.

Fig. 1. Procedure to identify candidate variants. The Exome Aggregation Consortium (ExAC) provided a list of human genes and annotated them with a value that measures the probability of being loss-of-function intolerant (pLI). The names of the genes were associated to the names of the corresponding proteins. Such list was used as the input to look within the Scopus database for papers containing any of the protein names and the term “SARS-CoV*” in the article’s title, abstract or keywords. “SARS-CoV* referable genes”, encompassing frequent variants, missense or loss of function, were identified setting the threshold to allele frequency to 1%, and excluding variants that did not pass quality control or occur in non-canonical transcripts. The genes covered by the GWAS analysis (Group, 2020) were excluded. Among these genes, those encompassing frequent deleterious variants were identified.
Table 1: Likely deleterious frequent variants (allele frequency >1%) in genes referable to the term “SARS-CoV-” in the literature. Variant frequency is referred to gnomAD total exome frequency obtained from wAnnovar annotation. For variants covered in the GWAS study, GUST90000255 Fisher p-values are provided.

| Gene name | Protein name/Mim phenotype | Reference SNP ID number | Amino acid change | Variant frequency | GWAS p-value |
|-----------|----------------------------|-------------------------|------------------|------------------|-------------|
| ACE       | Angiotensin-converting enzyme | rs3790025 | Tyr244Cys | 0.0107 |            |
| AHSG      | Alpha-2-HS-glycoprotein | rs140827890 | Ala164Thr | 0.0114 | 0.05 |
| CD27      | CD27 antigen #615122 | rs25680 | Ala59Thr | 0.1912 |            |
| GFFA      | Gial fibrillary acidic protein | rs59291670 | Asp157Asn | 0.11 |            |
| IFI30     | Gamma-interferon-inducible lysosomal thiol reductase | rs11554159 | Arg76Gln | 0.2281 |            |
| MBL2      | Mannose-binding protein C #614372 | rs1800450 | Gly54Asp | 0.141 |            |
| MEFV      | Pyrin #134610 #249100 | rs3743930 | Glu148Glu | 0.0708 |            |
| TLR2      | Toll-like receptor 2 #607948 | rs5743704 | Pro631His | 0.0281 | 0.02 |
| TLR3      | Toll-like receptor 3 #609423 | rs5743708 | Arg753Gln | 0.0176 |            |
| TMPRSS2   | Transmembrane protease serine 2 | rs12329760 | Val197Met | 0.2452 | 0.01528 |

**MBL2** encodes a mannose-binding protein C that binds mannose, fucose, and N-acetylglucosamine on different microorganisms and activates the lectin complement pathway. The deficiency of mannose-binding protein C (MIM #614372) is associated with increased susceptibility to infection (Degen et al., 2011). MBL2 binds and activates the lectin complement pathway in viral infections such as HIV (Garred et al., 1997) and influenza (Kase et al., 1999). We tested three frequent missense mutations in MBL2: Gly54Asp (rs1800450), Arg52Cys (rs5030737), and Gly57Glu (rs1800451). All these mutations affect the ability of the protein to bind carbohydrate surfaces and MASp-2, the mannann-binding lectin serine protease 2 which activates the complement system (Turner, 2003); the lowest insertion is observed with Arg52Cys (Larsen et al., 2004). In a study conducted in China, it was proven that Gly54Asp (rs1800450) is significantly associated with susceptibility to SARS-CoV infection but not with disease severity (Jp et al., 2005; Zhang et al., 2005); unfortunately, the other two variants were not tested. We found that the allele frequency per population of Arg52Cys (rs5030737) correlates positively with the number of cases per country [p-value 0.0394] of COVID-19, but not with the number of deaths (Fig. 2, panels A and B), while the allele frequency per population of Gly54Asp (rs1800450) correlates positively both with the number of cases [p-value 0.0018] and the number of deaths [p-value 0.02] (Fig. 2, panels C and D). This finding supports the hypothesis that mannose-binding protein C plays a protective role and that MBL2 inactivation is a risk factor for SARS-CoV-2 infection, and that the immune response plays a role in the course of the disease.

Unexpectedly, the allele frequency per population of Gly57Glu (rs1800451) was found to be negatively correlated with the number of cases [p-value 0.0122] (Fig. 3, panel A). The correlation with the number of deaths was significant when the data from Iran were excluded [p-value 0.0241] (Fig. 3, panel B).

**TMPRSS2** encodes transmembrane protease serine 2. Mouse models have shown that SARS, MERS CoVs and Influenza viruses use this protease to cleave the Spike protein and achieve viral internalization (Iwata-Yoshikawa et al., 2019). A negative correlation between the allele frequency of Val197Met and the number of cases [p-value 0.0012] and deaths [p-value 0.0057] per country (Fig. 4 panels A and B) was found. We observed the number of cases and deaths as a function of prevalence clusters in two groups and the high mortality group is associated with a low prevalence of the variant, indicating that genetic TMPRSS2 inactivation is a protective factor against SARS-CoV-2 infection and progression towards severe cases. A lower frequency of Val197Met in Italian COVID-19 patients compared to the allelic frequency of this variant in GnomAD for the EUR reference population has recently been reported (Latini et al., 2020). Val197Met is a SNP covered by the GWAS study (Group, 2020) and it is significantly associated with susceptibility to (p-value < 0.05). TMPRSS2 is a druggable protein that can be inhibited by camostat mesylate, a drug approved in Japan, by bromhexine and nafamostat. Clinical trials are ongoing for their usage in COVID-19 (ClinicalTrials.gov Bromhexine: NCT04355026, NCT04273763, NCT04334039, Camostat: NCT04321096, NCT04355052, NCT04353284, NCT04338906, NCT04470544, Nafamostat: NCT04352400; NCT04473053). Our data provide epidemiologic evidence of the role of the protease in SARS-CoV-2 infection and encourages the usage of protease inhibitors for the therapy (Hoffmann et al., 2020).

**CD27** receptor is thought to have an essential role in connecting the CD4 and CD8 branches of the immune system. Unbalances between the two branches has been associated with a worse prognosis in SARS-CoV infections (Li et al., 2008). Known pathological mutations of CD27 exist, and they are associated with cytokine storms (Amberger et al., 2009) and hemophagocytic lymphohistiocytosis (HLH) like syndromes that frequently characterize severe SARS-CoV-2 infections (Shoenfeld, 2020). We observed a statistically significant negative correlation between the allele frequency of Ala59Thr and the number of deaths per country [p-value 0.0337] (Fig. 5, panel B), but not with the numbers of cases (Fig. 5 panel A).

### 2.3. Description of the other genes potentially related to COVID-19 response

As discussed above, the association between six inactivating variants (Table 1) and SARS-CoV-2 susceptibility could not be demonstrated because of the absence of data on population distribution. Nevertheless, literature data support a possible role of some of them in COVID-19.

The role of **ACE** in innate immunity has been reviewed (Bernstein et al., 2018). A very frequent intronic insertion/deletion I/D polymorphism influences ACE expression (Rigat et al., 1990). A recent paper has suggested the usage of ACE I/D polymorphisms for the identification of high-risk COVID-19 patients (Delanghe et al., 2020). The hypothesis that variants of ACE are associated with COVID-19 risks is very interesting and is in line with several observations. Although ACE2, rather than ACE, has been identified to be one of the receptors of the SARS-CoV-2, it has been proposed that the unbalance between ACE and...
ACE2 expression leads to lung injury (Guo et al., 2020). Children, who have higher levels of ACE than adults, are much less affected by COVID-19 (Guo et al., 2020). ACE inhibitors have been proven to upregulate ACE2, and some authors speculate a possible role of these extremely commonly used drugs in infection. Their role in COVID-19 disease, however, is very controversial, because upregulation of \( \text{ACE2} \) may have a protective role against lung inflammation and fibrosis (Guo et al., 2020). Publications on \( \text{ACEI/D} \) polymorphisms and SARS-CoV are of little help because they provide different results (Chan et al., 2005; Itoyama et al., 2004). A final word on this subject could come from a very recent paper by Richards and coworkers (Butler-Laporte et al., 2020). Although the authors did not include Tyr244Cys (rs3730025) in their study, they found that variants which decrease the expression of \( \text{ACE} \) do not increase COVID-19 susceptibility or severity.

\( \text{AHSG} \) encodes a protein known as human fetuin which is thought to be an immune modulator, mainly acting on cytokine release from macrophages. Macrophages play a central role in the IL1, 6, TNF alfa, INF gamma dependent cytokine storm. Its deletion has already been linked to worse outcomes in SARS-CoV infected patients (Zhu et al., 2011). The SNP Ala164Thr is covered by the GWAS study (Group, 2020) and significantly associates with susceptibility to COVID-19 (\( p \)-value = 0.05).

The role of \( \text{GFAP} \) in COVID-19 might be prevalently associated to the severity of the disease and in particular to the development of neurological symptoms. \( \text{GFAP} \) is a TGF-beta induced gene that participates in the ERG dependent TGF-beta triggering of fibrosis consequent to ROS production. It has been proven that SARS-CoV viruses possess a deubiquitinating protein named PLpro that stimulates such pathway, inducing lung fibrosis in infected patients. Pulmonary fibrosis has been described in cases of SARS-CoV infections, but so far it has not been described to be a relevant feature of the disease (Yang et al., 2006). Very little is known so far about COVID-19 pathology, and future studies could demonstrate the importance of fibrosis since it is a central feature in uncontrolled activation of macrophage-mediated inflammation.

Fig. 2. \( \text{MBL2} \) allele frequencies per population correlation with COVID-19 cases and deaths. rs5030737 (Arg52Cys) allele frequencies in different countries from Verdu et al. were correlated with the respective number of cases (panel A; \( p \)-value 0.0394) and deaths (panel B; n. s.) of COVID-19. rs1800450 (Gly54Asp) allele frequencies were correlated both with cases (panel C; \( p \)-value 0.0018) and deaths (panel D; \( p \)-value 0.0199) of COVID-19 per country. Different symbols relate to different populations associated with the same country [Purple circle: GBP; purple square: GBB.]. Details about populations and their association to countries are provided in supplementary file S1. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)
IFI30 (Gamma-interferon-inducible lysosomal thiol reductase, GILT) has a role in protecting against the internalization of SARS-CoV viruses in the lung tissue (Chen, D. et al., 2019). Polymorphisms in this gene may be correlated with a predisposition to infection by SARS-CoV-2, as well as with a worse prognosis due to higher initial viral charge. The viral charge has, in fact, already been proven to be linked to worse clinical outcomes (Chen, D. et al., 2019).

MEFV (Pyrin) is associated to Familial Mediterranean Fever whose symptoms, fever, abdominal/chest pain, elevation of C-reactive protein, and leucocytosis, overlap in part with those of COVID-19. Deleterious variants in this gene might influence the severity of the infection. MEFV is thought to be involved in the NLRP3 dependant IL-1 beta modulation (Manukyan and Aminov, 2016). The capability of SARS-CoV viruses to activate the NLRP3 inflammasome has already been described (Chen, I.-Y. et al., 2019). There are some known pathological mutations of MEFV that cause autoinflammatory disease due to a IL1, 6, TNF- alfa INF-gamma cytokine storm (Manukyan and Aminov, 2016).

Toll like receptors activate Type-I interferon (alpha and beta) response that enhance innate antiviral responses such as the limitation of cell replication, the induction of apoptosis, and the synthesis of

![Fig. 3. MBL2 allele frequencies per population correlation with COVID-19 cases and deaths.](image)

rs1800451 (Gly57Glu) allele frequencies were correlated with cases (panel A; p-value 0.0122) and deaths (panel B; p-value 0.0241) of COVID-19 per country. The outsider Iran was excluded to observe a significant correlation with the number of deaths. Different symbols relate to different populations associated with the same country [Purple circle: GBP; purple square: GBB.] Details about populations and their association to countries are provided in supplementary file S1. . (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

![Fig. 4. TMPRSS2 allele frequencies per population correlation with COVID-19 cases and deaths.](image)

rs12329760 (Val197Met) allele frequencies in different countries from ALFRED were correlated with the respective number of cases (panel A; p-value 0.0012) and deaths (panel B; p-value 0.0057) of COVID-19. Different symbols relate to different populations associated with the same country [Blu circle: SA004049R; blu square: SA001504K; light brown circle: SA001503J; light green circle: SA004050J; red triangle SA001505L. Light green square: SA001508O.] Details about populations and their association to countries are provided in supplementary file S1. . (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

IFI30 (Gamma-interferon-inducible lysosomal thiol reductase, GILT) has a role in protecting against the internalization of SARS-CoV viruses in the lung tissue (Chen, D. et al., 2019). Polymorphisms in this gene may be correlated with a predisposition to infection by SARS-CoV-2, as well as with a worse prognosis due to higher initial viral charge. The viral charge has, in fact, already been proven to be linked to worse clinical outcomes (Chen, D. et al., 2019).

MEFV (Pyrin) is associated to Familial Mediterranean Fever whose symptoms, fever, abdominal/chest pain, elevation of C-reactive protein, and leucocytosis, overlap in part with those of COVID-19. Deleterious variants in this gene might influence the severity of the infection. MEFV is thought to be involved in the NLRP3 dependant IL-1 beta modulation (Manukyan and Aminov, 2016). The capability of SARS-CoV viruses to activate the NLRP3 inflammasome has already been described (Chen, I.-Y. et al., 2019). There are some known pathological mutations of MEFV that cause autoinflammatory disease due to a IL1, 6, TNF- alfa INF-gamma cytokine storm (Manukyan and Aminov, 2016).

Toll like receptors activate Type-I interferon (alpha and beta) response that enhance innate antiviral responses such as the limitation of cell replication, the induction of apoptosis, and the synthesis of

![Fig. 4. TMPRSS2 allele frequencies per population correlation with COVID-19 cases and deaths.](image)

rs12329760 (Val197Met) allele frequencies in different countries from ALFRED were correlated with the respective number of cases (panel A; p-value 0.0012) and deaths (panel B; p-value 0.0057) of COVID-19. Different symbols relate to different populations associated with the same country [Blu circle: SA004049R; blu square: SA001504K; light brown circle: SA001503J; light green circle: SA004050J; red triangle SA001505L. Light green square: SA001508O.] Details about populations and their association to countries are provided in supplementary file S1. . (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

IFI30 (Gamma-interferon-inducible lysosomal thiol reductase, GILT) has a role in protecting against the internalization of SARS-CoV viruses in the lung tissue (Chen, D. et al., 2019). Polymorphisms in this gene may be correlated with a predisposition to infection by SARS-CoV-2, as well as with a worse prognosis due to higher initial viral charge. The viral charge has, in fact, already been proven to be linked to worse clinical outcomes (Chen, D. et al., 2019).

MEFV (Pyrin) is associated to Familial Mediterranean Fever whose symptoms, fever, abdominal/chest pain, elevation of C-reactive protein, and leucocytosis, overlap in part with those of COVID-19. Deleterious variants in this gene might influence the severity of the infection. MEFV is thought to be involved in the NLRP3 dependant IL-1 beta modulation (Manukyan and Aminov, 2016). The capability of SARS-CoV viruses to activate the NLRP3 inflammasome has already been described (Chen, I.-Y. et al., 2019). There are some known pathological mutations of MEFV that cause autoinflammatory disease due to a IL1, 6, TNF- alfa INF-gamma cytokine storm (Manukyan and Aminov, 2016).

Toll like receptors activate Type-I interferon (alpha and beta) response that enhance innate antiviral responses such as the limitation of cell replication, the induction of apoptosis, and the synthesis of
enzymes with direct antiviral effect (Kawasaki and Kawai, 2014). Some viruses have been described to elude the Type I interferon response by shifting the reply towards the Type II (gamma) interferon branch, which favors viral diffusion through tissues by activating a cytokine storm (McGonagle et al., 2020).

Toll-like receptor 2 interacts with TICAM2 (Stack et al., 2014) that in turn contributes to SARS-CoV pathogenesis in mice (Gralinski et al., 2017). Recent papers on TLR2 have linked the function of this gene to viral as well as to the already known bacterial targets (Carty and Bowie, 2010). TLR2 might bind SARS-CoV-2 proteins and have a key role in directing the immune response to this pathogen. Arg753Gln (rs5743708) and Pro631His (rs5743704) in TLR2 are relatively frequent in the European population, approximately 3%. A few papers confirm the role of Arg753Gln (rs5743708) in respiratory diseases (Patarcić et al., 2015) (Smelaya et al., 2016). The effects of Arg753 substitution by glutamine on TLR2 were proven in vitro. The mutation impairs agonist-induced phosphorylation hetero-dimerization with TLR6, and recruitment of myeloid differentiation primary response protein 88 (MYD88) and MyD88 adapter-like (MAL) (Xiong et al., 2012). The SNP Arg753Gln (rs5743708) is covered in the GWAS study (Group, 2020), although these SNPs were not covered by the study.

Thr199Met in the gene SLC6A20 encoding Sodium- and chloride-dependent transporter XTRP3 and Asp182Asn in the gene LZTFL1 encoding Leucine zipper transcription factor-like protein 1 are relatively frequent in the European population but not in African and East Asian populations.

Glue994Lys and Arg1140Gln are in the gene FYCO1 encoding FYVE and coiled-coil domain-containing protein 1 that participate in the transportation and maturation of autophagosomes.

3. Conclusion

COVID-19 pandemic upset the world. The fast spread of SARS-CoV-2 through over the world and some peculiar novelties of the virus, such as the high number of asymptomatic individuals, rendered it a difficult challenge from different aspects. A strong effort was put in the discovery of risk factors associated with a more serious phenotype. Besides this, different countries experienced different impacts caused by SARS-CoV-2. As well described by Yamamoto and Bauer (2020), many factors contribute to these differences, e.g. socio-behavioral habits, political management, co-existence with other viruses, genetic factors. Genetic factors can play a major role in defining the outcome of the infection. This correlation is hard to prove due to all the side-factors described above, that certainly influence the official data from the different countries. Moreover, the globalized world we live in certainly complicates the association between population variants and countries.

Fig. 5. CD27 allele frequencies per population correlation with COVID-19 cases and deaths. rs25680 (Aia59Thr) allele frequencies in different countries from ALFRED were correlated with the respective number of cases (panel A; n. s.) and deaths (panel B; p-value 0.0337) of COVID-19. Different colors relate to different countries while different symbols relate to different populations associated with the same country [Blue circle: SA004377V; blue square: SA004049R. Red circle: SA002255O; red square: SA004057Q; red triangle SA001505L.] Details about populations and their association to countries are provided in supplementary file S1. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)
Nevertheless, our data showed that a statistically significant correlation could be proved for some common variants in three genes, namely MBL2, TMPRSS2, and CD27, with the number of cases or deaths observed per country. These data suggest that the genetic background is of utmost importance in the evaluation of COVID-19 susceptibility and that the discovery of target genes could be useful in the treatment and prevention of the infection. Moreover, the evaluation of the role of these polymorphisms might be necessary for the management of very common drugs during a pandemic.

Acknowledgements

This work is dedicated to our friend and colleague Maria Malanga.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ejmg.2021.104227.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

Amberger, J., Bocchini, C.A., Scott, A.F., Hamasch, A., 2009. McKusick’s online Mendelian inheritance in man (OMIM®). Nucleic Acids Res. 37 (Suppl. 1), D979–D986.

Bernstein, K.E., Khan, Z., Gianì, J.F., Cao, D.Y., Bernstein, E.A., Shen, X.Z., 2018. Polymorphisms in the toll-like receptor 3 (TLR3) gene are associated with the natural course of hepatitis B virus infection in Caucasian population. Sci. Rep. 8 (1), 181.

Cobb, R., Dowall, S., Fotheringham, S.A., 2020. Prophylactic intranasal administration of a TLR2/6 agonist reduces upper respiratory tract viral shedding in healthy volunteers. Efficacy and safety of a TLR1/2 agonist in severe acute respiratory syndrome coronavirus 2 infection. Journal of the American Medical Association 7 (6), 1663.

Garcia, L.F., 2020. Immune response, inflammation, and the clinical spectrum of COVID-19. Front. Immunol. 11, 1441.

Guo, J., Huang, Z., Lin, L., lv, J., 2020. Coronavirus disease 2019 (covid-19) and cardiovascular disease: a viewpoint on the potential influence of angiotensin-converting enzyme inhibitors/angiotensin receptor blockers on onset and severity of severe acute respiratory syndrome coronavirus 2 infection. Journal of the American Heart Association 9 (7), e016219.

Husain, M., Jabeen, N., Raza, F., Shabbir, S., Baig, A.A., Amannaul, A., Aziz, B., 2020. Structural variations in human ACE2 may influence its binding with SARS-CoV-2 spike protein. J. Med. Virol.

Iwata-Yoshikawa, N., Okamura, T., Shimizu, Y., Hasegawa, H., Takeda, M., Hasegawa, H., Takeda, M., Nagata, N. 2019. TMPRSS2 contributes to virus spread and immunosuppression in the airways of murine models after coronavirus infection. J. Virol. 93 (6).

Kaye, T., Suzuki, Y., Kawai, T., Sakamoto, T., Ohtani, K., Eda, S., Maeda, A., Okuno, Y., Kurimu, T., Wakayama, N., 1999. Human mannann-binding lectin inhibits the infection of influenza A virus without complement. Immunology 97 (3), 385.

Kawasaki, T., Kawai, T., 2014. Toll-like receptor signaling pathways. Front. Immunol. 5, 640.

Larsen, F., Madsen, H.O., Sim, R.B., Koch, G., Garrard, P., 2004. Disease-associated mutations in human mannose-binding lectin compromise oligomerization and activity of the final protein. J. Biol. Chem. 279 (20), 17302–17311.

Mancini, O., Galiotto, A., Borgiani, P., Gianini, R., Gravina, P., Smarrazzo, A., Dauri, M., Andreoni, M., Rogliani, P., 2020. COVID-19 and genetic variants of protein involved in the SARS-CoV-2 entry into the host cells. Genes 11 (9), 1010.

Nakamura, Y., 2020. Corona (COVID-19) time musings: our involvement in COVID-19 related Neuroweaver. Advanced Science Research 22 (2), 327.

Nettles, J.G., 2015. Toll-like receptors in respiratory tract infections: systematic review, meta-analyses and quantitative synthesis. conceived in 2006. J. Med. Virol. 86 (4), 141–423.

Nicolò, G., Peroni, D., Scarpa, R., Zamboni, P., Formenti, S., 2017. Potential role of genetic susceptibility and airborne transmission. Hum. Genom. 7 (1) 104227.
Xiong, Y., Song, C., Snyder, G.A., Sundberg, E.J., Medvedev, A.E., 2012. R753Q polymorphism inhibits Toll-like receptor (TLR) 2 tyrosine phosphorylation, dimerization with TLR6, and recruitment of myeloid differentiation primary response protein 88. J. Biol. Chem. 287 (45), 38327–38337.

Yamamoto, N., Bauer, G., 2020. Apparent difference in fatalities between Central Europe and East Asia due to SARS-COV-2 and COVID-19: four hypotheses for possible explanation. Med. Hypotheses 144, 110160.

Yan, R., Zhang, Y., Li, Y., Xia, L., Guo, Y., Zhou, Q., 2020. Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. Science 367 (6485), 1444–1448.

Yang, Z., Suzuki, R., Brunquell, C.B., Sala, C.J., Nishiyama, A., 2006. NG2 glial cells provide a favorable substrate for growing axons. J. Neurosci. 26 (14), 3829–3839.

Zhang, H., Zhou, G., Zhi, L., Yang, H., Zhai, Y., Dong, X., Zhang, X., Gao, X., Zhu, Y., He, F., 2005. Association between mannose-binding lectin gene polymorphisms and susceptibility to severe acute respiratory syndrome coronavirus infection. J. Infect. Dis. 192 (8), 1355–1361.

Zhang, Q., Bastard, P., Liu, Z., Le Pen, J., Moncada-Velez, M., Chen, J., Ogishi, M., Sahli, I.K., Hodeib, S., Korol, C., 2020a. Inborn errors of type I IFN immunity in patients with life-threatening COVID-19. Science 370 (6515).

Zhang, S.-Y., Zhang, Q., Casanova, J.-L., So, H.C., 2020b. Severe COVID-19 in the young and healthy: monogenic inborn errors of immunity? Nat. Rev. Immunol. 20 (8), 455–456.

Zheng, Y., Xu, H., Yang, M., Zeng, Y., Chen, H., Liu, R., Li, Q., Zhang, N., Wang, D., 2020. Epidemiological characteristics and clinical features of 32 critical and 67 noncritical cases of COVID-19 in Chengdu. J. Clin. Virol. 104366.

Zhu, X., Wang, Y., Zhang, H., Liu, X., Chen, T., Yang, R., Shi, Y., Cao, W., Li, P., Ma, Q., 2011. Genetic variation of the human α-2-Heremans-Schmid glycoprotein (AHSG) gene associated with the risk of SARS-CoV infection. PloS One 6 (8), e23730.